PHYSICAL REVIEW LETTERS

VOLUME 60 | 18 JANUARY 1988 | NUMBER 3

Geometrical Phases from Global Gauge Invariance of Nonlinear Classical Field Theories

John C. Garrison
Lawrence Livermore National Laboratory, University of California, Livermore, California 94550

and

Raymond Y. Chiao
Department of Physics, University of California, Berkeley, California 94720
(Received 29 September 1987)

We show that the geometrical phases recently discovered in quantum mechanics also occur naturally in the theory of any classical complex multicomponent field satisfying nonlinear equations derived from a Lagrangean which is invariant under gauge transformations of the first kind. Some examples are the paraxial wave equation for nonlinear optics, and Ginzburg-Landau equations for complex order parameters in condensed-matter physics.

PACS numbers: 03.65.Bz, 11.15.Kc, 42.65.Jx

In the past few years, a great deal of work has been done on applying and generalizing the concept of geometrical phases. This notion was first introduced by Berry[1] in the context of the adiabatic approximation in quantum mechanics, but it was clearly recognized from the beginning that the ideas involved were applicable to any linear wave theory. Its first experimental confirmation was the observation by Tomita and Chiao[2] of geometrically induced optical activity. Its mathematical structure was elucidated by Simon.[3] The theory was generalized to the non-Abelian case by Wilczek and Zee,[4] and the restriction to the adiabatic approximation was removed by Aharonov and Anandan.[5] In all of this work the fact that the Schrödinger time-evolution operator is a linear isometry (i.e., is unitary) was used. Our objective in the present Letter is to show that linearity is not necessary. We do this by exhibiting a large class of nonlinear evolution equations which possess a geometrical phase. To this end, we study complex multicomponent classical fields. For the sake of clarity we will only consider fields $\psi=\mathrm{col}(\psi_1,\ldots,\psi_n)$, $\psi^\dagger=(\psi_1^*,\ldots,\psi_n^*)$, described by a Lagrangean density of the form

$$\mathcal{L}=i\left(\psi^\dagger\frac{\partial\psi}{\partial t}-\frac{\partial\psi^\dagger}{\partial t}\psi\right)+\mathcal{L}_1(\psi^\dagger,\psi,\nabla\psi^\dagger,\nabla\psi). \quad (1)$$

We assume that the Lagrangean (1) is real and that it is invariant under gauge transformations of the first kind, $\psi(x,t)\rightarrow\exp(i\alpha)\psi(x,t)$, with constant α. The assumption that the Lagrangean density is linear in the time derivatives is not necessary, and a more general discussion will be given elsewhere. It is the condition of *global* gauge invariance which is really essential to the following argument. The field equations derived from (1) can be written as

$$i\,\partial\psi/\partial t=G(\psi^\dagger,\psi,\nabla\psi^\dagger,\nabla\psi), \quad (2)$$

and the global gauge invariance of $\mathcal{L}$ implies that G satisfies the homogeneity condition

$$G(e^{-i\alpha}\psi^\dagger,e^{i\alpha}\psi,e^{-i\alpha}\nabla\psi^\dagger,e^{i\alpha}\nabla\psi)=e^{i\alpha}G(\psi^\dagger,\psi,\nabla\psi^\dagger,\nabla\psi). \quad (3)$$

In the case of quantum mechanics, $G=H\psi$, where H is the Hamiltonian. The global gauge invariance of the Lagrangean also implies, by Noether's theorem,[6] the existence of a conserved four-current density $(\rho,\mathbf{j})$. The assumption (1) for the Lagrangean gives the simple expression $\rho=\psi^\dagger\psi$ for the "charge density," but the explicit form of $\mathbf{j}$ cannot be obtained without specification of $\mathcal{L}_1$. The continuity equation for $(\rho,\mathbf{j})$ immediately yields a conservation law for the total charge Q:

$$\frac{dQ}{dt} \equiv \int dv \frac{\partial}{\partial t}\rho = -\int dv \nabla\cdot \mathbf{j} = -\int d\sigma\cdot \mathbf{j} = 0, \quad (4)$$

where the first integral extends over the configuration space appropriate to the problem and the last integral extends over the boundary. If the configuration space is finite, the last equality is a consequence of the boundary conditions on the field, but if the configuration space is unbounded, the finiteness of Q means that the field is square integrable (L_2). Since any physically reasonable field can be approximated by an L_2 field, we will restrict our attention to L_2 fields. The last equality then follows from the vanishing of the field at spatial infinity. The set of all such fields forms a Hilbert space $\mathcal{H}$ with Q as the norm, and the conservation law (4) shows that the evolution generated by G preserves this norm, i.e., the evolution is an isometry. The physical meaning of this fact depends on the theory in question; the interpretation of (4) as conservation of probability is peculiar to quantum mechanics. The vector properties of $\mathcal{H}$ play no particular role when G is nonlinear; in particular the inner product of two distinct vectors is not generally conserved.

The combination of the homogeneity property (3) and the isometry property (4) is sufficient to allow us to repeat *mutatis mutandis* the argument of Aharonov and Anandan demonstrating the existence of a geometrical phase associated with any cyclic evolution. For the sake of completeness, we give our version of their argument here. We first define a *cyclic* vector for the evolution law G as a field, $\psi(\mathbf{x},0)$, for which there is a time T and a phase Φ such that $\psi(\mathbf{x},T) = \exp(i\Phi)\psi(\mathbf{x},0)$. In quantum mechanics, the existence of cyclic vectors is assured; they are the eigenvectors of the unitary evolution operator $U(0,T)$, and the corresponding eigenvalue is $\exp(i\Phi)$. In the nonlinear case, no such general argument is available. A cyclic vector corresponds to a solution in which the spatial form is periodically replicated up to a phase, e.g., a soliton. The cyclic vector can be converted to a periodic form by the introduction of the modified field $\tilde\psi(\mathbf{x},t) = \exp[-if(t)]\psi(\mathbf{x},t)$, where $f(t)$ is any smooth function satisfying $f(T)-f(0)=\Phi$, so that $\tilde\psi(\mathbf{x},T) = \tilde\psi(\mathbf{x},0)$. Since G does not involve any time derivatives, the homogeneity property (3) yields the modified field equation

$$i\,\partial\tilde\psi/\partial t = \dot f\tilde\psi + G(\tilde\psi^\dagger,\tilde\psi,\nabla\tilde\psi^\dagger,\nabla\tilde\psi), \quad (5)$$

where $\dot f$ is the time derivative of f. Let (ψ,ξ) denote the inner product in $\mathcal{H}$; then forming the inner product of $\tilde\psi$ with (5) yields

$$(\tilde\psi, i\,\partial\tilde\psi/\partial t) = \dot f(\tilde\psi,\tilde\psi) + (\tilde\psi,\tilde G), \quad (6)$$

where $\tilde G = G(\tilde\psi,\dots)$. The definition of Q and the homogeneity property (3) show that $(\tilde\psi,\tilde\psi)=(\psi,\psi)=Q$, and $(\tilde\psi,\tilde G)=(\psi,G)$. Integration of (6) over the interval $(0,T)$ then gives the central result of this paper, $\Phi = \delta+\gamma$, where the *dynamical* phase δ and the *geometrical* phase γ are given by

$$\delta = -\frac{1}{Q}\int_0^T dt(\psi,G), \quad (7)$$

$$\gamma = \frac{1}{Q}\int_0^T dt\left(\tilde\psi, i\frac{\partial\tilde\psi}{\partial t}\right). \quad (8)$$

The phase δ is called dynamical because it depends explicitly on the form of G. To understand the geometrical nature of γ, we first remark that the cyclic evolution describes a curve, $t\to\psi(t)$, in $\mathcal{H}$ that begins and ends on the same ray, where a ray is the set of all constant multiples of some chosen vector. The set of all rays constitutes the projective space $\mathcal{P}(\mathcal{H})$, and the curve in $\mathcal{H}$ induces a closed curve C in $\mathcal{P}(\mathcal{H})$. The geometry of this projective space has been discussed by Zumino.[7] According to Aharonov and Anandan,[5] γ depends only on C. Their argument also works in our case. Consider an alternative field equation described by the function G' with solution ψ', period T', and phase shift Φ'. Since G and G' are both isometries and the evolutions start at the same ray, the value of the norm Q is the same for both. Now suppose that the two evolutions give the same curve C in $\mathcal{P}(\mathcal{H})$; then for any time $t\in(0,T)$ there is a time $t'\in(0,T')$ such that the vectors $\psi(t)$ and $\psi'(t')$ lie on the same ray. In fact, the two vectors can differ only by a phase factor, because of the equality of their norms. The same is obviously true for the modified fields $\tilde\psi(t)$ and $\tilde\psi'(t')$ so that there is a function $h(t')$ satisfying $\tilde\psi'(t')=\exp[ih(t')]\tilde\psi(t)$, and $h(T')=h(0)+2\pi n$, where n is an integer and the last equality follows from the periodicity of the modified fields. We next calculate γ' by using the primed version of (8) and the relation between the two fields to obtain

$$\gamma' = \frac{1}{Q}\int_0^{T'} dt'\left(\exp[ih(t')]\tilde\psi(t), i\frac{\partial}{\partial t'}\{\exp[ih(t')]\tilde\psi(t)\}\right) = \frac{1}{Q}\int_0^{T'} dt'\left\{-(\tilde\psi(t),\tilde\psi(t))\frac{dh}{dt'} + \left(\tilde\psi(t), i\frac{\partial\tilde\psi(t)}{\partial t'}\right)\right\}. \quad (9)$$

Since $(\tilde\psi,\tilde\psi)=(\psi,\psi)=Q$ is a constant, the periodicity of $h(t')$ [mod 2π] guarantees that the first term will contribute the phase $(-2\pi n)$, and the identity $dt'(\partial/\partial t') = dt(\partial/\partial t)$ shows that the second term is just γ. Thus $\gamma'=\gamma-2\pi n$ and the geometrical phase depends only on the curve C [mod 2π]. Thus it is the same for a large equivalence class of evolutions $\{G\}$ which generate the same closed curve C in $\mathcal{P}(\mathcal{H})$.

As an example, we consider the propagation of an intense, elliptically polarized light wave in a Kerr-active medium.[8]

The electric field can be expressed as $\mathbf{E}=\mathrm{Re}\{\mathcal{E}(\mathbf{x},t)\times\exp[i(kz-\omega t)]\}$, where $\mathcal{E}(\mathbf{x},t)$ is a slowly varying complex amplitude, which is transverse. In terms of the variables z and $\mathbf{x}_T=(x,y)$, the paraxial wave equation for $\mathcal{E}$ can be written as

$$i\frac{\partial\mathcal{E}}{\partial z}=-\frac{1}{2k}\nabla_T^2\mathcal{E}-G_0\{(\mathcal{E}\cdot\mathcal{E})\mathcal{E}^*+\tfrac{1}{3}(\mathcal{E}^*\cdot\mathcal{E})\mathcal{E}\}, \tag{10}$$

where $G_0=3kn_2/8n_0$, and n_0 and n_2 are respectively the linear and second-order indices of refraction. Thus the longitudinal variable z will play the role of time. (Note that since Maxwell's equations are first order in time, the paraxial approximation is unessential.) The paraxial wave equation (10) can be derived from a gauge-invariant Lagrangean satisfying (1), so that the complex vector $\mathcal{E}$ is an example of the complex n-component field of the general discussion. It is formally identical to a Ginzburg-Landau equation for a vector complex order parameter of an isotropic medium. In order to simplify the explicit calculation further we impose the plane-wave approximation, i.e., we ignore diffraction effects. In this limit there is no dependence on $\mathbf{x}_T$ and the Hilbert space $\mathcal{H}$ reduces to the space $\mathcal{C}^2$ of complex, two-component vectors. If an elliptically polarized wave with eccentricity ϵ is injected into the medium, the initial field can be written as $\mathcal{E}(0)=E_0\mathrm{col}(1,i\epsilon)$. The solution of (10) is then[8]

$$\mathcal{E}(z)=\exp(i\kappa z)R(\Gamma z)\mathcal{E}(0), \tag{11}$$

where $R(\eta)$ is the matrix representing a rotation around the z axis through the angle η and the constants κ and Γ are given by $\kappa=4G_0E_0^2(1+\epsilon^2)/3$, and $\Gamma=-2G_0E_0^2\epsilon$. Thus the field rotates at a uniform rate and acquires a uniformly increasing phase shift. This self-induced precession of the polarization ellipse has been observed.[9] The solution returns to the ray containing $\mathcal{E}(0)$ after a period $\Lambda=2\pi/|\Gamma|$, and the total phase shift Φ is given by $\Phi=2\pi\kappa/|\Gamma|$. The nonlinear function G is defined by (10), and (7) then gives the dynamical phase δ as

$$\delta=\frac{G_0}{Q}\int_0^\Lambda dz\{|\mathcal{E}\cdot\mathcal{E}|^2+\tfrac{1}{3}(\mathcal{E}^*\cdot\mathcal{E})^2\}$$
$$=\frac{4\pi(1-\epsilon^2+\epsilon^4)}{3|\epsilon|(1+\epsilon^2)}, \tag{12}$$

where we have used the values of Λ and Γ given above and also the expressions

$$|\mathcal{E}(z)\cdot\mathcal{E}(z)|^2=|\mathcal{E}(0)\cdot\mathcal{E}(0)|^2=E_0^4(1-\epsilon^2)^2,$$
$$[\mathcal{E}^*(z)\cdot\mathcal{E}(z)]^2=Q^2=E_0^4(1+\epsilon^2)^2,$$

which are easily derived from the explicit solution (11). We have shown above that the geometrical phase depends only on the closed curve C in the projective space $\mathcal{P}(\mathcal{H})$. Here $\mathcal{P}(\mathcal{H})=\mathcal{P}_1(\mathcal{C})$,[7] since $\mathcal{H}=\mathcal{C}^2$, and is the familiar Poincaré sphere used in the description of polarization states.[10] The natural coordinates for $\mathcal{P}(\mathcal{H})$ are the polar angle θ and the azimuthal angle ϕ, and the closed curve C corresponding to the solution (11) is simply a circle at constant latitude, $\theta=\theta_0$, where θ_0 is related to the eccentricity ϵ by $\cos\theta_0=-2\epsilon/(1+\epsilon^2)$. One period of the solution (11) corresponds to circumnavigation of the Poincaré sphere *twice* along this circle. Thus the polarization ellipse processes uniformly with no change in ellipticity. The coordinates on the Poincaré sphere are defined by the construction of a 2×2 Hermitean matrix M which has $\mathcal{E}(z)$ as an eigenvector with unit eigenvalue. An application of Stoke's theorem allows the line integral to be evaluated as

$$\gamma=-2\int_{\sigma(C)}V, \tag{13}$$

where V is a two-form related to the matrix M and $\sigma(C)$ is any surface bounded by C.[11] The factor of 2 in (13) comes from the double circumnavigation of the Poincaré sphere. This phase is closely related to that first derived by Pancharatnam[12] for a sequence of discrete polarization-changing linear optical elements which cause the polarization to undergo a cyclic evolution on the Poincaré sphere. However, it differs from the phase of Chiao and Wu[13] for spinning photons whose direction is changed cyclically. In our case, the result of this calculation is

$$\gamma=\mathrm{sgn}(\epsilon)\Omega(C)=\frac{4\pi|\epsilon|}{1+\epsilon^2}(\mathrm{mod}\,2\pi), \tag{14}$$

where $\Omega(C)$ is the solid angle subtended by C. The total phase shift is

$$\Phi=\frac{2\pi\kappa}{|\Gamma|}=\frac{4\pi}{3}\frac{1+\epsilon^2}{|\epsilon|}, \tag{15}$$

which is indeed the sum of (12) and (14).

In quantum mechanics, the dynamical phase δ is well understood: It is simply the time integral of the expectation value of the Hamiltonian. The meaning of δ is not so clear in the nonlinear case, and so it is of some interest to suggest an experiment which allows δ to be measured. For this purpose, let the elliptically polarized beam of our example be split, with one beam going through the nonlinear Kerr-active medium and the other through a linear medium in which there are suitably oriented static electric and magnetic fields. Berry has shown that this latter arrangement can be used to create a linear optical medium which is both gyrotropic and birefringent.[11] For a medium which is solely gyrotropic, i.e., optically active, the eigenpolarizations are circular-polarization states of opposite senses. For a medium which is solely birefringent, the eigenpolarizations are orthogonal linear-polarization states. When one combines both gyrotropy and birefringence in the same medium, the eigenpolarizations are elliptical-polarization states, in general. By slowly rotating the axis of the birefringence, e.g., by slowly rotating a static electric field applied to the Kerr medium around the direction of propagation of the light, one can force the semimajor axis of the polarization ellipse to corotate with the birefringence axis. Thus the orientation and strength of the fields can be chosen so that the polarization ellipse rotates in the same way in both arms, i.e., the curve traversed on the Poincaré spheres will be identical for both linear and nonlinear evolutions. According to the argument given above, the geometrical phase shift will be the same in both arms, but the dynamical phase shifts, which depend on the details of the dynamics, will differ. Thus recombination of the two beams will result in an interference pattern with a fringe shift depending on the difference in the dynamical phases. In other words, the dynamical phase for the nonlinear system can be measured relative to that of a properly chosen linear system. Since the total phase Φ is a physical observable, and the dynamical phase δ is a physical observable, one infers that their difference, the geometrical phase γ, is also a physical observable.

We conclude with some remarks on possible directions for future work. The restriction (1) on the form of the Lagrangean can be lifted at the expense of some formal complications in the theory, but the restriction to classical fields is a more difficult matter. Since some of the theories covered by our results are correspondence-principle limits of quantum field theories, e.g., quantum optics, the possibility of generalizing these considerations to quantized field theories is of great interest. A second important point arises from the fact that the line integral defining the geometrical phase γ is left invariant when the integrand is changed by the addition of the gradient [with respect to the coordinates of $\mathcal{P}(\mathcal{H})$] of any scalar function. Thus gauge invariance of the first kind for fields defined on ordinary space-time *induces* gauge invariance of the second kind in the projective space $\mathcal{P}(\mathcal{H})$. This constitutes a novel connection between the *global* gauge principle and the *local* gauge principle in different spaces. This remark holds for quantum mechanics as well as the classical field theories considered above. Therefore investigation of its significance should be of general interest.

We thank M. V. Berry for sending us his preprint on Pancharatnam's phase. This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore Laboratory under Contract No. W-7405-Eng-48, and National Science Foundation Grants No. ECS 86-13773 and No. ECS 86-20323.

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

[2]A. Tomita and R. Y. Chiao, Phys. Rev. Lett. **57**, 937 (1986).

[3]B. Simon, Phys. Rev. Lett. **51**, 2176 (1983).

[4]F. Wilczek and A. Zee, Phys. Rev. Lett. **52**, 2111 (1984).

[5]Y. Aharonov and J. Anandan, Phys. Rev. Lett. **58**, 1593 (1987).

[6]N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience, New York, 1959), p. 19ff.

[7]B. Zumino, LBL Report No. LBL-23056 UCB/PTH-87/13, 1987 (to be published).

[8]R. Y. Chiao and J. Godine, Phys. Rev. **185**, 430 (1969).

[9]P. D. Maker, R. W. Terhune, and C. M. Savage, Phys. Rev. Lett. **12**, 507 (1964).

[10]M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1975), p. 28ff.

[11]M. V. Berry, in *Fundamental Aspects of Quantum Theory,* edited by V. Gorini and A. Frigerio (Plenum, New York, 1986), cf. Eq. (A-5).

[12]S. Pancharatnam, Proc. Indian Acad. Sci., Sect. A **44**, 247 (1956), and in *Collected Works of S. Pancharatnam,* edited by G. W. Series (Oxford Univ. Press, Oxford, 1975); M. V. Berry, to be published.

[13]R. Y. Chiao and Y. S. Wu, Phys. Rev. Lett. **57**, 933 (1986).