Berry phases near degeneracies: Beyond the simplest case

Anupam Garg
Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208

(Received 23 January 2009; accepted 11 March 2010)

The Berry phase is reviewed with emphasis on the Berry curvature and the Chern number. The behavior of these quantities and the analytic properties of adiabatically continued wave functions in the vicinity of degeneracies are discussed. An example of a spin Hamiltonian is given in which the Chern numbers associated with the states involved in a double degeneracy are ± 2 rather than ± 1, as is usually the case. Degeneracies in the spectrum of magnetic molecular solids are discussed.

[DOI: 10.1119/1.3377135]

I. INTRODUCTION

The Berry phase entered the lexicon of physics some 25 years ago.[1] Since then, numerous physical applications and experimental confirmations of this phase have been found. There is an enormous literature on the subject, of which I cannot give a complete account because only a small part is known to me. A resource letter[2] gives a much more complete reference list up to 1996. A reprint volume[3] with commentary contains the pioneering papers including anticipations of the phase and related physics by Pancharatnam,[4] Herzberg and Longuet-Higgins,[5,6] Stone,[7] and Mead and Trulhar.[8] We may nevertheless mention some other papers in which the physical consequences of the phase are described. Among the experiments, we note the rotation of the polarization of light in a helical optical fiber,[9] neutron spin rotation in a helical magnetic field,[10] and an NMR measurement on protons.[11,12] There are also many connections with effects in solid state physics such as the quantization of the Hall conductance in a periodic potential,[13] the polarization of ferroelectrics,[14] and the anomalous velocity in semiclassical electron dynamics.[15] A more recent example is that of the half-integer shift in the quantization condition for the quantum Hall effect in graphene.[16–18]

There are several excellent expositions of the Berry phase, or geometrical phase as it is also called. Berry's original paper[1] is exceptionally lucid and strongly recommended. A good explanation has been given by Holstein.[19] Another superb explanation is given by Shankar,[20] who especially clarifies the role and effects of this phase in the Born–Oppenheimer approximation.

The purpose of this paper is to discuss a slightly more intricate example of Berry's phase than is usually encountered. Of the articles intended for a general audience of which I am aware, none go beyond the example of a spin in a slowly time-dependent magnetic field. Further, these papers focus mainly on the Berry potential **A**, whose line integral in parameter space gives the geometrical phase.[21] We shall attend more to the Berry curvature given by the generalized curl or exterior derivative of this potential. Also, the mathematical and physical structures that Berry's phase entails are very rich, and the simple spin example does not capture them fully. For example, the integral of the Berry curvature over a closed surface is guaranteed to be an integer multiple of 2π. The integer, known as the Chern number, is a topological invariant. For the simple example, it turns out to be ± 1. In so far as one learns physics more effectively through example and counterexample, anyone who encounters the general theorem about the Chern number is bound to be curious about problems where higher numbers arise.[22] I have not been able to find any such example, however.

It was this motivation that led me to the example presented here. In this example the degeneracy is twofold, but the Chern number is ± 2. In discussing it, I found myself invoking basic Berry phase concepts repeatedly, and so I have included a review of these concepts. One could equally well read the papers of Berry[1] or Holstein,[19] for example, but it is useful to have all of the results in one place in uniform notation. I also discuss related issues such as the codimensions of degeneracies and some of the subtle points about single-valuedness and analyticity of the wave functions, as students often find these confusing.

The paper is written at a level suitable for graduate students or advanced undergraduates. There is much quantum mechanics to be learned. In addition to the Berry phase notions themselves, the example entails perturbation theory for the case where a degeneracy is not lifted until the second order. The energies are needed to second order, and the wave functions are needed to first order. This case tends be discussed only in relatively advanced texts.[23,24]

The plan of the paper is as follows. The review is contained in Sec. II. Expert readers can skip this section. The example with Chern number ± 2 is given in Sec. III. In Sec. IV I briefly discuss magnetic molecular solids, a class of materials which have been the subject of much recent study, and which display degeneracies and associated Berry phase physics in abundance. The spin Hamiltonians that arise in studying these systems are similar to the example in Sec. III. In Sec. V I present a few exercises that students may enjoy. The more difficult ones could even form the basis of short study projects. Alternative treatments of the codimensions of degeneracies and of the Berry curvature for the example in Sec. III are given in the Appendixes.

II. REVIEW OF BASIC BERRY-PHASE CONCEPTS

In this section we review the key ideas behind Berry's phase using a notation close to his original one.[1] We consider a quantum system with Hamiltonian $\mathcal{H}(\mathbf{R})$, which depends parametrically on variables, $R_1, R_2, \ldots$, denoted collectively by the vector **R**. We shall write most formulas as if **R** were three dimensional, but this is done merely for convenience, and the arguments hold equally well for higher dimensional **R**.[25]

A. Adiabatic transport and the Berry phase

Let us now suppose that $\mathbf{R}$ can be varied with time, and this variation is so slow as to allow the adiabatic approximation. Let $|n(\mathbf{R})\rangle$ denote the eigenstates of $\mathcal{H}(\mathbf{R})$,

$$\mathcal{H}(\mathbf{R})|n(\mathbf{R})\rangle = E_n(\mathbf{R})|n(\mathbf{R})\rangle. \tag{1}$$

We shall refer to this as the snapshot basis because it is found by freezing $\mathbf{R}$ at a particular value, as in a snapshot. Berry showed that if $\mathbf{R}$ is taken around a closed loop C so that $\mathbf{R}(t=0)=\mathbf{R}(t=T)$, and if the initial state is $|n(\mathbf{R}(0))\rangle$, then provided the system does not pass through any degeneracies, the final state at time T is

$$e^{i\gamma_n}e^{i\phi_n}|n(\mathbf{R}(T))\rangle. \tag{2}$$

Because $\mathbf{R}(T)=\mathbf{R}(0)$, the system returns to its original state, as mandated by the adiabatic theorem, but modulo the phase $\gamma_n+\phi_n$. The part ϕ_n is the dynamical phase,

$$\phi_n = -\frac{1}{\hbar}\int_0^T E_n(\mathbf{R}(t))dt. \tag{3}$$

The other part is the Berry phase,

$$\gamma_n = i\int_0^T \langle n(\mathbf{R}(t))|\frac{d}{dt}|n(\mathbf{R}(t))\rangle dt. \tag{4}$$

The key point is that γ_n is purely geometrical, independent of how slowly the loop in $\mathbf{R}$ space is traversed. To see this, we write

$$\frac{d}{dt}|n(\mathbf{R}(t))\rangle = \sum_i \frac{\partial}{\partial R_i}|n(\mathbf{R})\rangle\frac{dR_i}{dt} = \boldsymbol{\nabla}_{\mathbf{R}}|n(\mathbf{R})\rangle\cdot\frac{d\mathbf{R}}{dt}, \tag{5}$$

where we have adopted three-dimensional vector notation in the last form. It is further convenient to put the gradient operator inside the ket and write

$$\boldsymbol{\nabla}_{\mathbf{R}}|n(\mathbf{R})\rangle \equiv |\boldsymbol{\nabla}_{\mathbf{R}}n(\mathbf{R})\rangle. \tag{6}$$

By substituting this result into Eq. (4), we obtain

$$\gamma_n = i\int_0^T \langle n(\mathbf{R})|\boldsymbol{\nabla}_{\mathbf{R}}n(\mathbf{R})\rangle\cdot\frac{d\mathbf{R}}{dt}dt. \tag{7}$$

But now we can cancel the dt's and write γ_n purely in terms of an integral in parameter space,

$$\gamma_n(C) = i\oint_C \langle n(\mathbf{R})|\boldsymbol{\nabla}_{\mathbf{R}}n(\mathbf{R})\rangle\cdot d\mathbf{R}. \tag{8}$$

We show the dependence of γ_n on the loop in the $\mathbf{R}$ space explicitly in Eq. (8).

The second point is that γ_n cannot be gauged away. If we multiply $|n(\mathbf{R})\rangle$ by $\exp(i\alpha(\mathbf{R}))$, where $\alpha(\mathbf{R})$ is single-valued, a term $i\boldsymbol{\nabla}_{\mathbf{R}}\alpha(\mathbf{R})$ is added to the integrand for γ_n, but this term integrates to zero. This point can be stated in a more familiar language if we introduce the vector-potential-like object,

$$\mathbf{A}_n(\mathbf{R}) = i\langle n(\mathbf{R})|\boldsymbol{\nabla}_{\mathbf{R}}n(\mathbf{R})\rangle. \tag{9}$$

Equation (7) then resembles the Aharonov–Bohm phase for a particle moving around a closed loop in a magnetic field described by a vector potential $\mathbf{A}_n(\mathbf{R})$,

$$\gamma_n(C) = \oint_C \mathbf{A}_n(\mathbf{R})\cdot\frac{d\mathbf{R}}{dt}dt. \tag{10}$$

Vector potentials are not gauge invariant, but the Aharonov–Bohm phase is. The quantity $\mathbf{A}_n(\mathbf{R})$ is now known as the *Berry vector potential* or the *Berry connection*.[26] The latter terminology arises because $\mathbf{A}_n(\mathbf{R})$ describes how to relate or connect the kets $|n(\mathbf{R})\rangle$ and $|n(\mathbf{R}+d\mathbf{R})\rangle$ at two nearby points in parameter space. The phase $\gamma_n(C)$ is viewed as the *anholonomy* associated with this connection.[27]

B. The Berry curvature

It is now clear how to write γ_n in a manifestly gauge invariant way. If we define the magnetic-field-like object,

$$B_n(\mathbf{R}) = \boldsymbol{\nabla}_{\mathbf{R}}\times\mathbf{A}_n(\mathbf{R}), \tag{11}$$

which is known as the *Berry curvature*, then by Stokes' theorem we may write γ_n as a surface integral,

$$\gamma_n(C) = \int\int_S B_n(\mathbf{R})\cdot d\mathbf{S}, \tag{12}$$

where S is any surface spanning C. Since $\mathbf{B}_n$ is gauge invariant, so is $\gamma_n(C)$. Why $\mathbf{B}_n$ is called a curvature is discussed in the following. Also, note that the dimensions of $\mathbf{B}_n$ are $[R]^{-2}$, where $[R]$ is the dimension of $\mathbf{R}$.

It thus follows that the Berry curvature is an intrinsic property of the way in which the entire ray in Hilbert space associated with $|n(\mathbf{R})\rangle$ twists and turns as $\mathbf{R}$ is varied. An extremely interesting question now arises. If we think of $\mathbf{B}_n$ as a magnetic field, what are the sources of this field? Equation (11) shows that $\boldsymbol{\nabla}_{\mathbf{R}}\cdot\mathbf{B}_n(\mathbf{R})=0$ just as for a true magnetic field, so the sources must have physical significance. To find them, we need to write $\mathbf{B}_n$ in two other forms. We first note that we may also write (employing an obvious abbreviated notation)

$$\mathbf{A}_n = -\,\mathrm{Im}\langle n|\boldsymbol{\nabla}n\rangle. \tag{13}$$

The imaginary part restriction can be understood as follows. Since $|n(\mathbf{R})\rangle$ is normalized for all $\mathbf{R}$, applying the gradient to the relation $1=\langle n|n\rangle$ gives

$$0 = \boldsymbol{\nabla}\langle n|n\rangle = \langle\boldsymbol{\nabla}n|n\rangle + \langle n|\boldsymbol{\nabla}n\rangle. \tag{14}$$

Here, $\langle\boldsymbol{\nabla}n|$ stands for $\boldsymbol{\nabla}\langle n|$ in analogy to Eq. (6). More explicitly, let $|m\rangle$ be a complete set of *fixed* ($\mathbf{R}$-independent) basis states. Then,

$$|n(\mathbf{R})\rangle = \sum_m c_m(\mathbf{R})|m\rangle, \quad c_m(\mathbf{R}) = \langle m|n(\mathbf{R})\rangle. \tag{15}$$

Therefore,

$$|\boldsymbol{\nabla}n\rangle = \sum_m \boldsymbol{\nabla}c_m(\mathbf{R})|m\rangle. \tag{16}$$

Similarly,

$$\langle\boldsymbol{\nabla}n| = \sum_m \boldsymbol{\nabla}c_m^*(\mathbf{R})\langle m|. \tag{17}$$

The two terms on the right-hand side of Eq. (14) are seen to be complex conjugates, from which it follows that each of them is pure imaginary, and Eq. (13) follows in turn. Equations (9) and (13) then show that $\gamma_n(C)$ is real, as it should

be. It follows that the Berry potential and curvature are also real.

The curl of Eq. (13) yields

$$\mathbf{B}_n = -\,\mathrm{Im}\langle\boldsymbol{\nabla} n| \times |\boldsymbol{\nabla} n\rangle \tag{18}$$

since $\boldsymbol{\nabla}\times|\boldsymbol{\nabla} n\rangle=0$. This is the first form for $\mathbf{B}_n$. It is useful for calculations, and it also shows that $\mathbf{B}_n$ is explicitly gauge invariant. For if we change $|n\rangle$ to $e^{i\alpha}|n\rangle$, the additional terms involving $\boldsymbol{\nabla}\alpha$ vanish due to Eq. (14). Lastly, it also shows that $\boldsymbol{\nabla}\cdot\mathbf{B}_n=0$.

To obtain the second form for $\mathbf{B}_n$, let us insert a resolution of unity as a sum over the complete set of snapshot basis states in Eq. (18). This gives

$$\mathbf{B}_n = -\,\mathrm{Im}\sum_{n'\neq n}\langle\boldsymbol{\nabla} n|n'\rangle\times\langle n'|\boldsymbol{\nabla} n\rangle. \tag{19}$$

The term with $n'=n$ has been omitted from the sum, as it has no imaginary part. To find $\langle n'|\boldsymbol{\nabla} n\rangle$, let us take the gradient of Eq. (1) and project onto $|n'(\mathbf{R})\rangle$. This yields

$$\langle n'|\,\boldsymbol{\nabla}\,\mathcal{H}|n\rangle + \langle n'|\mathcal{H}|\boldsymbol{\nabla} n\rangle = \langle n'|\,\boldsymbol{\nabla}\,E_n|n\rangle + \langle n'|E_n|\boldsymbol{\nabla} n\rangle. \tag{20}$$

We invoke $\langle n'|\mathcal{H}=E_{n'}\langle n'|$ and simplify and obtain

$$\langle n'|\boldsymbol{\nabla} n\rangle = \frac{\langle n'|\,\boldsymbol{\nabla}\,\mathcal{H}|n\rangle}{E_{n'}-E_n} \quad (n'\neq n). \tag{21}$$

Feeding this result into Eq. (19) leads to the second form

$$\mathbf{B}_n = -\,\mathrm{Im}\sum_{n'\neq n}\frac{\langle n|\,\boldsymbol{\nabla}\,\mathcal{H}|n'\rangle\times\langle n'|\,\boldsymbol{\nabla}\,\mathcal{H}|n\rangle}{(E_{n'}-E_n)^2}. \tag{22}$$

This form shows that $\mathbf{B}_n$ is singular at points of degeneracy in parameter space where the energy denominators vanish since the off-diagonal matrix elements of $\boldsymbol{\nabla}\mathcal{H}$ will generally not vanish at the same points. These points are like magnetic monopole sources of $\mathbf{B}$. They are not the only sources because $\boldsymbol{\nabla}\times\mathbf{B}\neq 0$, so there must also be currents flowing through the parameter space. They are the most interesting, however, as there are strong topological constraints on them.

C. Codimensions of degeneracies

The first problem that therefore arises is to find an $\mathbf{R}$ where a degeneracy exists. It is not immediately obvious that such points are rare. At first sight, the condition $E_1(\mathbf{R})=E_2(\mathbf{R})$ would seem to require the variation of just one parameter, and so if $\mathbf{R}$ lives in a three-dimensional space, degeneracies would seem to lie on two-dimensional surfaces. This conclusion is incorrect. A classic theorem due to von Neumann and Wigner[28] and Teller[29] states that to find a double degeneracy other than one allowed by a symmetry of the Hamiltonian, we must tune two parameters if the Hamiltonian is real symmetric, and three if it is complex Hermitian. We give Teller's argument for the latter case.[30,31] Let us seek a degeneracy between two states $|1(\mathbf{R})\rangle$ and $|2(\mathbf{R})\rangle$. Suppose that the states $|n(\mathbf{R})\rangle$, $n\geqslant 3$, orthogonal to the first two are already known. Let $|\alpha(\mathbf{R})\rangle$ and $|\beta(\mathbf{R})\rangle$ be two fixed states orthogonal to each other and to $|3(\mathbf{R})\rangle, |4(\mathbf{R})\rangle, \ldots$. The problem of finding the energy eigenstates $|1(\mathbf{R})\rangle$ and $|2(\mathbf{R})\rangle$ then reduces to diagonalization of the 2×2 matrix

$$\mathcal{H} = \begin{pmatrix} H_{\alpha\alpha}(\mathbf{R}) & H_{\alpha\beta}(\mathbf{R}) \\ H_{\beta\alpha}(\mathbf{R}) & H_{\beta\beta}(\mathbf{R}) \end{pmatrix}. \tag{23}$$

Naturally, $H_{\beta\alpha}(\mathbf{R})=H^*_{\alpha\beta}(\mathbf{R})$. For the eigenvalues of this matrix to be equal, it must be similar to an identity matrix (times a constant). In other words, we must satisfy the three conditions

$$\begin{aligned} &H_{\alpha\alpha}(\mathbf{R}) = H_{\beta\beta}(\mathbf{R}),\\ &\mathrm{Re}\,H_{\alpha\beta}(\mathbf{R}) = 0,\\ &\mathrm{Im}\,H_{\beta\alpha}(\mathbf{R}) = 0, \end{aligned} \tag{24}$$

which requires, in general, that at least three tunable parameters be available. If $H_{\alpha\beta}(\mathbf{R})$ is real, at least two parameters are needed. More formally, the *codimensions* of a degeneracy are 3 and 2 for the two cases, respectively. In particular, for three-dimensional $\mathbf{R}$, and a complex Hermitian Hamiltonian, degeneracies will occur only at isolated points in the $\mathbf{R}$-space.

The previous argument implicitly assumes that there is no symmetry which ensures that $H_{\alpha\beta}(\mathbf{R})$ vanishes for general $\mathbf{R}$. It does not, however, preclude the existence of a symmetry at the degeneracy point itself.

D. The diabolo

The simplest and most generic degeneracy involves just two states. It follows from Eq. (22) that near the degeneracy point, we may ignore all other states and truncate the Hamiltonian to a 2×2 matrix. [Or, we can invoke the argument used to arrive at Eq. (23).] Without any loss of significance, we can shift the origin of $\mathbf{R}$ to the degeneracy, shift and scale the units of energy, and rotate and scale axes in $\mathbf{R}$ so that the Hamiltonian reads

$$\mathcal{H} = -\frac{1}{2}\begin{pmatrix} Z & X-iY \\ X+iY & -Z \end{pmatrix}. \tag{25}$$

Here, X, Y, and Z are the components of $\mathbf{R}$. This form is the most general because the off-diagonal elements must be complex conjugates and is independent of the diagonal elements, which can be taken to add to zero by a shift in the zero of energy. Equation (25) is the Hamiltonian of a spin $-1/2$ in a magnetic field. If the energy levels are denoted by $E_\pm=\pm|\mathbf{R}|/2$, then it is easy to show that

$$\mathbf{B}_+ = -\,\mathbf{B}_- = -\frac{\mathbf{R}}{2R^3}. \tag{26}$$

(The derivation involves simple Pauli matrix algebra and is given by Berry.[1])

Degeneracies of the type just discussed have been termed *diabolical points* by Berry and Wilkinson[32] because the energy surfaces as a function of X and Z form a double cone, which reminded them of a yo-yo-like Italian toy called the *diabolo*. At one time, the term *conical intersections* was used (see the title of Ref. 8, for example), but it has fallen out of favor.[33] We shall refer to all double degeneracies as diabolical even when the energy surfaces are of different form.

The field (26) is that of a monopole at $\mathbf{R}=0$. Away from the monopole, $\boldsymbol{\nabla}\cdot\mathbf{B}_+=0$. To find the strength of the monopole, we find the flux over any closed surface surrounding it. If we choose a sphere (simplest), we obtain

$$Q_+ = -\frac{1}{2\pi}\int\int \frac{-\mathbf{R}}{2R^3}\cdot d\mathbf{S} = 1. \tag{27}$$

The prefactor of $-1/2\pi$ is conventional.

E. The Chern number

The result (27) illustrates one of the key properties of the Berry phase. It is a remarkable fact that the monopole strength,

$$Q_n = -\frac{1}{2\pi}\oiint B_n(\mathbf{R})\cdot d\mathbf{S}, \tag{28}$$

must be an integer (known as the Chern number) for any n (that is, for any of the snapshot states) and for any surface. Secondly, for the states involved in any degeneracy,

$$\sum_n Q_n = 0. \tag{29}$$

These facts clearly hold for the simple example we have discussed. Another example (also given by Berry) for which they are again easily verified is

$$\mathcal{H} = -\mu\mathbf{J}\cdot\mathbf{H}, \tag{30}$$

where μ is a magnetic moment, $\mathbf{J}$ is a (dimensionless) spin-j angular momentum and $\mathbf{H}$ is a (true) magnetic field. In this case,

$$B_m(\mathbf{H}) = -m\frac{\mathbf{H}}{H^3}, \tag{31}$$

where $m=-j,-j+1,\ldots,j$. We now get $Q_m=2m$, which is an integer as promised.

It should be noted that the Chern number is dimensionless, irrespective of the dimensions of $\mathbf{R}$. In the example (30), $\mathbf{R}$ is a real magnetic field, and the Berry magnetic field (or curvature) has dimensions of the inverse square of the real magnetic field. The flux of the former through a two-dimensional surface in the space of the latter is therefore dimensionless. In many examples and applications of the Berry phase, the parameter is a magnetic field, and keeping these points in mind helps distinguish the Berry magnetic field from the true one.

We now discuss why the Chern number must be an integer. The argument is basically the same as that given by Dirac in his proof that the existence of a (true) magnetic monopole implies the quantization of electric charge.[34] The first point is that the vector potential describing the monopole must be singular at at least one point on any surface surrounding the monopole. Let us consider an infinitesimal disk shaped patch on the surface and suppose that $\mathbf{A}$ is nonsingular on it. The flux through this patch is given by the line integral of $\mathbf{A}$ over C, the loop bounding the patch. We now make C bigger and assume that no singularities of $\mathbf{A}$ are encountered (see Fig. 1). The flux through the loop will grow. If we keep making the patch bigger and bigger until it essentially becomes the entire surface, C will become an infinitesimal loop on the opposite side of the surface from where we started. Since the flux through the patch is now essentially equal to the total flux emanating from the monopole, and the length of the loop is very small, $\mathbf{A}$ must be very large in magnitude. As the loop length shrinks to zero, $\mathbf{A}$ must diverge. It is, of course, possible for $\mathbf{A}$ to be singular in other ways, and the argument we have given concentrates the singularity at one point. By joining together the singular points on every surface surrounding the monopole, we obtain the famous Dirac string. This singularity can be moved around by changing the gauge, but it cannot be completely eliminated.

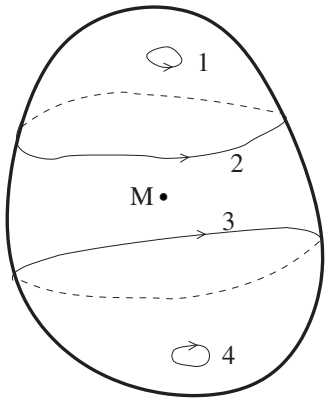

Fig. 1. Why the Chern number is an integer. An initially infitesimal loop 1 is expanded to cover more and more of the surface (loops 2 and 3) surrounding the degeneracy (M) until it covers essentially the entire surface (4). The Berry phase, given by the flux through the loop grows from 0 to an integer multiple of 2π because in the final configuration (4), adiabatic transport of the wave function around the loop must not lead to any change in the wave function.

Let us now consider the loop C when it has been shrunk to an infinitesimal one around the string. Then, by the argument just given,

$$\gamma_n(C) = \oiint_S \mathbf{B}_n\cdot d\mathbf{S} = -2\pi Q_n. \tag{32}$$

But the string is a fiction; it is not a physical object. And it is also obvious that adiabatic transport around an infinitesimal loop cannot really change the state. Therefore, it must be that the phase is unobservable (even by interferometric means), and we must have

$$e^{i\gamma_n(C)} = 1, \tag{33}$$

which implies that

$$Q_n = \text{integer}. \tag{34}$$

F. Analyticity of adiabatic wave functions

An issue closely related to the singularity of $\mathbf{A}_n$ is that it is not possible to write a single expression for the snapshot basis $|n(\mathbf{R})\rangle$ in a way that is nonsingular everywhere in $\mathbf{R}$. Take, for example, the Hamiltonian (25). If we introduce spherical polar coordinates in $\mathbf{R}$-space ($Z=R\cos\theta$, $X=R\sin\theta\cos\varphi$, $Y=R\sin\theta\sin\varphi$), we find that one choice for the state $|-,\mathbf{R}\rangle$ is

$$|-,\mathbf{R}\rangle = \begin{pmatrix} \cos\frac{1}{2}\theta \\ e^{i\varphi}\sin\frac{1}{2}\theta \end{pmatrix}. \tag{35}$$

The term $e^{i\varphi}\sin\frac{1}{2}\theta$ is singular at the south pole, and if we calculate $\mathbf{A}_-$, we will discover a Dirac string there. We can multiply this state by a globally analytic phase factor $e^{i\alpha(\mathbf{R})}$, but if we are to eliminate the singularity at the south pole, we must have $\alpha(\mathbf{R})=-\varphi(\mathbf{R})+\alpha'(\mathbf{R})$, where $\alpha'(\mathbf{R})$ is nonsingu-

lar, but now the other element of the column vector is singular at the north pole.

We note in passing that Eq. (35) gives the spin-coherent state for a spin-$\frac{1}{2}$ particle with maximal spin projection along the direction (θ,φ). Our argument shows that a single analytic expression cannot cover the entire sphere. The modern approach is to divide the sphere into two patches, one surrounding each pole, and extending past the equator into the other hemisphere up to some latitude short of the other pole. We can then define states analytic in each patch, and related by a gauge transformation (or transition function) in the overlap of the two patches. Requiring the gauge transformation to be analytic is another way of seeing that the Chern number must be quantized.

The need for more than one coordinate patch also shows up in an older argument of Herzberg and Longuet-Higgins[5,6] concerning real symmetric Hamiltonians. They proved that if the adiabatic wave function of a state reverses sign as a closed contour is traversed, that contour necessarily contains a degeneracy. The conditions of this theorem are met in example (25) if we set $Y=0$. Then under a circuit enclosing the origin in the XZ plane, the Berry phase is $\pm\pi$ because the flux of $\mathbf{B}_+$ or $\mathbf{B}_-$ through a hemispherical surface (or one with the topology of a hemisphere) is half the flux through a closed surface by symmetry, and thus equal to $\pm\pi$. Since $e^{\pm i\pi}=-1$, the theorem is verified. To see this explicitly, we can use the state (35) and see how it changes as we go in a circle in the x-z plane. Suppose we start along the $+z$ axis, so the initial state has the bra vector (1 0). We go clockwise, so that x starts out becoming positive. That is, θ increases and $\varphi=0$, and the state is $(\cos\frac{1}{2}\theta \sin\frac{1}{2}\theta)$. As we approach the $-z$ axis, $\theta\to\pi$, so the state approaches (0 1). As the $-z$ axis is crossed, x becomes negative, φ jumps to π, so the state must be taken in the form $(-\cos\frac{1}{2}\theta - e^{-i\varphi}\sin\frac{1}{2}\theta) = (-\cos\frac{1}{2}\theta \sin\frac{1}{2}\theta)$ to maintain continuity at $\theta=\pi$. As we keep going around the circle, θ decreases, until it approaches 0 as we return to the $+z$ axis. The state however returns to $(-1\ 0)$, showing that the sign has reversed.

We note that the sign need not change for degeneracies with higher Chern numbers, but a globally analytic wave function is still not possible.

G. Why "curvature"?

Next, we explain why $\mathbf{B}_n$ is called a curvature by making an analogy with curved surfaces. Let us consider an arbitrary point on an ordinary surface embedded in three-dimensional Euclidean space. Let us choose a coordinate system with its origin at this point, and with the outward normal $\hat{\mathbf{n}}$ aligned with the z axis. Further, let us align the x and y axes with the two directions of principal curvature. Then, near the origin, the equation of the surface is given by

$$z = -\tfrac{1}{2}(\kappa_1 x^2 + \kappa_2 y^2), \tag{36}$$

where κ_1 and κ_2 are the principal curvatures. Their product gives the Gaussian curvature, K, of the surface at the point in question,

$$K = \kappa_1\kappa_2. \tag{37}$$

But, we can also derive K as follows. For very small x and y, the normal to the surface is given by

$$\hat{\mathbf{n}} = \hat{\mathbf{z}} + \kappa_1 x\hat{\mathbf{x}} + \kappa_2 y\hat{\mathbf{y}}. \tag{38}$$

Therefore,

$$\frac{d\hat{\mathbf{n}}}{dx} = \kappa_1\hat{\mathbf{x}}, \quad \frac{d\hat{\mathbf{n}}}{dy} = \kappa_2\hat{\mathbf{y}} \tag{39}$$

and

$$\frac{d\hat{\mathbf{n}}}{dx}\times\frac{d\hat{\mathbf{n}}}{dy} = K\hat{\mathbf{z}}. \tag{40}$$

The similarity between this equation and Eq. (18) is why $\mathbf{B}_n$ is called a curvature.[35] Further, the quantization of the Chern number is analogous to the Gauss–Bonnet theorem, according to which the integral of the Gaussian curvature over any closed surface equals 2π times the Euler characteristic of that surface.[36] We recall that the latter depends only on the topology of the surface: it is 2 for a sphere, 0 for a torus, -2 for a sphere with two handles, and so on. The Chern number is similarly topological.

III. THE EXAMPLE

The example Hamiltonian we study is

$$\bar{\mathcal{H}} = k(1 - J_z^2) - g\mu_B\mathbf{J}\cdot\mathbf{H}. \tag{41}$$

It describes a spin-1 degree of freedom, such as a magnetic ion in a solid. $\mathbf{J}=(J_x,J_y,J_z)$ are dimensionless spin-1 operators, g is a g-factor, μ_B is the Bohr magneton, and $\mathbf{H}$ is an external magnetic field. We imagine that because of the solid environment, different spin orientations are not equal in energy and k is a constant that describes this anisotropy. We take $k>0$, which makes the anisotropy of the easy axis or Ising type. Our goal is to find how the eigenstates of $\bar{\mathcal{H}}$ change as $\mathbf{H}$ is varied and thus find the Berry curvature $\mathbf{B}(\mathbf{H})$. The role of the parameter $\mathbf{R}$ in Sec. II is now played by $\mathbf{H}$, which is also three dimensional.

To avoid clutter in the formulas, we measure energy in units of k, define $\mathcal{H}=\bar{\mathcal{H}}/k$, and the scaled field

$$\mathbf{R} = (g\mu_B/k)\mathbf{H} = (X,Y,Z). \tag{42}$$

This notation makes it easy to apply the formulas of Sec. II.

Because $\bar{\mathcal{H}}$ is rotationally symmetric about the z axis, it suffices to solve the eigenvalue problem for $\mathbf{R}$ in the x-z plane, that is, set $Y=0$. In the standard J_z basis, the Hamiltonian matrix is then

$$\mathcal{H} = \begin{pmatrix} -Z & -\frac{X}{\sqrt{2}} & 0 \\ -\frac{X}{\sqrt{2}} & 1 & -\frac{X}{\sqrt{2}} \\ 0 & -\frac{X}{\sqrt{2}} & Z \end{pmatrix}. \tag{43}$$

There are three diabolical points in the magnetic field space, all on the z axis. When $Z=1$, the states $|0\rangle$ and $|-1\rangle$ are degenerate, where $|m\rangle$ denotes the eigenstate of J_z with eigenvalue m. Likewise, when $Z=-1$, the states $|0\rangle$ and $|1\rangle$ are degenerate. Finally, when $Z=0$, the states $|\pm 1\rangle$ are degener-

ate. The first two diabolical points are of the simplest type, with $Q=\pm 1$. It is the third we wish to focus on, because it has $Q=\pm 2$.

Our goal, therefore, is to find the states and energies by perturbation theory when both X and Z are small. As we shall see in the following, we need the states up to first order and the energies up to second order in the perturbation. That we have two small parameters in which we may expand would appear to be an advantage, but the difficulty is that the degeneracy of the states $|\pm 1\rangle$ is broken only in second order in X, and it does not pay to assume $X \ll Z$ or $Z \ll X$ as we will eventually need to consider all orientations of $\mathbf{R}$ around $\mathbf{R}=0$.

To first order in X and Z, the states are (we denote the perturbative state developing from $|m\rangle$ by $|m^*\rangle$)

$$|1^*\rangle = |1\rangle + \frac{X}{\sqrt{2}}|0\rangle,$$

$$|0^*\rangle = |0\rangle - \frac{X}{\sqrt{2}}(|1\rangle + |-1\rangle), \tag{44}$$

$$|-1^*\rangle = |-1\rangle + \frac{X}{\sqrt{2}}|0\rangle,$$

and the energies are

$$E_0 = 1, \quad E_{\pm 1} = \mp Z. \tag{45}$$

Because Z is assumed small, however, the degeneracy of $|\pm 1\rangle$ is essentially not resolved, and the next order of perturbation theory will mix these states strongly. Thus, the previous results for $|\pm 1^*\rangle$ and $E_{\pm 1}$ are not of much use. The correct procedure is to diagonalize the second-order secular matrix added to the first-order one.[23] We reprise the main formula for a general situation. Let n, n', n'', etc., label a group of states that remain degenerate or nearly so to first order. Let the perturbation be denoted by V. Then up to second order, the secular matrix is given by

$$V^{(2)}_{nn'} = V_{nn'} + \sum_{m \neq n,n',\ldots} \frac{V_{nm}V_{mn'}}{E_n - E_m}. \tag{46}$$

The sum excludes not only the states n and n', but all states in the (nearly) degenerate group, so the energy denominator, which is formed from the unperturbed energies, never becomes small or zero. In our case the secular matrix is easily found to be

$$V^{(2)} = -\tfrac{1}{2}X^2 1 + \begin{pmatrix} -Z & -\frac{1}{2}X^2 \\ -\frac{1}{2}X^2 & Z \end{pmatrix}. \tag{47}$$

The eigenvalues of $V^{(2)}$ give the energies of the states to second order. We have

$$E_\pm = -\tfrac{1}{2}X^2 \pm D, \tag{48}$$

where the state labels have been changed to $+$ and $-$, and where

$$D = \left(Z^2 + \tfrac{1}{4}X^4\right)^{1/2}. \tag{49}$$

The states themselves are the eigenvectors of $V^{(2)}$. Writing

$$Z = D\cos\zeta, \quad \tfrac{1}{2}X^2 = D\sin\zeta, \tag{50}$$

we have

$$|-\rangle = \cos\tfrac{1}{2}\zeta|1^*\rangle + \sin\tfrac{1}{2}\zeta|-1^*\rangle, \tag{51}$$

$$|+\rangle = \sin\tfrac{1}{2}\zeta|1^*\rangle - \cos\tfrac{1}{2}\zeta|-1^*\rangle. \tag{52}$$

The key point is that on the right-hand side we must use the first-order states $|\pm 1^*\rangle$ from Eq. (45), and not the bare states $|\pm 1\rangle$. One way to see this is that otherwise $|\pm\rangle$ would not be orthogonal to $|0^*\rangle$. (A more systematic way is given in the suggested exercises.)

Combining Eqs. (45) and (52), we obtain the states in the original J_z basis. To save writing, we introduce the notation

$$c = \cos\tfrac{1}{2}\zeta, \quad s = \sin\tfrac{1}{2}\zeta. \tag{53}$$

Then,

$$|-\rangle = c|1\rangle + \frac{X}{\sqrt{2}}(c+s)|0\rangle + s|-1\rangle, \tag{54}$$

$$|0^*\rangle = -\frac{X}{\sqrt{2}}|1\rangle + |0\rangle - \frac{X}{\sqrt{2}}|-1\rangle, \tag{55}$$

$$|+\rangle = s|1\rangle - \frac{X}{\sqrt{2}}(c-s)|0\rangle - c|-1\rangle. \tag{56}$$

For completeness, we also give E_0 to second order

$$E_0 = 1 + X^2. \tag{57}$$

We now have all the ingredients needed to calculate the Berry curvatures $\mathbf{B}_\pm$ and $\mathbf{B}_0$. We calculate $\mathbf{B}_+$ using Eq. (18) below and using Eq. (22) in Appendix B.

To use Eq. (18) we need to find $|\boldsymbol{\nabla}+\rangle$, for which we need $|+\rangle$ at points outside the x-z plane. To this end, let us introduce the azimuthal angle φ in the x-y plane such that $\varphi=0$ on the x axis, and jumps from $+\pi$ to $-\pi$ as we cross the $-x$ axis in the anticlockwise sense. Then, since

$$e^{-iJ_z\varphi}J_x e^{iJ_z\varphi} = \cos\varphi J_x + \sin\varphi J_y, \tag{58}$$

it follows that

$$\mathcal{H}(X,Y,Z) = e^{-iJ_z\varphi}\mathcal{H}(R_\perp,0,Z)e^{iJ_z\varphi}, \tag{59}$$

where $R_\perp = (X^2+Y^2)^{1/2}$. The energy eigenstates for $Y\neq 0$ can thus be obtained by acting on those for $Y=0$ with the operator $e^{-iJ_z\varphi}$. The energies are, of course, unchanged. In this way we get

$$|+\rangle = se^{-i\varphi}|1\rangle - \frac{R_\perp}{\sqrt{2}}(c-s)|0\rangle - ce^{i\varphi}|-1\rangle. \tag{60}$$

Taking the gradient now gives

$$\langle 1|\boldsymbol{\nabla}+\rangle = \left(-is\,\boldsymbol{\nabla}\varphi + \tfrac{1}{2}c\,\boldsymbol{\nabla}\zeta\right)e^{-i\varphi}, \tag{61}$$

$$\langle 0|\boldsymbol{\nabla}+\rangle = -\frac{1}{\sqrt{2}}(c-s)\,\boldsymbol{\nabla}R_\perp + \frac{R_\perp}{2\sqrt{2}}(c+s)\,\boldsymbol{\nabla}\zeta, \tag{62}$$

$$\langle -1|\boldsymbol{\nabla}+\rangle = -\left(ic\,\boldsymbol{\nabla}\varphi - \tfrac{1}{2}s\,\boldsymbol{\nabla}\zeta\right)e^{i\varphi}. \tag{63}$$

The contributions to $\mathrm{Im}\langle\boldsymbol{\nabla}+|\times|\boldsymbol{\nabla}+\rangle$ from the various $|m\rangle$ states may now be calculated and are

$$m=1: \quad cs\,\boldsymbol{\nabla}\varphi\times\boldsymbol{\nabla}\zeta, \tag{64}$$

$$m=0: \quad 0, \tag{65}$$

$$m=-1: \quad cs\,\boldsymbol{\nabla}\varphi\times\boldsymbol{\nabla}\zeta. \tag{66}$$

Thus,

$$\mathbf{B}_+=-2cs\,\boldsymbol{\nabla}\varphi\times\boldsymbol{\nabla}\zeta. \tag{67}$$

It is enough to evaluate this in the x-z plane. Near this plane, $\varphi\approx Y/X$, so on it,

$$\boldsymbol{\nabla}\varphi=\frac{1}{X}\hat{\mathbf{y}}. \tag{68}$$

Further, since $2cs=\sin\zeta$,

$$2cs\,\boldsymbol{\nabla}\zeta=-\boldsymbol{\nabla}(\cos\zeta)=-\boldsymbol{\nabla}\frac{Z}{D}=\frac{X^4\hat{\mathbf{z}}-2ZX^3\hat{\mathbf{x}}}{4D^3}. \tag{69}$$

Hence,

$$\mathbf{B}_+=\frac{X^2}{4D^3}(X\hat{\mathbf{x}}+2Z\hat{\mathbf{z}}). \tag{70}$$

We stress that this formula is only valid sufficiently close to the degeneracy at $\mathbf{R}=0$. We leave it to the reader to verify that $\mathbf{B}_0=0$ and $\mathbf{B}_-=-\mathbf{B}_+$.

We can now find the Chern number Q_+ for $|+\rangle$, that is, the flux of $\mathbf{B}_+$ through a surface surrounding the origin. We take the surface to be a cylinder with axis along $\hat{\mathbf{z}}$, of radius R, and extending from $Z=-L$ to $Z=L$. Denoting the distance from the z axis by r, we find the contribution to Q_+ through the top end of the cylinder to be

$$Q_{+,\mathrm{top}}=-\frac{1}{2\pi}\int_0^R\frac{r^2L}{2\left(L^2+\frac{1}{4}r^4\right)^{3/2}}2\pi r dr. \tag{71}$$

If we change the variable of integration to $r^4/4$, the integral can be easily done and yields

$$Q_{+,\mathrm{top}}=-1+\frac{L}{\left(L^2+\frac{1}{4}R^4\right)^{1/2}}. \tag{72}$$

The contribution from the bottom of the cylinder is the same by symmetry. From the sides of the cylinder, we get

$$Q_{+,\mathrm{sides}}=-\frac{1}{2\pi}\int_{-L}^{L}\frac{R^3}{4\left(Z^2+\frac{1}{4}R^4\right)^{3/2}}2\pi R dZ. \tag{73}$$

This integral is also elementary and yields

$$Q_{+,\mathrm{sides}}=-\left.\frac{Z}{\left(Z^2+\frac{1}{4}R^4\right)^{1/2}}\right|_{-L}^{L}=-2\frac{L}{\left(L^2+\frac{1}{4}R^4\right)^{1/2}}. \tag{74}$$

Adding up all the pieces, we get, as advertised,

$$Q_+=-2. \tag{75}$$

Again, we leave it to the reader to verify that $Q_0=0$ and $Q_-=2$.

We now discuss the analyticity of the state $|+\rangle$ in the $\mathbf{R}$-space. We reproduce the two wave functions (56) and (60) for this state for ready reference:

$$|+_1\rangle=s|1\rangle-\frac{X}{\sqrt{2}}(c-s)|0\rangle-c|-1\rangle, \tag{76}$$

$$|+_2\rangle=se^{-i\varphi}|1\rangle-\frac{R_\perp}{\sqrt{2}}(c-s)|0\rangle-ce^{i\varphi}|-1\rangle. \tag{77}$$

The form $|+_1\rangle$ only applies when $\mathbf{R}$ is in the x-z plane, while $|+_2\rangle$ applies for all $\mathbf{R}$. Let us now consider two points in the x-z plane at the same value of $Z\neq0$, but opposite values of X, and call them A and B. Let $X>0$ at A, so that at A, $|+_2\rangle=|+_1\rangle$. If we employ $|+_2\rangle$ and go from A to B via a 180° rotation about $\hat{\mathbf{z}}$, we get $-|+_1\rangle$, with an extra minus sign. To avoid this sign, we can take the wave function as

$$|+_3\rangle=s|1\rangle-e^{i\varphi}\frac{R_\perp}{\sqrt{2}}(c-s)|0\rangle-ce^{2i\varphi}|-1\rangle, \tag{78}$$

which is identical to $|+_1\rangle$ in the x-z plane and applies for all $\mathbf{R}$.

Both $|+_2\rangle$ and $|+_3\rangle$ are singular at both poles, illustrating the general comments in Sec. II F. Readers are invited to find the corresponding Berry potentials and see how these singularities show up there.

We can also try and execute a Herzberg and Longuet-Higggins circuit in the x-z plane because the Hamiltonian is then always real. The form $|+_1\rangle$ is the one to use. As in Sec. II F, we go around in a clockwise circle starting on the z axis. On this axis, $|+_1\rangle=-|-1\rangle$. As we keep increasing X, the state changes until at A it has the form

$$|+_1(A)\rangle=a|1\rangle+b|0\rangle+c|-1\rangle, \tag{79}$$

with some coefficients a, b, and c. When we reach the $-z$ axis, $\zeta=\pi$, so the state is $|1\rangle$. There is no jump in the form of $|+_1\rangle$ as we cross this axis. Since ζ has the same value at A and B, at B $|+_1\rangle$ has the form

$$|+_1(B)\rangle=a|1\rangle-b|0\rangle+c|-1\rangle, \tag{80}$$

and it returns to $-|-1\rangle$ as we return to the z axis. In other words, there is no sign reversal, consistent with $Q_+=-2$. There is no contradiction with the general theorem because sign reversal is only a sufficient condition for there to be a degeneracy inside the circuit.

IV. MAGNETIC MOLECULAR SOLIDS

Spin Hamiltonians of the same general flavor as discussed in Sec. III arise in many magnetic molecular solids. These systems have been intensively studied in the past decade or so.[37] Broadly speaking, these are molecular solids of organic molecules in which there is a core of magnetic ions (Mn^{3+}, Fe^{2+}, etc.), giving a net spin (and magnetic moment) to the entire molecule. The magnetic interactions between the molecules are weak and may be neglected so that one is justified in studying the one-body or single molecule problem. The solid state environment is anisotropic, giving rise to Hamiltonians such as Eqs. (41) and (81).

It is not the purpose of this paper to discuss the many fascinating physical phenomena that these systems display. They are mentioned here because diabolical points have been experimentally observed in several such solids. The best studied material is based on the spin 10 molecule $[(\mathrm{tacn})_6\mathrm{Fe}_8\mathrm{O}_2(\mathrm{OH})_{12}]^{8+}$ (abbreviated to Fe_8) in which a whole array of such points is seen.[38,39] An approximate model Hamiltonian for this system is

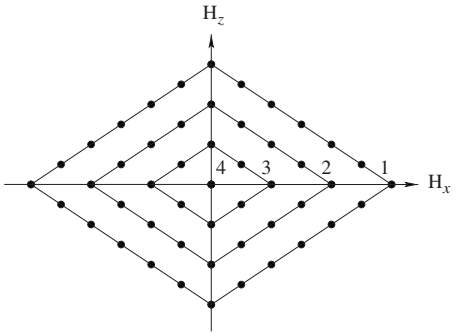

Fig. 2. Diabolical points for the Hamiltonian (81) for $J=7/2$. All points on the outermost rhombus are singly diabolical, those on the next one are doubly diabolical, and so on. The origin is quadruply diabolical.

$$\mathcal{H} = -k_2 J_z^2 + (k_1 - k_2) J_x^2 - g\mu_B \mathbf{J}\cdot\mathbf{H}, \tag{81}$$

in which $k_1 > k_2 > 0$ are measured anisotropy energies, and g and μ_B are the g-factor and Bohr magneton. This model possesses a remarkable set of diabolical points in the H_x-H_z plane (see Fig. 2). The points lie on a series of concentric rhombi, forming a perfect centered rectangular lattice.[40] All the diabolical points are of the simplest type with Chern numbers ± 1. However, many of the points are multiply diabolical; that is, more than one pair of levels is simultaneously degenerate at the same value of H_x and H_z. The diabolical points on the H_x or H_z axes can be understood in terms of symmetry, but the others cannot. For a given spin magnitude J, there are

$$\tfrac{2}{3}J(J+1)(2J+1) \tag{82}$$

diabolical points in all. It is tempting to speculate that the Hamiltonian possesses an additional "hidden" symmetry, but that is so far unproven.

A more accurate model for Fe_8 includes a correction term to Eq. (81) that is proportional to $(J_+^4 + J_-^4)$. This term destroys the beautiful lattice structure we have described, but it gives rise to other equally beautiful phenomena. For example, as the strength of this term is varied, diabolical points merge and resplit,[41–43] but the total number of points in the three-dimensional $\mathbf{H}$ space does not change. Alas, the experiments[38] can see only a handful (15–20) of the 1540 predicted by theory!

Some of these features are explored for $J=1$ in problems (6) and (7).

V. SUGGESTED PROBLEMS

(1) Since the integral of the approximation (70) for $\mathbf{B}_+$ over the cylinder is independent of the dimensions of the cylinder, this approximation must be divergence-free. Verify directly that $\nabla\cdot\mathbf{B}_+ = 0$.

(2) Equation (45) defines a unitary transformation whose matrix elements are $U_{m'm} = \langle m' | m^* \rangle$. Construct the Hamiltonian in the $|m^*\rangle$ basis by calculating $U\mathcal{H}U^\dagger$ and neglecting terms higher than second order in X and Z. Show that the transformed Hamiltonian is block diagonal with $m^*=0$ being one (1×1) block, and $m^*=1$ and $m^*=-1$ forming a 2×2 block. You should find that the 2×2 block is identical to Eq. (47). The eigenvectors of this block must naturally be written as linear combinations of $|\pm 1^*\rangle$, yielding Eq. (52).

(3) Find the codimension of a triple degeneracy using both Teller's argument and that in Appendix A. Do this for the real symmetric and complex Hermitean cases separately.

(4) The Hamiltonian (43) also has degeneracies at $Z=\pm 1$, $X=Y=0$. Show that these degeneracies are of the simple diabolical type and find the associated Berry curvatures and Chern numbers.

(5) Consider the Hamiltonian (43) for $|\mathbf{R}| \gg 1$ and find the Berry curvature of the highest energy level by treating the $(1-J_z^2)$ term as a perturbation. (It will pay to rotate the coordinate system so that the z axis is along $\mathbf{R}$ and to use symmetry.) Find the Chern number by integrating the curvature so found over a sphere with $|\mathbf{R}| \gg 1$. Is your answer the same as that found by adding the Chern numbers for the same level around the degeneracies in problem (4)? Is it the same as for the highest level of the Hamiltonian (30)? If yes, why, and if not, why not?

(6) (More difficult). Add to the Hamiltonian (43) another term

$$\mathcal{H}' = k'(J_x^2 - J_y^2), \tag{83}$$

with $k' > 0$. If you perform second-order perturbation theory as in Sec. III with k' as an additional small parameter, you should discover the following condition for diabolicity,

$$Z = 0, \quad X^2 + Y^2 = \sqrt{2k'}. \tag{84}$$

But this cannot be correct. It violates the von Neumann–Wigner–Teller theorem on the codimension of a double degeneracy, as the total Hamiltonian is not invariant about the z-axis rotations. The problem might be resolved by higher order perturbation theory, but that is rather difficult. Solving the characteristic equation (a cubic) is also very difficult. A better approach is as follows. Rewrite the Hamiltonian on the J_x basis (or, equivalently, rotate by 90° about $\hat{\mathbf{y}}$) and set $Y=Z=0$. All three eigenvalues are now easily found and you should find a degeneracy at

$$X^* = [2k'(1+k')]^{1/2}, \tag{85}$$

with an energy $E^* = -k'$. This result is exact, and the degeneracy is guaranteed to be an isolated point by the theorem on codimensions. (a) Is the degeneracy so found allowed by a symmetry of the Hamiltonian? (b) Where is the degeneracy when $k' < 0$? (c) Draw the degenerate manifold in the k', X, Y space with $Z=0$. Is the theorem on codimensions of degeneracies being violated?

(7) (More difficult still). This problem is for readers not satisfied that the degeneracy point in problem (6) is isolated. Do not set Y and Z to zero, but treat them as perturbations along with $\delta X = X - X^*$. Do degenerate perturbation theory in the space of the crossing energy levels. You should find that the energies are given by

$$E_\pm = E^* + a(\delta X) \pm (b(\delta X)^2 + cY^2 + dZ^2)^{1/2}, \tag{86}$$

where a, b, c, and d are constants, all of which except a are positive. This is the characteristic form of the diabolo.

ACKNOWLEDGMENT

This work was supported in part by the NSF via Grant No. PHY-0854896.

APPENDIX A: ALTERNATIVE ARGUMENT FOR THE CODIMENSION OF A DIABOLO

The argument given in Sec. II C requires one to use a representation for the Hamiltonian in which all but two of the off-diagonal elements vanish. Many people find this unsettling because it seems to imply that we know how to solve the eigenvalue problem in the first place. We therefore give an alternative argument, which extends Arnold's proof[30] to the case of a complex Hermitian matrix.

Let the matrix be of order n. The general Hermitian matrix of this order has n real numbers on the diagonal, and $(n^2-n)/2$ pairs of complex conjugate numbers off the diagonal. The total number of real numbers needed to specify the matrix is therefore

$$n+2\times\frac{1}{2}(n^2-n)=n^2. \tag{A1}$$

We can also say that the space of complex Hermitian matrices of order n has dimension n^2.

Let us now ask how many free parameters there are if the matrix is to have one double degeneracy. We first look at the eigenvectors, numbering them from 1 to n, with the degenerate ones being numbers $n-1$ and n. The first eigenvector is specified by giving $2n$ real numbers, but normalization and an overall phase remove 2 of them, leaving us with $2n-2$ free parameters. In the same way, the second eigenvector also gives $2n-2$ parameters, but 2 are constrained by the need for it to be orthogonal to the first, leaving us with $2n-4$ free parameters. The third vector has $2n-6$ free parameters, and so on, until we come to the third last eigenvector which has 4 free parameters. There is no freedom to choose the last two eigenvectors because they span the degenerate space, and any linear combination is equally good. Thus, the number of parameters we may vary in specifying the eigenvectors is

$$\begin{aligned}2(n-1)+2(n-2)+\cdots+4&=\tfrac{1}{2}(n-2)(2n-2+4)\\&=n^2-n-2.\end{aligned} \tag{A2}$$

Next, let us consider the eigenvalues. There are $n-2$ nondegenerate eigenvalues, and one degenerate eigenvalue, leading to $n-2+1=n-1$ free parameters. Thus, the dimensionality of the space of matrices with one degeneracy, which is the total number of parameters that may be varied freely without destroying the degeneracy is

$$(n^2-n-2)+(n-1)=n^2-3. \tag{A3}$$

Hence, the codimension of the degeneracy, the number of parameters that must be tuned, is $n^2-(n^2-3)$, which is equal to 3, as found in Sec. II C.

APPENDIX B: ALTERNATIVE CALCULATION OF BERRY CURVATURE

It is interesting to calculate the Berry curvature $\mathbf{B}_+$ for the example in Sec. III using Eq. (22). The sum on n' runs over the states $|0^*\rangle$ and $|-\rangle$, but since E_+-E_0 does not vanish at $\mathbf{R}=0$, whereas E_+-E_- does, only the term with $n'=-$ need be considered as it dominates the sum. Since $\boldsymbol{\nabla}\mathcal{H}=-\mathbf{J}$, we need the $+-$ matrix element of $\mathbf{J}$,

$$\mathbf{J}_{+-}\equiv\langle+|\mathbf{J}|-\rangle. \tag{B1}$$

The matrix for the vector operator $\mathbf{J}$ is

$$\mathbf{J}=J_x\hat{\mathbf{x}}+J_y\hat{\mathbf{y}}+J_z\hat{\mathbf{z}}=\begin{pmatrix}\hat{\mathbf{z}} & \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}-i\hat{\mathbf{y}}) & 0\\ \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}+i\hat{\mathbf{y}}) & 0 & \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}-i\hat{\mathbf{y}})\\ 0 & \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}+i\hat{\mathbf{y}}) & -\hat{\mathbf{z}}\end{pmatrix}. \tag{B2}$$

Therefore, feeding in Eq. (56), we have

$$\mathbf{J}_{+-}=\begin{pmatrix}s & -\frac{X}{\sqrt{2}}(c-s) & -c\end{pmatrix}\begin{pmatrix}\hat{\mathbf{z}} & \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}-i\hat{\mathbf{y}}) & 0\\ \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}+i\hat{\mathbf{y}}) & 0 & \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}-i\hat{\mathbf{y}})\\ 0 & \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}+i\hat{\mathbf{y}}) & -\hat{\mathbf{z}}\end{pmatrix}\begin{pmatrix}c\\ \frac{X}{\sqrt{2}}(c+s)\\ s\end{pmatrix}=2cs\hat{\mathbf{z}}+X(s^2-c^2)\hat{\mathbf{x}}-iX(s^2+c^2)\hat{\mathbf{y}}. \tag{B3}$$

Since $s^2+c^2=1$, $c^2-s^2=\cos\zeta=Z/D$, and $2cs=\sin\zeta=X^2/2D$, Eq. (B3) simplifies to

$$\mathbf{J}_{+-}=-\frac{XZ}{D}\hat{\mathbf{x}}-iX\hat{\mathbf{y}}+\frac{X^2}{2D}\hat{\mathbf{z}} \tag{B4}$$

and

$$\mathrm{Im}(\mathbf{J}_{+-}\times\mathbf{J}_{+-}^{*})=-\frac{X^2}{D}(X\hat{\mathbf{x}}+2Z\hat{\mathbf{z}}). \tag{B5}$$

Finally, since $E_+-E_-=2D$, we get

$$\mathbf{B}_{+}=\frac{X^2}{4D^3}(X\hat{\mathbf{x}}+2Z\hat{\mathbf{z}}), \tag{B6}$$

in agreement with Eq. (70).

[1] M. V. Berry, "Quantal phase factors accompanying adiabatic changes," Proc. R. Soc. London, Ser. A **392**, 45–57 (1984).
[2] J. Anandan, J. Christian, and K. Wanelik, "Resource Letter GPP-1: Geometric phases in physics," Am. J. Phys. **65**, 180–185 (1997).
[3] *Geometric Phases in Physics*, edited by A. Shapere and F. Wilczek (World Scientific, Singapore, 1989).
[4] S. Pancharatnam, "Generalized theory of interference, and its applications," Proc. Indian Acad. Sci., Sect. A **44**, 247–262 (1956).
[5] G. Herzberg and H. C. Longuet-Higgins, "Intersection of potential energy surfaces in polyatomic molecules," Discuss. Faraday Soc. **35**, 77–82 (1963).
[6] H. C. Longuet-Higgins, "The intersection of potential energy surfaces in polyatomic molecules," Proc. R. Soc. London, Ser. A **344**, 147–156 (1975).
[7] A. J. Stone, "Spin-orbit coupling and the intersection of potential energy surfaces in polyatomic molecules," Proc. R. Soc. London, Ser. A **351**, 141–150 (1976).
[8] C. A. Mead and D. Trulhar, "On the determination of Born-Oppenheimer nuclear motion wave functions including complications due to conical intersections and identical nuclei," J. Chem. Phys. **70**, 2284–2296 (1979).
[9] A. Tomita and R. Y. Chiao, "Observation of Berry's topological phase by use of an optical fiber," Phys. Rev. Lett. **57**, 937–940 (1986).
[10] T. Bitter and D. Dubbers, "Manifestation of Berry's topological phase in neutron spin rotation," Phys. Rev. Lett. **59**, 251–254 (1987).
[11] J. Moody, A. Shapere, and F. Wilczek, "Realizations of magnetic-monopole gauge fields: Diatoms and spin precession," Phys. Rev. Lett. **56**, 893–896 (1986).
[12] D. Suter, G. C. Chingas, R. A. Harris, and A. Pines, "Berry's phase in magnetic resonance," Mol. Phys. **61**, 1327–1340 (1987).
[13] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, "Quantized Hall conductance in a two-dimensional periodic potential," Phys. Rev. Lett. **49**, 405–408 (1982).
[14] R. D. King-Smith and D. Vanderbilt, "Theory of polarization of crystalline solids," Phys. Rev. B **47**, 1651–1654 (1993).
[15] G. Sundaram and Q. Niu, "Wave-packet dynamics in slowly perturbed crystals: Gradient corrections and Berry-phase effects," Phys. Rev. B **59**, 14915–14925 (1999).
[16] Y. Zheng and T. Ando, "Hall conductivity of a two-dimensional graphite system," Phys. Rev. B **65**, 245420-1–11 (2002).
[17] V. P. Gusynin and S. G. Sharapov, "Unconventional integer quantum Hall effect in graphene," Phys. Rev. Lett. **95**, 146801-1–4 (2005).
[18] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, "Experimental observation of the quantum Hall effect and Berry's phase in graphene," Nature (London) **438**, 201–204 (2005).
[19] B. Holstein, "The adiabatic theorem and Berry's phase," Am. J. Phys. **57**, 1079–1084 (1989).
[20] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Springer, New York, 1994), Chap. 21.
[21] This potential goes by many other names, such as the Berry connection, the geometrical vector potential, and the molecular Aharonov–Bohm potential.
[22] Another example discussed by Berry, which we review briefly in Sec. II, is that of a spin-j in a magnetic field. In this case the Chern number is $2m$ for the Zeeman level with spin projection along the field equal to m. However, we now have a $(2j+1)$-fold degeneracy at zero field. This state of affairs is still unsatisfactory, as the example is highly special, and we would like a twofold degeneracy with Chern number other than ±1.
[23] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed. (Pergamon, New York, 1977), see Sec. 39, Eq. (39.4) in particular.
[24] K. Gottfried, *Quantum Mechanics* (Benjamin, Reading, MA, 1974), Vol. I, Chap. VI, Problem 1.
[25] In this case we must generalize the more familiar vector calculus concepts of gradient, curl, and divergence using the language of differential forms and exterior calculus. See B. Simon, "Holonomy, the quantum adiabatic theorem, and Berry's phase," Phys. Rev. Lett. **51**, 2167–2170 (1983). A quick introduction to the mathematical language is given in the very readable monograph by H. Flanders, *Differential Forms with Applications to the Physical Sciences* (Dover, New York, 1963).
[26] In applying Berry's ideas to the Born–Oppenheimer approximation, there also arises a scalar potential. See Ref. 20.
[27] An analogy from spherical geometry may be useful here. It is well known that parallel transport of a vector around a geodesic triangle on the surface of the Earth does not bring one back to the original vector. The angle by which the vector has turned is the anholonomy associated with the connection defined by parallel transport.
[28] J. von Neumann and E. P. Wigner, "Über merkwürdige diskrete eigenwerte," Z. Phys. **30**, 467–470 (1929).
[29] E. Teller, "The crossing of potential energy surfaces," J. Phys. Chem. **41**, 109–116 (1937).
[30] For the real symmetric case a very instructive alternative proof is given by V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1978), Appendix 10. We apply the same idea to the complex Hermitian case in Appendix A.
[31] Teller's argument is put in a slightly different way that some readers may find more appealing in Ref. 23, Sec. 79, and Fig. 28.
[32] M. V. Berry and M. Wilkinson, "Diabolical points in the spectra of triangles," Proc. R. Soc. London, Ser. A **392**, 15–43 (1984).
[33] The etymology of the word *diabolo* is interesting and indicates that there is no devilish connotation to it. Professor Michael Widom, in a conversation with the author a few years ago, observed the pleasing assonance with the names of other conic sections. The connection may go beyond the sound of the word, however. According to Wikipedia, the toy originated in China, and a French engineer who promoted it in Europe called it the diabolo because in Greek dia+bolo would mean *throw* (bolo) *across* (dia). Because the hyperbola also is derived from Greek with the rough meaning of *throw* (bolo) *beyond* (hyper), Professor Widom was more correct than either of us knew.
[34] P. A. M. Dirac, "Quantised singularities in the electromagnetic field," Proc. R. Soc. London, Ser. A **133**, 60/1–13 (1931).
[35] The similarity is even stronger in the language of differential forms. We obtain $d\hat{\mathbf{n}}\times d\hat{\mathbf{n}}=K(dx\wedge dy)\hat{\mathbf{n}}$, where $dx\wedge dy$, is the area two-form in the surface with local orthogonal coordinates x and y. See, for example, Sec. 4.5 in Flanders, Ref. 25.
[36] A delightful discussion of this theorem is given by D. H. Gottlieb, "All the way with Gauss-Bonnet and the sociology of mathematics," Am. Math. Monthly **103**, 457–469 (1996).
[37] For a comprehensive review, see D. Gatteschi, R. Sessoli, and J. Villain, *Molecular Nanomagnets* (Oxford U. P., Oxford, 2006).
[38] W. Wernsdorfer and R. Sessoli, "Quantum phase interference and parity effects in magnetic molecular clusters," Science **284**, 133–135 (1999).
[39] A. Garg, "Topologically quenched tunnel splitting in spin systems without Kramers' degeneracy," Europhys. Lett. **22**, 205–210 (1993).
[40] E. Keçecioğlu and A. Garg, "Diabolical points in magnetic molecules: An exactly solvable model," Phys. Rev. B **63**, 064422-1–4 (2001).
[41] E. Keçecioğlu and A. Garg, "SU(2) instantons with boundary jumps and spin tunneling in magnetic molecules," Phys. Rev. Lett. **88**, 237205-1–4 (2002).
[42] E. Keçecioğlu and A. Garg, "Spin tunneling in magnetic molecules: Quasisingular perturbations and discontinuous SU(2) instantons," Phys. Rev. B **67**, 054406-1–13 (2003).
[43] P. Bruno, "Berry phase, topology, and degeneracies in quantum nanomagnets," Phys. Rev. Lett. **96**, 117208-1–4 (2006).