PHYSICAL REVIEW LETTERS

VOLUME 60 | 6 JUNE 1988 | NUMBER 23

General Setting for Berry's Phase

Joseph Samuel and Rajendra Bhandari

Raman Research Institute, Bangalore 560080, India

(Received 30 November 1987)

It is shown that Berry's phase appears in a more general context than realized so far. The evolution of the quantum system need be neither unitary nor cyclic and may be interrupted by quantum measurements. A key ingredient in this generalization is the use of some ideas introduced by Pancharatnam in his study of the interference of polarized light, which, when carried over to quantum mechanics, allow a meaningful comparison of the phase between any two nonorthogonal vectors in Hilbert space.

PACS numbers: 03.65.Bz, 02.40.+m

Three years ago, Berry[1] made a rather perceptive and interesting observation regarding the behavior of quantum-mechanical systems in a slowly changing environment. If the system is initially in an eigenstate of the instantaneous Hamiltonian, the adiabatic theorem guarantees that it remains so. This, however, determines the state of the system only up to a phase. Berry asked the question "What is the phase of the system?" and got a somewhat unexpected answer. If the environment (more precisely, the Hamiltonian) returns to its initial state, the system also does, but it acquires an extra phase over and above the dynamical phase, which can be calculated and allowed for. This effect has been studied and measured in various contexts.

Simon[2] gave a simple geometrical interpretation of Berry's phase. If one regards the space of normalized states as a fiber bundle over the space of rays[3] (a ray is defined as an equivalence class of states differing only in phase), then this bundle has a natural connection. This connection permits a comparison of the phases of states on two neighboring rays. Simon observed that when the dynamical phase factor is removed, the evolution of the system as determined by the Schrödinger equation is a parallel transport of the phase of the system according to this natural connection. Berry's phase is then a consequence of the curvature of this connection.

Recently, Aharanov and Anandan[4] generalized Berry's results by giving up the assumption of adiabaticity. The key step in this work is their identification of the integral of the expectation value of the Hamiltonian as the dynamical phase. Once this dynamical phase is removed, the evolution of the phase of the system is again determined by the natural connection and one recovers Berry's phase for any cyclic evolution of the quantum system.

The purpose of this Letter is to point out that Berry's phase appears in a still more general context. The evolution of the system need be neither unitary (norm preserving) nor cyclic (returning to the original ray). This generalization is based on the work of Pancharatnam[5] on the interference of polarized light. Carrying Pancharatnam's ideas over to quantum mechanics yields a fairly general setting for a discussion of Berry's phase. We briefly describe Pancharatnam's work before developing the subject of the present paper.

Pancharatnam posed the following question: Given two beams of polarized light, is there a natural way to compare the phases of these beams? His physically motivated answer was to cause interference of these two polarized beams and regard them as "in phase" when the resultant intensity is maximum. This provides a "connection" (a rule for the comparison of phases) between any two states of polarization which are not orthogonal. (This rule breaks down for orhtogonal states. These do not interfere and the resultant intensity is insensitive to the relative phase of the two beams.) Consider three (nonorthogonal) states of polarization represented by three points 1, 2, and 3 on the Poincaré sphere. Suppose now that 1 and 2 are "in phase" and 2 and 3 are "in phase"; then 1 and 3 are not necessarily in phase. Pancharatnam showed that the excess phase of 3 over 1 is given by half the solid angle subtended by the spherical triangle 123 at the center of the Poincaré sphere. Thus the Pancharatnam connection has curvature. While Pancharatnam's studies, both theoretical and experimental, were carried out in the 1950's, the relation between his work and Berry's was pointed out only recently by Ramaseshan and Nityananda.[6] They, and subsequently Berry,[7] observed that Pancharatnam's excess phase is in fact an early example of Berry's phase. A laser interferometer experiment demonstrating Pancharatnam's excess phase has earlier been reported by us.[8] In the rest of this paper we show that carrying over Pancharatnam's ideas to quantum mechanics leads to a fruitful generalization of the Berry's phase.

Consider a quantum system whose state vector $|\psi\rangle$ (an element of a Hilbert space $\mathcal{H}$) evolves according to the Schrödinger equation $i(d/dt)|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle$. ($\hat{H}$ is a linear operator, possibly non-Hermitean.) Let us define a new state vector $|\phi(t)\rangle$, which differs from $|\psi(t)\rangle$ only in that it has had a dynamical phase factor removed:

$$|\phi(t)\rangle = \exp\left[i\int_0^t h(t')dt'\right]|\psi(t)\rangle,$$

where

$$h(t') = \langle\psi|\psi\rangle^{-1}\operatorname{Re}\langle\psi(t')|\hat{H}(t')|\psi(t')\rangle.$$

Clearly, $|\phi(t)\rangle$ satisfies the equation

$$i(d/dt)|\phi(t)\rangle = [\hat{H}(t) - h(t)]|\phi(t)\rangle.$$

Contracting this with $\langle\phi(t)|$ yields the parallel-transport law

$$\operatorname{Im}\langle\phi(t)|(d/dt)|\phi(t)\rangle = 0. \tag{1}$$

While this law has its origin in the Schrödinger equation, it is purely geometric, as are the considerations in the rest of this paper.

Let $\mathcal{N}$ denote the set of normalizable states in $\mathcal{H}$:

$$\mathcal{N} = \{|\psi\rangle \in \mathcal{H} \mid \langle\psi|\psi\rangle \neq 0\}.$$

Let $\mathcal{R}$ be the space of rays: $\mathcal{R} = \mathcal{N}/\sim$, where $\sim$ denotes that elements of $\mathcal{N}$ which differ only by a phase are regarded as equivalent. There is a natural projection map $\pi:\mathcal{N}\to\mathcal{R}$, which maps each vector to the ray on which it lies. The triplet $(\mathcal{N},\mathcal{R},\pi)$ forms a principal fiber bundle over the base space $\mathcal{R}$ [with structure group U(1)] and the parallel-transport law (1) defines a natural connection on this fiber bundle. A connection[9] is an assignment of a "horizontal subspace" in the tangent space of each point in $\mathcal{N}$. Horizontal vectors are those that satisfy (1). Given a curve in $\mathcal{R}$, one can lift this curve up to $\mathcal{N}$ so that its tangent vector is horizontal. However, the horizontal lift of a closed curve in $\mathcal{R}$ may be open in $\mathcal{N}$. This is referred to as holonomy of the connection and provides a geometric picture of Berry's phase.

Let $|\phi(s)\rangle$ be a curve in $\mathcal{N}$. Let $|u\rangle = (d/ds)|\phi(s)\rangle$ denote the tangent vector to this curve. Let us define[10]

$$A_s = \operatorname{Im}\langle\phi|u\rangle/\langle\phi|\phi\rangle. \tag{2}$$

Under transformations of the kind $|\phi(s)\rangle \to \exp[i \times\alpha(s)]|\phi(s)\rangle$ (referred to as gauge transformations), A_s transforms inhomogeneously,

$$A_s \to A_s + d\alpha/ds, \tag{3}$$

like the vector potential in electrodynamics. The parallel-transport law (1) states that A_s vanishes along the actual curve $|\phi(s)\rangle$ followed by the quantum system.

Let us first consider $|\psi(t)\rangle$, a solution of the Schrödinger equation which is cyclic, i.e., returns to the initial ray at some time τ. This defines a curve in $\mathcal{N}$. The "shadow" of this curve under projection map π is a closed curve in $\mathcal{R}$. Given the closed curve $r(s)$ in $\mathcal{R}$, let us ask for the curve $|\phi(s)\rangle$ in $\mathcal{N}$ traced out by the state vector (with the dynamical phase removed). Using (1), we find that the curve is determined by the condition $A_s = 0$ along the curve. Consider the integral

$$\gamma = \oint A_s\, ds \tag{4}$$

along the curve $|\phi(s)\rangle$ in $\mathcal{N}$ closed by the vertical curve joining $|\phi(\tau)\rangle$ to $|\phi(0)\rangle$. The segment $|\phi(s)\rangle$ represents the actual evolution of the system and along this, $A_s = 0$. The vertical contributes the phase difference between $|\phi(0)\rangle$ and $|\phi(\tau)\rangle$ and represents Berry's phase. However, the integral (4) is gauge invariant because of (3) and can be regarded as an integral on $\mathcal{R}$. With use of Stokes's theorem, γ can be expressed as

$$\gamma = \int_S F, \tag{5}$$

where S is any surface in $\mathcal{R}$ bounded by the closed curve $r(s)$ in $\mathcal{R}$ and F is the gauge-invariant ("field strength") two-form representing the curl of A. γ depends only on the geometric curve $r(s)$ and not on the rate at which it is transversed in time. This gives the formula for Berry's phase in a (possibly nonunitary) cyclic evolution of a quantum system.

In a general evolution, the state vector may not return to the initial ray. In order to handle this situation, we need a method of comparing states on different rays for phase. This is provided by the *Pancharatnan connection*: Let $|\phi_1\rangle$ and $|\phi_2\rangle$ by any two elements of $\mathcal{N}$ which are not orthogonal. Interference of these two states by superposition yields

$$\| |\phi_1\rangle + |\phi_2\rangle \|^2 = \langle\phi_1|\phi_1\rangle + \langle\phi_2|\phi_2\rangle + 2\operatorname{Re}\langle\phi_1|\phi_2\rangle.$$

The modulus of the resultant vector is clearly a maximum when $\langle\phi_1|\phi_2\rangle$ is real and positive. Under this condition, $|\phi_1\rangle$ and $|\phi_2\rangle$ are said to be "in phase." More generally, if one writes the complex number $\langle\phi_1|\phi_2\rangle$ in polar form, $\rho\exp i\beta$, $\rho>0$, then the phase difference between $|\phi_1\rangle$ and $|\phi_2\rangle$ is β. The Pancharatnam connection has a clear physical basis and is more general than the natural connection since it permits a comparison of *any* two (nonorthogonal) states for phase and not just neighboring ones. In the particular case where $|\phi_1\rangle$ and $|\phi_2\rangle$ are on neighboring rays, the Pancharatnam connection reduces to the natural connection.

We now go on to express the Pancharatnam phase difference in terms of the natural connection. In order to do this, we need to explore some geometrical properties of the ray space $\mathcal{R}$. We observe that $\mathcal{R}$ has a natural metric on it, which comes from the (positive definite) inner product $\langle\,|\,\rangle$ on $\mathcal{H}$. Since each point of $\mathcal{R}$ is an entire equivalence class, it is convenient to take representative elements from $\mathcal{N}$ and make sure our considerations are gauge invariant. Let $|\phi(s)\rangle$ be a curve in $\mathcal{N}$ and $|u\rangle$ its tangent vector. Under gauge transformations, $|u\rangle$ does not transform covariantly. But its projection orthogonal to the fiber,

$$|u'\rangle=|u\rangle-|\phi\rangle[\langle\phi|u\rangle-\langle u|\phi\rangle](2\langle\phi|\phi\rangle)^{-1}$$

is gauge covariant. $|u'\rangle$ is in fact the covariant derivative

$$|u'\rangle=(d/ds)|\phi(s)\rangle-iA_s|\phi(s)\rangle.$$

$\langle u'|u'\rangle$ is gauge invariant and can be used to define a metric on $\mathcal{R}$: $dl^2=\langle u'|u'\rangle ds^2$. dl^2 is the square of the distance between points $\pi(|\phi(s)\rangle)$ and $\pi(|\phi(s+ds)\rangle)$. This metric can also be expressed with use of the density matrix $\rho=|\psi\rangle\langle\psi|$, which contains information only about the ray and not the phase. Its form is

$$dl^2=(\mathrm{Tr}\rho)^{-1}[\mathrm{Tr}(d\rho\, d\rho)-\tfrac{1}{2}(\mathrm{Tr}d\rho)^2].$$

This metric then determines geodesics in $\mathcal{R}$. These can be found by variation of $\int\langle u'|u'\rangle dl$, where l is an affine parameter. This yields the geodesic equation

$$\frac{D^2}{dl^2}|\phi(l)\rangle=\frac{d}{dl}|u'\rangle-iA_s|u'\rangle=0. \tag{6}$$

Curves in $\mathcal{N}$ which satisfy this equation project down to geodesics in $\mathcal{R}$. Notice that (6) is gauge covariant and so the geodesic nature is a property of the "shadow" of the curve and not the curve itself.

The importance of geodesic curves in $\mathcal{R}$ stems from the fact that one can express the Pancharatnam phase difference as a line integral of A_s with use of the *geodesic rule*: Let $|\phi_1\rangle$ and $|\phi_2\rangle$ be any two (nonorthogonal) states in $\mathcal{N}$, with phase difference β according to the Pancharatnam connection. Let $|\phi(s)\rangle$ be any *geodesic* curve connecting $|\phi_1\rangle$ to $|\phi_2\rangle$: $|\phi(0)\rangle=|\phi_1\rangle$, $|\phi(1)\rangle=|\phi_2\rangle$. Then β is given by

$$\beta=\int A_s\,ds, \tag{7}$$

where A_s is given by (2).

Proof: Let $r(s)$ be a geodesic curve in $\mathcal{R}$ joining $\pi(|\phi_1\rangle)$ to $\pi(|\phi_2\rangle)$. Consider the horizontal lift $|\tilde\phi(s)\rangle$ of this curve, which starts from $|\phi_1\rangle$ [$|\tilde\phi(0)\rangle=|\phi_1\rangle$, $\tilde A_s=0$]. The geodesic equation (6) reduces to $(d^2/ds^2)|\tilde\phi(s)\rangle=0$, whose solution is a straight line in $\mathcal{N}$. Further, $|\tilde\phi(s)\rangle$ is "in phase" with $|\tilde\phi(0)\rangle$. To see this, define $g(s)=\mathrm{Im}\langle\tilde\phi(0)|\phi(s)\rangle$. Clearly, $g(0)=0$ and $\dot g(0)=0$ since $|\tilde\phi(s)\rangle$ is a horizontal curve. Now $\ddot g(s)$ can be worked out from the geodesic equation $\ddot g(s)=\mathrm{Im}\langle\tilde\phi(0)|(d^2/ds^2)|\phi(s)\rangle=0$, so that $g(s)$ is identically zero along the horizontal curve $|\tilde\phi(s)\rangle$; hence $\langle\phi_1|\tilde\phi(s)\rangle$ is real,[11] and so $|\tilde\phi(s)\rangle$ and $|\phi_1\rangle$ are "in phase." To prove (7), we simply perform a gauge transformation $|\phi(s)\rangle=\exp[i\alpha(s)]|\tilde\phi(s)\rangle$, where $\alpha(s)$ is chosen so that $\alpha(0)=0$, $\alpha(1)=\beta$. Then $|\phi(s)\rangle$ is still a geodesic curve (since the geodesic equation is gauge covariant) and connects $|\phi_1\rangle$ to $|\phi_2\rangle$. Using the fact that $\tilde A_s$ was zero along the horizontal curve $|\tilde\phi(s)\rangle$ and the behavior (3) of A_s under gauge transformations, we find that the right-hand side of (7) becomes $\int_0^1(d\alpha/ds)ds=\beta$ and (7) is verified. The geodesic rule[12] (7) is the main result of this paper.

We are now ready to show how Berry's phase also appears in a noncyclic evolution. Let the state vector $|\phi(t)\rangle$ (with the dynamical phase removed as always) evolve from $|\phi(0)\rangle$, initially, to $|\phi(\tau)\rangle$. If $|\phi(\tau)\rangle$ is not orthogonal to $|\phi(0)\rangle$, it is meaningful to ask, "What is the phase difference between them?" If we use the geodesic rule, this can be expressed as in (7). Now add to this integral the quantity $\int A_s\,ds$ integrated along the actual curve determined by the Schrödinger equation. Because of (1), this vanishes and the phase difference γ between $|\phi(\tau)\rangle$ and $|\phi(0)\rangle$ is expressed as an integral (4) where the contour C is given by the actual evolution $|\phi(t)\rangle$ from $|\phi(0)\rangle$ to $|\phi(\tau)\rangle$ and *back along any geodesic* curve joining $|\phi(\tau)\rangle$ to $|\phi(0)\rangle$. This expression for γ is gauge invariant and so can be regarded as defined on the base $\mathcal{R}$. So γ can be expressed as the integral (5) of the two-form F over a surface bounded by the closed curve $\pi(C)$. γ is clearly a gauge-invariant quantity and measureable. It has a purely geometric origin and depends only on the geometric path that the system traces in $\mathcal{R}$ and not on its rate of traversal.

Let us next consider a quantum system undergoing a nonunitary evolution, as happens, for example, when the system is subjected to measurements. According to the collapse postulate, the effect of the measurement on the system is described by the projection operator $\hat P=|\psi\rangle\langle\psi|$ onto the eigenstate corresponding to the eigenvalue (the outcome of the measurement) of the operator measured. Consider a system initially in the state $|\psi_1\rangle$ on which three successive measurements are made. If the effects of these measurements are to project the state of the system onto $|\psi_2\rangle$, then onto $|\psi_3\rangle$ and back onto $|\psi_1\rangle$, the final state of the system is given by $|\psi_1\rangle\langle\psi_1|\psi_3\rangle\langle\psi_3|\psi_2\rangle\langle\psi_2|\psi_1\rangle$. (We ignore the time evolution, i.e., set $\hat{H}=0$, so that we can concentrate on the effects due to projection that we are interested in.) The final and initial states have a well-defined phase difference, given by the phase of the complex number $\langle\psi_1|\psi_3\rangle\langle\psi_3|\psi_2\rangle\langle\psi_2|\psi_1\rangle$. Using the geodesic rule, we see that the phase is given by (5), where now the surface is bounded by the geodesic triangle connecting rays 1, 2, and 3. Thus Berry's phase also appears in systems subjected to quantum measurements. For a spin-$\frac{1}{2}$ (two-state) system, this formula for γ reduces to half the solid angle subtended at the center of the Poincaré sphere by the rays 1, 2, and 3. γ can be experimentally measured.

In summary, Berry's phase appears to be more general than the context in which it was discovered by Berry, i.e., for an adiabatic, cyclic, and unitary evolution. Our discussion uses a new ingredient—the Pancharatnam connection—and contains previous work as special cases. Since Berry's phase is being studied and applied in many different contexts, this generalization may be of interest.

It is a pleasure to thank S. Ramaseshan and Rajaram Nityananda for exciting our interest in Pancharatnam's work and S. Sridhar for useful discussions.

[1] M. V. Berry, Proc. Roy. Soc. London **392**, 45 (1984).

[2] B. Simon, Phys. Rev. Lett. **51**, 2167 (1983); M. S. Narasimhan and S. Ramanan, Am. J. Math. **83**, 563 (1961).

[3] Actually, Simon used the induced bundle on a finite-dimensional parameter space mapped injectively into the ray space. Y. Aharonov and J. Anandan, Phys. Rev. Lett. **58**, 1593 (1987), work directly with the ray space. See also Don Page, Phys. Rev. A **36**, 3479 (1987).

[4] Aharonov and Anandan, Ref. 3.

[5] S. Pancharatnam, Proc. Indian Acad. Sci. **A44**, 247 (1956), reprinted in *Collected Works of S. Pancharatnam* (Oxford Univ. Press, London, 1975).

[6] S. Ramaseshan and R. Nityananda, Curr. Sci. **55**, 1225 (1986).

[7] M. V. Berry, J. Mod. Opt. **34**, 1401 (1987). A proof of the geodesic rule is given here for two-state quantum systems.

[8] R. Bhandari and J. Samuel, Phys. Rev. Lett. **60**, 1211 (1988).

[9] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Wiley, New York, 1963).

[10] The connection can be given by our specifying a connection one-form A on $\mathcal{N}$. The contraction of this one-form with a tangent vector $|u\rangle$ is denoted by A_s in the text. Note that A is a linear functional on tangent vectors over the reals (not over $\mathbb{C}$).

[11] In fact, $\langle\tilde{\phi}(0)|\tilde{\phi}(s)\rangle$ is also positive if $|\tilde{\phi}(s)\rangle$ is the shortest geodesic connecting $|\tilde{\phi}(0)\rangle$ with $|\tilde{\phi}(1)\rangle$. Along this curve, $\langle\tilde{\phi}(0)|\tilde{\phi}(s)\rangle$ never vanishes and, since it was positive to start with, remains so.

[12] This rule breaks down for orthogonal states, since these are connected by a continuous infinity of geodesics, and this prescription does not give a definite answer.