PHYSICAL REVIEW A VOLUME 36, NUMBER 1 JULY 1, 1987

Photon Berry's phase as a classical topological effect

Jan Segert*
Physics Department, Princeton University, Princeton, New Jersey 08544
(Received 19 December 1986)

We discuss Berry's phase rotation for electromagnetic radiation propagating in an optical fiber as a classical effect. We show that the evolution of the polarization vector is determined by a connection on the tangent bundle of the two-dimensional sphere. We use a topological argument to show that there exists only one rotationally invariant connection on the tangent bundle of the two sphere. The arguments apply to any classical transverse wave: for example, transverse vibrations propagating in a bent solid rod. The analogous effect in quantum electrodynamics, namely, Berry's phase for a single photon propagating in an optical fiber, is a simple consequence of the classical effect described. We argue that this should be viewed as a classical, rather than a quantum, effect. We also comment on recent related work of Haldane and of Berry.

I. INTRODUCTION

Berry's phase rotation effects[1,2] for electromagnetic radiation propagating in a waveguide (optical fiber) have been the subject of recent articles. Chiao and Wu[3] have given a theoretical analysis and Tomita and Chiao[4] have provided experimental confirmation. Chiao and Wu consider these effects as "topological features of classical Maxwell theory which originate at the quantum level but survive the correspondence principle limit ($\hbar \rightarrow 0$) into the classical level."[3] In this paper we take the opposite view. We shall argue that this should be viewed as a classical effect. The geometrical and topological setting for these effects is then quite clear.

We shall show that an analogous phase rotation occurs for any classical transverse wave whose direction of propagation is changed adiabatically, for example, transverse vibrations propagating down a bent metal rod. This is, of course, restricted to waveguides, optical fibers, and rods of circular section. The phase rotation property is not just a feature of the classical Maxwell theory, but occurs for all transverse classical waves. Although classical Maxwell theory is merely a limiting case of quantum electrodynamics, and, similarly, a classical vibrational wave in a solid is a limiting case of a quantum phonon process, the simple geometrical origin of the effect is much clearer in the classical theory. The rotation of phase, or polarization vector, is a simple consequence of the topology of a certain fiber bundle over the two-dimensional sphere. Classically, this is the tangent bundle of the sphere. In the quantum theory, this is a complex one-dimensional vector bundle over the sphere, which is isomorphic as a real bundle to the tangent bundle. There is a unique rotationally invariant connection in either case. This is much easier to prove for the tangent bundle. The torsion of the connection, which is defined only for a connection on the tangent bundle and not for a connection on an arbitrary bundle, plays a central role. The proof for the complex one-dimensional bundle, using the theory of homogeneous spaces, is given elsewhere.[5] We shall derive the quantum effect very easily from the classical effect. The quantum effect can, of course, be derived purely in the quantum theory, as was done by Chiao and Wu.[3]

We now discuss the classical effect. The transverse nature of the wave is essential. This is what allows us to define and measure polarization. We assume that the propagation is adiabatic, in the sense that the wave remains polarized and does not undergo partial reflection along the way. This is valid if the direction of the wave is changed slowly enough, i.e., if the curvature of the fiber is kept small. Chiao and Wu made the same assumption in their paper,[3] and it is verified in the experiments of Tomita and Chiao.[4] There is a classical analogue of the adiabatic theorem, which demonstrates the validity of this assumption. For the optical fiber, this follows easily from the quantum adiabatic theorem, when one considers the classical limit of many photons. One can give a more satisfactory proof entirely within the classical theory, but we shall not discuss this here. We refer the reader to Berry's recent work.[6] We will in the following concentrate on linearly polarized waves and on determining the direction of the polarization vector. By the expression "phase rotation" we will mean the angle through which the polarization vector rotates.

We now outline the central argument of this paper. The direction of the optical fiber at each point is specified by a unit vector tangent to the fiber. The parameter space is then the space of all unit vectors in $\mathbb{R}^3$, which is the two-dimensional sphere S^2. The unit tangent vectors along a fiber in $\mathbb{R}^3$ determine a path γ on S^2. The polarization vector is always orthogonal to the direction of propagation and so can be identified with a vector in the tangent vector bundle of S^2. We first give a heuristic argument for determining the evolution of the polarization vector, based on an analogue of Fermi-Walker transport. We shall show that the Fermi-Walker transport is equivalent to parallel transport along γ in a certain connection on the tangent bundle of S^2. We shall then demonstrate that the evolution of the polarization vector is indeed determined by parallel transport along γ in some connection on the tangent bundle. The argument is completed by showing that there is only one connection on the tangent bundle of S^2 which is consistent with the rotational symmetry. This proves that the evolution of the polarization vector is given by the Fermi-Walker transport. We easily recover the formula relating the phase rotation around a closed path on S^2 to the area enclosed by the path.

In the concluding section we discuss the quantum description of light propagating in an optical fiber. We shall see that the photon Berry's phase effect is a simple consequence of the classical effect. Although this may be derived as a purely quantum effect, we argue that this is more appropriately viewed as a classical effect.

Comments on the recent work by Haldane[7] and by Berry[6] may be found throughout the paper. Haldane[7] calculated the phase rotation using the Fermi-Walker prescription. Haldane's justification for the Fermi-Walker prescription is the fact that any fiber can be approximated by a piecewise planar fiber. Berry[6] showed that the Fermi-Walker prescription is the correct adiabatic limit of the Maxwell equations. There is some overlap in the results of these two papers and the present paper, but the methods are quite different. I reduce the problem to a geometric problem on the sphere, which is then solved using topological and group-theoretical methods.

II. ELEMENTARY ANALYSIS

We parametrize the path of the waveguide in three-dimensional space $\mathbb{R}^3$ by t, where t varies between 0 and l, the length of the path. It is convenient to take t at a point to be the distance along the path to the initial point. We denote the unit tangent vector by $\mathbf{k}_t$. There is a unit vector $\mathbf{e}_t$ orthogonal to $\mathbf{k}_t$ which characterizes the polarization. Given a path and some initial polarization vector $\mathbf{e}_0$, we wish to find $\mathbf{e}_t$, the polarization vector at a point farther along the path. We will in the following wish to compare vectors at different points of $\mathbb{R}^3$. To do so, we rigidly translate them to the same point. Note that if $\mathbf{k}_l=\mathbf{k}_0$, which by the above convention means that $\mathbf{k}_0$ and $\mathbf{k}_l$ are parallel, then all possible vectors $\mathbf{e}_0$ and $\mathbf{e}_l$ lie in the plane orthogonal to $\mathbf{k}_0=\mathbf{k}_l$.

Using time-reversal invariance of the equations of motion, it is clear that any final polarization is possible with the appropriate choice of initial polarization. Consider two distinct initial polarizations $\mathbf{e}_0$ and $\tilde{\mathbf{e}}_0$. From the linearity of the equations of motion, and from the fact that $\mathbf{e}_0\cdot\mathbf{e}_0=\mathbf{e}_l\cdot\mathbf{e}_l$, we conclude that $\mathbf{e}_0\cdot\tilde{\mathbf{e}}_0=\mathbf{e}_l\cdot\tilde{\mathbf{e}}_l$. Now consider a path with $\mathbf{k}_0=\mathbf{k}_l$. Then there is a natural identification of the space of initial polarization vectors with the space of final polarization vectors, given by rigid translation. So we can measure the angle between the initial and final polarization. This angle is the same for any choice of initial polarization, since dot products are preserved. So for a path with $\mathbf{k}_0=\mathbf{k}_l$, there is a well-defined notion of phase rotation associated with the path. Of course this might well be zero. Now suppose that $\mathbf{k}_0\neq\mathbf{k}_l$. Then there is no such natural notion of phase rotation, since there is no longer a natural identification between the space of initial polarization vectors and the space of final polarization vectors.

We will define a notion of phase rotation in a way which involves some arbitrariness. We take a field of unit vectors $\mathbf{u}_t$ perpendicular to $\mathbf{k}_t$, varying smoothly with t, such that if for any t_1 and t_2 between 0 and l, $\mathbf{k}_{t_1}=\mathbf{k}_{t_2}$ then $\mathbf{u}_{t_1}=\mathbf{u}_{t_2}$. We call a field of unit vectors satisfying these conditions admissible. We now define phase rotation in the following way. Measure the angle between $\mathbf{e}_0$ and $\mathbf{u}_0$. Then measure the angle between $\mathbf{e}_l$ and $\mathbf{u}_l$. Subtract the two angles. This notion clearly does not depend on $\mathbf{e}_0$. Since the angle is measured relative to the arbitrarily chosen admissible $\mathbf{u}$, we get a result which depends on the choice of $\mathbf{u}$. Note, however, that all choices of $\mathbf{u}$ give the same result for the case $\mathbf{k}_0=\mathbf{k}_l$.

We now construct such a $\mathbf{u}_t$ for a large class of paths. Fix an arbitrary unit vector $\mathbf{w}$. If $\mathbf{k}_t\neq\pm\mathbf{w}$, define

$$\mathbf{u}_t=\frac{\mathbf{k}_t\times\mathbf{w}}{[(\mathbf{k}_t\times\mathbf{w})\cdot(\mathbf{k}_t\times\mathbf{w})]^{1/2}} .$$

Complete this to an orthonormal basis at each point of the path by defining $\mathbf{v}_t=\mathbf{k}_t\times\mathbf{u}_t$. Note that $\mathbf{u}$ and $\mathbf{v}$ are both admissible. Any linear combination of unit length is also admissible.

Given an initial polarization vector $\mathbf{e}_0$, we wish to determine $\mathbf{e}_t$. The normal notion of parallel transport in $\mathbb{R}^3$, i.e., rigid translation, is clearly not what we want here, since it will not respect the condition that $\mathbf{e}_t\cdot\mathbf{k}_t=0$. The notion we need is analogous to the Fermi-Walker transport of vectors familiar from relativity.[8,9] We will defer the proof that this is the correct notion. This will follow easily from the results of the later sections. We shall show later that this rule for transport is the only rule compatible with the rotational invariance of the system. We refer also to Haldane[7] and Berry[6] for alternate arguments.

Choose t between 0 and l. Choose a vector $\mathbf{e}_t$. We wish to transport this vector to $t+\delta$. We subtract from $\mathbf{e}_t$ the component in the direction of $\mathbf{k}_{(t+\delta)}$. The resulting vector is perpendicular to $\mathbf{k}_{(t+\delta)}$, and of length $1+O(\delta^2)$. To transport from $t=0$ to $t=l$, we divide the path into segments of length δ and perform this operation successively for each segment. Of course we are interested in the limit $\delta\to 0$. We now calculate the phase rotation. Transport $\mathbf{u}_t$ in the above manner to $t+\delta$. The result can be expressed as a linear combination of $\mathbf{u}_{(t+\delta)}$ and $\mathbf{v}_{(t+\delta)}$. By definition, the phase rotation is to first order in δ the coefficient of $\mathbf{v}_{(t+\delta)}$, which is equal to $\mathbf{u}_t\cdot\mathbf{v}_{(t+\delta)}$. To find the phase rotation for the whole path, we sum over the successive segments and take the limit $\delta\to 0$. The phase rotation angle Ψ is given by

$$\Psi=\int_0^l \mathbf{u}_t\cdot\mathbf{v}_t'\,dt \ , \tag{1}$$

where the prime denotes derivative with respect to t. For our specific choice of $\mathbf{u}$, a straightforward calculation yields

$$\mathbf{u}\cdot\mathbf{v}'=\frac{(\mathbf{k}'\cdot(\mathbf{k}\times\mathbf{w}))(\mathbf{k}\cdot\mathbf{w})}{1-(\mathbf{k}\cdot\mathbf{w})^2} . \tag{2}$$

III. CONNECTIONS ON THE TANGENT BUNDLE

We still have not justified the statement that the polarization vector is determined by Fermi-Walker transport.

We shall now provide part of this justification. We shall first show that the evolution of the polarization vector is determined by parallel transport along the path γ in some connection on the tangent bundle. We then demonstrate that the Fermi-Walker transport determines a connection on the tangent bundle of S^2. It then remains to show that these two connections coincide. This will be done in Sec. V, wherein we prove that there is only one rotationally invariant connection on the tangent bundle. Since each of the above connections must be rotationally invariant, they coincide. Properties of connections on the tangent bundle are discussed in books on relativity[9,8] and differential geometry.[10,11]

We now demonstrate that the polarization vector is determined by parallel transport in some connection on the tangent bundle of S^2. Consider an optical fiber in $\mathbb{R}^3$, parametrized by length as before. The unit tangent vectors to the fiber determine a path γ on the parameter space S^2. $\mathbf{k}_0$ is the initial point and $\mathbf{k}_l$ the final point of the path γ. We now send polarized light down the fiber, with initial polarization vector $\mathbf{e}_0$, which satisfies $\mathbf{e}_0\cdot\mathbf{k}_0=0$. The final polarization is described by a vector $\mathbf{e}_l$, which satisfies $\mathbf{e}_l\cdot\mathbf{k}_l=0$. In other words, $\mathbf{e}_0$ is a tangent vector to S^2 at the point $\mathbf{k}_0$ and $\mathbf{e}_l$ is tangent at $\mathbf{k}_l$. For any path γ from $\mathbf{k}_0$ to $\mathbf{k}_l$, the evolution of the polarization vector gives a map of the unit tangent vectors at $\mathbf{k}_0$ to the unit tangent vectors at $\mathbf{k}_l$. We shall later verify that the polarization indeed depends only on the geometric path γ and not on the rate at which γ is traversed. We note again that the equations of motion are invariant under time reversal. Thus if we send polarized light down the path γ in the opposite direction, starting at $\mathbf{k}_l$ with polarization $\mathbf{e}_l$, the light emerges at $\mathbf{k}_0$ with polarization $\mathbf{e}_0$.

We also note that the polarization is defined at every intermediate point of the path. If we cut the optical fiber at any point, the light emerges polarized at that point. We call the tangent to the fiber at this intermediate point $\mathbf{k}_m$. The path γ on S^2 is now divided into two paths, γ_1 from $\mathbf{k}_0$ to $\mathbf{k}_m$, and γ_2 from $\mathbf{k}_m$ to $\mathbf{k}_l$. Suppose that the light sent down the fiber at the initial point, with polarization $\mathbf{e}_0$, emerges at the intermediate point with polarization $\mathbf{e}_m$, satisfying of course $\mathbf{e}_m\cdot\mathbf{k}_m=0$. If we now rejoin the two pieces of the fiber, we get light at the final point with polarization $\mathbf{e}_l$ as before. It is then clear that sending light down the piece of the fiber from the intermediate point to the final point, with initial polarization $\mathbf{e}_m$, we get light out with final polarization $\mathbf{e}_l$.

We now show that the polarization depends only on the geometric path γ on S^2, not on the speed at which it is traversed. This is easily verified by noting that inserting straight sections of fiber will not change the final polarization. Alternately, any path of the fiber in $\mathbb{R}^3$ can be approximated by a path which consists of straight sections alternating with curved sections. By shortening or lengthening the straight sections, the geometric path γ is unaltered, although it is traversed at a different speed. We have now reduced the question so it depends only on the geometric path on S^2. All further calculations will involve only the path γ on S^2, and not the path of the fiber in $\mathbb{R}^3$. This approach contrasts with that of Haldane.[7] Haldane, following Ross,[12] argues that any path in $\mathbb{R}^3$ can be approximated by a piecewise planar path, and thereby derives the Fermi-Walker prescription directly. He then derives the fact that the result depends only on the geometric path γ on S^2 from properties of the Fermi-Walker prescription. This results in a shorter derivation, but a somewhat less geometric picture.

We thus have the following situation. To every smooth path γ on S^2, we associate a map P_γ of the unit tangent vectors at the initial point $\mathbf{k}_0$ to the unit tangent vectors at the final point $\mathbf{k}_l$. We have just shown that this map has two properties.

(i) The inverse path γ^{-1} from $\mathbf{k}_l$ to $\mathbf{k}_0$ yields the inverse map

$$P_{\gamma^{-1}}=(P_\gamma)^{-1}\ .$$

(ii) The smooth composition of two paths γ_2 and γ_1 yields the composition of maps

$$P_{\gamma_2\circ\gamma_1}=P_{\gamma_2}\circ P_{\gamma_1}\ .$$

A rule P_γ for transporting tangent vectors along a path γ, satisfying the above properties, is the parallel transport in some connection on the tangent bundle.[10] We have discussed parallel transport for unit tangent vectors, but we can easily extend by linearity to all tangent vectors. The resulting connection preserves the lengths of vectors. The usual local expression of the connection, in terms of a connection form, or gauge potential, can be found by considering infinitesimally short paths. Conversely, from the local description of a connection, it is easy to verify that the parallel transport P_γ satisfies the properties above. We have seen that the evolution of the polarization vector depends on the path γ and satisfies properties (i) and (ii). Thus we conclude that the evolution of the polarization vector is determined by parallel transport along γ in some connection on the tangent bundle of S^2. This connection further preserves the lengths of vectors.

It is now easy to check that the Fermi-Walker prescription also determines a connection on the tangent bundle of S^2. Fermi-Walker transport clearly satisfies properties (i) and (ii) above. The Fermi-Walker transport also preserves the lengths of vectors. It remains to prove that the connection determined by the Fermi-Walker transport is the same as the connection which determines the polarization. This will be demonstrated in Sec. V, by showing that there exists only connection consistent with the rotational symmetry of the problem. We now return to the local analysis.

IV. LOCAL ANALYSIS ON THE SPHERE

We now express Eqs. (1) and (2) of Sec. II in an intrinsic way. The vector $\mathbf{k}'$ is orthogonal to $\mathbf{k},\mathbf{k}'\cdot\mathbf{k}=0$. It is thus tangent to the sphere at the point $\mathbf{k}$. The expression (2) defines a linear functional on the tangent vectors, in other words, a differential form of degree 1. We shall denote this form by α. Let $\mathbf{p}$ be a vector tangent to the point $\mathbf{k}$ on S^2. Then we have

$$\alpha(\mathbf{p})=\frac{(\mathbf{p}\cdot(\mathbf{k}\times\mathbf{w}))(\mathbf{k}\cdot\mathbf{w})}{1-(\mathbf{k}\cdot\mathbf{w})^2}\ .$$

The one-form α is defined everywhere on S^2 except at the poles, where $\mathbf{u}$ and $\mathbf{v}$ are not defined.

We express this form intrinsically. We use standard spherical coordinates, with $\mathbf{w}$ at $\theta=0$ and $-\mathbf{w}$ at $\theta=\pi$. We can express $\mathbf{u}$ and $\mathbf{v}$ as vector fields on the sphere, which are not defined at the poles. Specifically we have

$$\mathbf{u}=\hat{\mathbf{e}}_{\phi}, \quad \mathbf{v}=\hat{\mathbf{e}}_{\theta} .$$

Using $\mathbf{k}\cdot\mathbf{w}=\cos(\theta)$ and $\mathbf{k}\times\mathbf{w}=\mathbf{u}\sin(\theta)$, we find

$$\alpha(\mathbf{p})=\frac{(\mathbf{p}\cdot\mathbf{u})\sin(\theta)\cos(\theta)}{1-\cos^2(\theta)} .$$

The one form dual to $\mathbf{u}$ in the standard metric is $-\sin(\theta)d\phi$. We conclude that $\alpha=-\cos(\theta)d\phi$. It is easy to check, using the Cartan structural equations,[11,10] that α is the local expression relative to the orthonormal basis $(\mathbf{u},\mathbf{v})$ of the Levi-Civita connection associated with the standard metric on S^2. This will also become clear in Sec. V, when we show that there is only one rotationally invariant connection on the tangent bundle of S^2; hence, these two connections must coincide.

We now calculate the phase for a closed path γ on S^2. We can assume that γ is not self-intersecting. Recall that for a closed path, the phase is well defined and independent of the arbitrariness involved in choosing $\mathbf{u}$ and $\mathbf{v}$. We cannot, however, allow the path to hit either of the poles. This does not, of course, mean that the phase is not defined for a path that hits a pole, but only that the above calculation fails. The path γ divides the sphere into two regions. We denote by $\Omega(\gamma)$ the region whose boundary γ traverses in the counterclockwise direction. This is equivalent to giving the usual orientation on S^2 and saying that γ is the boundary of $\Omega(\gamma)$. First consider a path γ such that $\Omega(\gamma)$ excludes both poles. One can check that the integral of α around γ is just the area of $\Omega(\gamma)$. If we consider a path γ for which $\Omega(\gamma)$ includes one pole, then the integral of α around γ is equal to the area of $\Omega(\gamma)$ minus 2π. If γ includes both poles, we must subtract 4π. To check this, consider a small loop around one pole. This is analogous to the magnetic monopole of charge 2 with two Dirac strings. A waveguide which is planar in $\mathbb{R}^3$ corresponds to γ confined to the equator, so the phase is 0 or $\pm 2\pi$. Since the phase is defined (mod 2π), we recover the usual result.

One might worry about having to exclude paths that hit the poles. It is our choice of $\mathbf{u}$ and $\mathbf{v}$ which forces us to do this. Consider some different admissible field of vectors $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$, and calculate the corresponding one-form $\tilde{\alpha}$. Suppose that $\tilde{\alpha}$ is defined everywhere on S^2. The physical result, the phase around closed paths on S^2, is independent of whether we compute it using α or $\tilde{\alpha}$. We conclude that

$$\oint_{\gamma}\tilde{\alpha}=A(\Omega(\gamma))+n2\pi ,$$

where $A(\Omega(\gamma))$ is the area of the region $\Omega(\gamma)$ and n is some integer. If we contract γ to a point, then the integral of $\tilde{\alpha}$ around γ goes to zero and the area of $\Omega(\gamma)$ goes either to zero or to 4π. So depending on how we choose to contract γ, we find either $n=0$ or $n=-2$. This contradiction shows that $\tilde{\alpha}$ cannot be everywhere defined. We could also see this using Gauss's law,[11,13]

$$\oint_{\gamma}\tilde{\alpha}=\int_{\Omega(\gamma)}d\tilde{\alpha} .$$

Suppose $\tilde{\alpha}$ is everywhere defined. Gauss's law implies that $d\tilde{\alpha}$ is the area form. But the area form of a compact two-dimensional manifold M represents a nonzero element of $H^2(M,R)$, the second cohomology with real coefficients,[13,14] and thus cannot be an exact form. It is the fact that the tangent bundle of S^2 is topologically nontrivial[14] that prevents us from being able to globally define the form $\tilde{\alpha}$. We must thus take into account the topology of the tangent bundle.

V. GLOBAL ANALYSIS ON THE SPHERE

We now examine the global geometry and topology of S^2. We have shown that the Fermi-Walker prescription gives a connection on the tangent bundle. We have also seen that the evolution of the polarization vector is determined by parallel transport in some connection on the tangent bundle. It is clear that both the Fermi-Walker transport and the evolution of the polarization vector must yield rotationally invariant connections. This merely corresponds to the freedom of rotating the entire optical fiber rigidly about some point of $\mathbb{R}^3$. We shall now show that these two connections coincide, so that the evolution of the polarization is in fact determined by the Fermi-Walker transport. We do this by proving that there is only one rotationally invariant connection on the tangent bundle of S^2. This proof uses the fact that the tangent bundle of the sphere is topologically nontrivial, more specifically that the Euler characteristic of the sphere is nonzero.[14,13] The uniqueness can also be proved using the theory of homogeneous spaces.[5]

The two sphere has Euler characteristic $\chi(S^2)=+2$, so it does not admit any vector fields without zeros.[13,14] Specifically, there is no vector field on S^2 of unit length everywhere. This fact implies that there must be a phase rotation for some closed path γ. Suppose there were no phase rotation. Then we could define a unit vector field everywhere on the sphere by choosing a base point and a vector tangent to the base point. At any other point on the sphere, we define the vector by parallel transporting along a path from the base point. This is independent of the path, since transport around a closed path does not change the vector. Since no unit vector field can exist, we conclude that there must be a phase rotation at least for some path γ.

We can deduce the exact form of the phase rotation from invariance of the system under spatial rotation. We are looking for a connection on the tangent bundle of S^2 which preserves the standard metric and which is invariant under the SO(3) action on S^2. We first show that any invariant connection must be torsionless.[11,15,10] Recall that the torsion tensor can be expressed as two form taking values in the tangent bundle. By the transitivity of the SO(3) action, if the torsion vanishes at any point of S^2, it must vanish at all points. Now any ordinary two form on S^2 is equal to the area form multiplied by some function. Similarly, any tangent vector valued two form is equal to the area form tensored with some vector field.

We know that all vector fields must have a zero somewhere, so the torsion must be zero at some point, hence it must be zero at all points.

There exists a unique metric-preserving torsionless connection, the usual Levi-Civita connection,[11,15,10] familiar from general relativity. We thus conclude that there is a unique rotationally invariant connection on the tangent bundle of S^2. We recall that the polarization vector is determined by parallel transport in some connection on the tangent bundle. Since there is only one rotationally invariant connection, the polarization vector is determined by parallel transport in this connection. We also showed that the Fermi-Walker transport defines a connection on the tangent bundle of S^2. This connection must again be equal to the unique rotationally invariant connection. We thus conclude that the polarization vector is determined by Fermi-Walker transport.

Thus the phase rotation is determined by the parallel transport of tangent vectors using this connection. Given the path of the waveguide in $\mathbb{R}^3$ and the initial polarization, we find the final polarization in the following way. The unit vector tangent to the waveguide defines a path γ from $\mathbf{k}_0$ to $\mathbf{k}_l$ on S^2. The initial polarization $\mathbf{e}_0$ is a unit vector tangent to S^2 at the point $\mathbf{k}_0$. To find the final polarization $\mathbf{e}_l$, simply parallel transport this vector along the path γ to the point $\mathbf{k}_l$.

We now study closed paths on S^2. The curvature two-form of an invariant connection is also invariant and so must be a multiple of the area form. The curvature is easily found by explicit calculation, or by a simple application of the Gauss-Bonnet-Chern theorem.[11,15,10] In two dimensions, this theorem states that the integral of the curvature two-form over the manifold equals 2π times the Euler characteristic. Since the area of the unit sphere is 4π, we conclude that the curvature form is equal to the area form. Recall that the curvature form is defined by considering parallel transport around an infinitesimally small closed counterclockwise path γ.[9] By definition, the integral over $\Omega(\gamma)$ of the curvature form is equal to the phase rotation around γ. In the case of S^2, the phase rotation for a closed path which is not infinitesimally small can be built up from the infinitesimal case. Thus we recover the expression for the phase rotation in terms of the area of $\Omega(\gamma)$, but now we have no problems with singularities. The problems with singularities arose when we tried to define a global orthonormal frame. This cannot be done, because there are no vector fields without zeros. Note also that the previous local description in terms of the connection form α does not explicitly exhibit the rotational symmetry of the problem.

This derivation of the phase rotation is longer than Haldane's derivation, but it gives additional insight into the origin of the formula. The existence of some nontrivial phase rotation is a consequence of the topology of S^2, namely, the nonzero Euler characteristic. The exact form of the phase rotation is a consequence of the rotational invariance.

VI. THE QUANTUM CASE AND DISCUSSION

We first note that the above discussion is not specific to the Maxwell theory. We only require the transverse nature of the wave, so that the polarization vector, transverse to the direction of propagation, is defined. Another example of a classical transverse wave is a transverse vibrational wave propagating in a solid rod of circular section. It should be an easy matter to detect the polarization rotation experimentally in this case. This example is clearly classical, although the polarization rotation could presumably be derived as the classical limit of quantum Berry's phase for phonons. Such a derivation would, however, obscure the simple geometrical origin of the phase rotation effect.

We now discuss single photons passing through an optical fiber. This is no longer in the realm of the classical Maxwell theory. We must use quantum electrodynamics. We shall show that the phase or polarization rotation for a single photon follows easily from the classical effect. We refer to Sakurai[16] or Itzykson and Zuber[17] for a discussion of the quantization of electromagnetic radiation. We shall use natural units, where c and $\hbar$ are equal to 1.

We review the quantization of electromagnetic radiation in free space. First consider the classical vector potential $\mathbf{A}$. We use the radiation gauge $\nabla\cdot\mathbf{A}=0$. Each component (A_x, A_y, A_z) of the vector potential obeys the wave equation. This is schematically written as

$$\nabla^2\mathbf{A}-\frac{\partial^2\mathbf{A}}{\partial t^2}=0 \ .$$

Solutions of the wave equation are linear combinations of plane waves

$$\mathbf{u}_{\mathbf{k},\alpha}(\mathbf{x})=\mathbf{e}_\alpha\exp(i\mathbf{k}\cdot\mathbf{x}), \quad \alpha=1,2 \ .$$

The radiation gauge requires $\mathbf{e}_\alpha\cdot\mathbf{k}=0$. Conventionally, one chooses $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{k}$ mutually orthogonal, and $\mathbf{e}_\alpha$ unit vectors. It is convenient to assume periodic boundary conditions in a cube of side L. A general solution of the classical equations is given by[16]

$$\mathbf{A}(\mathbf{x},t)=\frac{1}{(L^3)^{1/2}}\sum_{\mathbf{k},\alpha}[c_{\mathbf{k},\alpha}(t)\mathbf{u}_{\mathbf{k},\alpha}+c^*_{\mathbf{k},\alpha}(t)\mathbf{u}^*_{\mathbf{k},\alpha}] \ .$$

Here, $c_{\mathbf{k},\alpha}(t)$ are time-dependent Fourier coefficients, satisfying

$$\frac{d^2c_{\mathbf{k},\alpha}}{dt^2}+\omega^2c_{\mathbf{k},\alpha}=0, \quad \omega^2=\mathbf{k}\cdot\mathbf{k} \ ,$$

and $\mathbf{u}^*_{\mathbf{k},\alpha}=\mathbf{e}_\alpha\exp(-i\mathbf{k}\cdot\mathbf{x})$ is the complex conjugate of $\mathbf{u}_{\mathbf{k},\alpha}$. The quantized radiation field is described by replacing the Fourier coefficients $c_{\mathbf{k},\alpha}$ and $c^*_{\mathbf{k},\alpha}$ with photon creation and annihilation operators $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$. Conventionally, one includes some factors in the definition of the creation and annihilation operators:

$$c_{\mathbf{k},\alpha}\rightarrow\left[\frac{1}{2\omega}\right]^{1/2}a_{\mathbf{k},\alpha}, \quad c^*_{\mathbf{k},\alpha}\rightarrow\left[\frac{1}{2\omega}\right]^{1/2}a^\dagger_{\mathbf{k},\alpha} \ .$$

The quantum field operator $\mathbf{A}$ takes the form

$$\mathbf{A}=\frac{1}{(L^3)^{1/2}}\sum_{\mathbf{k},\alpha}\frac{1}{2\omega}[a_{\mathbf{k},\alpha}(t)\mathbf{u}_{\mathbf{k},\alpha}+a^\dagger_{\mathbf{k},\alpha}(t)\mathbf{u}^*_{\mathbf{k},\alpha}] \ . \tag{3}$$

The vacuum state, the state with no photons, is denoted by $|0\rangle$. The state with n identical photons, each with momentum $\mathbf{k}$ and polarization $\mathbf{e}_\alpha$, is given by

$$\frac{1}{\sqrt{n!}}(a^\dagger_{\mathbf{k},\alpha})^n|0\rangle \ . \tag{4}$$

This is an eigenstate of the quantum Hamiltonian. The classical plane wave described in the limit n is very large.

We now discuss measurement of photon polarization. For simplicity, we consider a state of one photon, linearly polarized, with momentum $\mathbf{k}$ and polarization vector $\mathbf{e}=\mathbf{e}_1\cos(\theta)+\mathbf{e}_2\sin(\theta)$. This corresponds to the state

$$[a^\dagger_{\mathbf{k},1}\cos(\theta)+a^\dagger_{\mathbf{k},2}\sin(\theta)]\,|0\rangle \ .$$

This is clearly an eigenstate of the Hamiltonian. Consider measurements confined to some bounded region of space. Such a measurement can determine the polarization vector of the photon. Measurements of the polarization vector made in different regions of space agree. The polarization of the quantum photon is determined by polarization of the corresponding solution of the classical Maxwell equations; in this case, a linearly polarized plane wave.

Now consider the optical fiber. The classical solutions are no longer plane waves, but instead propagate along the direction of the fiber. The quantization of the electromagnetic fields proceeds exactly as in free space. The only difference is that the classical solutions are now modes of the fiber, not plane waves. Locally the fiber has a circular symmetry about the axis. We are interested in modes for which polarization is defined, hence, modes which are not circularly symmetric. Each such mode is specified by an index (β,α), where $\alpha=1,2$ specifies the direction of the linear polarization as before. In the discussion of free space, the mode index β corresponds to the momentum $\mathbf{k}$. The expression for the quantum field operator $\mathbf{A}$ takes the same form as above, Eq. (3), with $\mathbf{k}$ replaced by β. Consider the state of one linearly polarized photon of a certain mode propagating in the fiber, described by Eq. (4), with $\mathbf{k}$ replaced by β. This is an eigenstate of the quantum Hamiltonian. Consider measurements of the polarization vector. A measurement of the photon polarization at any point in the fiber yields the polarization vector of the classical solution corresponding to the photon. We have shown that the polarization vector of the classical solution undergoes rotation, depending on the path of the fiber. The polarization of the single photon undergoes exactly the same rotation as the classical solution.

We thus see that the quantum effect is a direct consequence of the classical effect. A photon is associated with a particular solution of the classical Maxwell equations. The polarization of the photon is the same as the polarization of the corresponding classical solution. The polarization of the classical solutions is determined by the uniqueness of the rotationally invariant connection on the tangent bundle of the sphere. The polarization of the photon is then determined by a classical topological effect.

ACKNOWLEDGMENT

I thank M. V. Berry for bringing Haldane's work to my attention, and for discussions on the earlier version of this paper and his own work.

---

*Address after September 1987: Department of Physics, California Institute of Technology, Pasadena, CA 91125.

[1]M. V. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984).
[2]B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).
[3]R. Y Chiao and Y. S. Wu, Phys. Rev. Lett. **57**, 933 (1986).
[4]A. Tomita and R. Y. Chiao, Phys. Rev. Lett. **57**, 937 (1986).
[5]J. Segert, J. Math. Phys. (to be published).
[6]M. V. Berry, Nature **326**, 277 (1987).
[7]F. D. M. Haldane, Opt. Lett. **11**, 730 (1986).
[8]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, 1973).
[9]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[10]S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Wiley, New York, 1969), Vols. 1 and 2.
[11]Y. Choquet-Bruhat and C. DeWitt-Morette, *Analysis, Manifolds, and Physics,* revised ed. (Elsevier, Amsterdam, 1982).
[12]J. N. Ross, Optical and Quantum Electronics **16**, 455 (1984).
[13]R. Bott and L. W. Tu, *Differential Forms in Algebraic Topology* (Springer-Verlag, Berlin, 1982).
[14]J. W. Milnor and J. D. Stasheff, *Characteristic Classes* (Princeton University, Princeton, N.J., 1974).
[15]T. Eguchi, P. B. Gilkey, and A. J. Hanson, Phys. Rep. **66**, 213 (1980).
[16]J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1982).
[17]C. Itzykson and J. B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).