Appearance of Gauge Structure in Simple Dynamical Systems

Frank Wilczek and A. Zee[a]
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106
(Received 9 April 1984)

Generalizing a construction of Berry and Simon, we show that non-Abelian gauge fields arise in the adiabatic development of simple quantum mechanical systems. Characteristics of the gauge fields are related to energy splittings, which may be observable in real systems. Similar phenomena are found for suitable classical systems.

PACS numbers: 03.65.Bz, 11.15.-q

Gauge fields, both Abelian and non-Abelian, figure prominently in modern theories of fundamental interactions. They also arise naturally in many geometrical contexts, and are central to much of modern mathematics. In this note we point out that gauge fields appear in a very natural way in ordinary quantum mechanical problems, whose initial formulation has no apparent relationship to gauge fields. We discuss some simple model problems in detail, and sketch in a general way how observable consequences of the gauge structures might be extracted for real physical systems. Finally, analogous behavior for classical oscillators is described.

It is, of course, potentially significant for models of elementary particles that gauge fields can arise "from nowhere," but we shall not attempt specific speculations along that line here.

Adiabatic problem.—Consider problems of the following general type: We are given a family of Hamiltonians $H(\vec{\lambda})$ depending continuously on parameters $\vec{\lambda}$, all of which have a set of n degenerate levels. By a simple renormalization of the energies, we can suppose that these levels are at $E=0$. Such degeneracies typically will occur when for each fixed value of the $\vec{\lambda}$ there is a symmetry; however, there need not be a single symmetry which is valid for all $\vec{\lambda}$. For example, the symmetry might be rotation around an axis whose direction is specified by $\vec{\lambda}$. More generally, the symmetry group H responsible for the degeneracy is embedded in a larger group G in a $\vec{\lambda}$-dependent way.

By the reasoning leading to the usual adiabatic theorem,[1] if the parameters are slowly varied from an initial value $\vec{\lambda}_i$ to some final value $\vec{\lambda}_f$ over a long time interval T, and the given space of degenerate levels does not cross other levels, then solutions of

$$H(\vec{\lambda}_i)\psi=0 \tag{1}$$

are mapped onto solutions of

$$H(\vec{\lambda}_f)\psi=0 \tag{2}$$

by solving the time-dependent Schrödinger equation

$$i\,\partial\psi/\partial t = H(\vec{\lambda}(t))\psi \tag{3}$$

with the boundary conditions $\vec{\lambda}(0)=\vec{\lambda}_i$, $\vec{\lambda}(T)=\vec{\lambda}_f$.

If $\vec{\lambda}_i=\vec{\lambda}_f$, so that the initial and final Hamiltonians are identical, then it becomes possible to formulate a more refined question: Given that the n degenerate levels are mapped back onto themselves by adiabatic development, is this mapping a nontrivial transformation? We find that it is, and that to describe such transformations gauge fields are the appropriate tool.

For $n=1$, a single level, the mapping is a simple phase multiplication, or for real wave functions, a sign. These situations, corresponding to U(1) or Z_2 gauge fields, were discussed by Berry[2] and by Simon.[3]

In the problem above, choose an arbitrary smooth set of bases $\psi_a(t)$ for the various spaces of degenerate levels, so that

$$H(\vec{\lambda}(t))\psi_a(t)=0. \tag{4}$$

Such a smooth choice can always be made locally, which is sufficient for our purposes. Let us write for the solutions of the Schrödinger problem (3), with the initial condition $\eta_a(0)=\psi_a(0)$,

$$\eta_a(t)=U_{ab}(t)\psi_b(t). \tag{5}$$

In writing (5) we have assumed the adiabatic limit, which can be justified to a sufficient degree of accuracy. Our task is to determine $U(t)$. We demand that the $\eta_a(t)$ remain normalized, so that

$$0=(\eta_b,\dot{\eta}_a)=(\eta_b,\dot{U}_{ac}\psi_c)+(\eta_b,U_{ac}\dot{\psi}_c) \tag{6}$$

which leads, in an evident notation, to the equation

$$(U^{-1}\dot{U})_{ba}=(\psi_b,\dot{\psi}_a)\equiv A_{ab}. \tag{7}$$

We will show that A, an anti-Hermitian matrix, plays the role of a gauge potential. Equation (7) is solved in terms of path-ordered integrals by

$$U(t)=P\exp\int_0^t A(\tau)\,d\tau. \tag{8}$$

It is remarkable that A depends only on the geometry of the space of degenerate levels. The specific form of A_{ab} computed from (7) depends, of course, upon the choice of bases $\psi_a(t)$. If one makes a different choice

$$\psi'(t)=\Omega(t)\psi(t), \tag{9}$$

then the A fields transform as

$$A'(t)=\dot{\Omega}\,\Omega^{-1}+\Omega A\,\Omega^{-1}, \tag{10}$$

i.e., as proper gauge potentials. As for ordinary gauge potentials, the path-ordered integral (8) around a closed loop transforms in a simple way under the gauge transformation (9), like (10) but without the inhomogeneous term. In particular, its eigenvalues are gauge invariant.

More generally, we can define the gauge potential A_μ everywhere on M, the space coordinatized by the parameters $\vec{\lambda}=\{\lambda^1,\ldots,\lambda^\mu,\ldots\}$. Explicitly,

$$A_\mu^{\mathrm{T}}=(\psi,\partial\psi/\partial\lambda^\mu). \tag{11}$$

The ordered integral

$$U(t)=P\exp\int_0^t A_\mu(\vec{\lambda}(t))\,d\lambda^\mu \tag{12a}$$

depends only on the path and not on its parametrization. In particular, for a closed path on M one obtains the Wilson loop

$$U=P\exp\oint A_\mu\,d\lambda^\mu. \tag{12b}$$

As a simple illustration of the preceding framework, consider the generic example of a system with $(n+1)$ levels, of which n levels are degenerate (at zero energy by normalization). Let the Hamiltonian be $H=R(t)H_0R^{-1}(t)$. Here H_0 denotes an $(n+1)$-dimensional matrix with the entries $(H_0)_{ij}=0$ unless $i=j=n+1$ and $R(t)=R(\theta(t))$ is the rotation

$$R=\exp(i\theta_n T_{n,n+1})\cdots\exp(i\theta_2 T_{2,n+1})\exp(i\theta_1 T_{1,n+1}).$$

The embedding of the relevant symmetry group $\mathrm{SO}(n)$ in $\mathrm{SO}(n+1)$ varies with time. The parameter space M is, of course, the coset space $\mathrm{SO}(n+1)/\mathrm{SO}(n)=S^n$. A simple evaluation of Eq. (11) gives the non-Abelian gauge potential and field strength

$$A_\mu=\pi R^{-1}(\partial R/\partial\theta^\mu)\pi, \tag{13a}$$

$$-F_{\mu\nu}=\pi R^{-1}\frac{\partial R}{\partial\theta^\mu}(1-\pi)R^{-1}\frac{\partial R}{\partial\theta^\mu}\pi-(\mu\leftrightarrow\nu), \tag{13b}$$

where π represents projection onto the first n components. Note that left and right projection π of the pure gauge $R^{-1}\partial_\mu R$ gives us a nontrivial $\mathrm{SO}(n)$ gauge field. For $n=3$ we find explicitly

$$A_1=0,\quad A_2=\sin\theta_1 T_{12},\qquad A_3=\sin\theta_1\cos\theta_2 T_{13}+\sin\theta_2 T_{23} \tag{14}$$

leading to the field strengths

$$F_{12}=\cos\theta_1 T_{12};\quad F_{12}^0=T_{12},$$
$$F_{23}=\cos^2\theta_1\cos\theta_2 T_{23};\quad F_{23}^0=T_{23}, \tag{15}$$
$$F_{13}=\cos\theta_1\cos\theta_2 T_{13};\quad F_{13}^0=T_{13}.$$

Since the metric structure

$$ds^2=d\theta_1^2+\cos^2\theta_1\,d\theta_2^2+\cos^2\theta_1\cos^2\theta_2\,d\theta_3^2$$

is diagonal we can define

$$F_{ij}^0\equiv-(g^{ii}g^{jj})^{1/2}F_{ij}, \tag{16}$$

the Cartesian tensor. As might have been anticipated from the simplicity of the starting Hamiltonians (13), the gauge structure is quite simple. In fact, the rotations induced by the ordered integrals (8) amount to parallel transport of tangent vectors to the sphere, with the obvious identifications. Nevertheless, this very fact shows that the example involves truly non-Abelian gauge structure.

The example can obviously be generalized. With the Hamiltonian suitably parametrized on the homogeneous space G/H, we can evaluate the gauge field at the "north pole," thus obtaining from Eq. (13) a simple expression in terms of the structure constant f of G:

$$F_{\mu\nu}=f_{\mu\nu a}\pi\lambda_a\pi. \tag{17}$$

(Here λ_a denote the generators of G; those generators not in H are labeled by a Greek index.) In particular, for the potentially physical example of a Hamiltonian with three levels, two of which are degenerate (at zero energy), and parametrized on $\mathrm{SU}(3)/\mathrm{SU}(2)\otimes\mathrm{U}(1)=CP_2$ we have, in the standard SU(3) notation, $F_{45}=(1+\tau_3)$, $F_{67}=(1-\tau_3)$ $F_{47}=-F_{56}=\tau_1$, and $F_{46}=F_{57}=\tau_2$.

Stationary states.—In many real systems there are fast and slow degrees of freedom, and then one may estimate the effect of the slow variables on the fast ones in the adiabatic approximation. An important familiar example is the Born-Oppenheimer

treatment of molecules, and we shall use the terminology of this example for definiteness, although we have not investigated any possible applications in realistic detail. In this context, an important problem is to find the stationary states, which in general requires, of course, that we treat the slow variables quantum mechanically. In order that our previous discussion, where these variables were, of course, treated classically, apply fairly directly, let us first discuss this in the correspondence or quasiclassical limit.

Suppose that the nuclei can be described quasiclassically as undergoing a motion with period $2\pi/\omega$; i.e., let them be in a quantum state of the type

$$|s\rangle = \int_0^{2\pi/\omega} e^{-ip\omega\tau} |\vec{\lambda}(\tau)\rangle \, d\tau, \tag{18}$$

where the label $\vec{\lambda}$ is periodic with period $2\pi/\omega$ and p is an integer. States labeled by different $\vec{\lambda}$ have negligible overlap, and $e^{-iHt}|\vec{\lambda}(t_0)\rangle = |\vec{\lambda}(t_0+t)\rangle$ to an adequate approximation, where H is the Hamiltonian for the nuclear motion calculated as if the electrons followed instantaneously. Within the stated approximations $|s\rangle$ is a stationary state, $e^{-iHt}|s\rangle = e^{-ip\omega t}|s\rangle$.

Let us suppose that for any fixed $\vec{\lambda}$ there is a symmetry guaranteeing degeneracy of two electronic levels, but that the symmetry cannot be defined independent of $\vec{\lambda}$, as in the previous discussion. As we have seen, the development of these levels in response to the motion of $\vec{\lambda}$ can involve nontrivial phases and mixings after a complete period of the motion of $\vec{\lambda}$. We can diagonalize the mixing matrix and thus find states which are multiplied by phases $\exp(i\gamma_1)$, $\exp(i\gamma_2)$, after a period. For these states, we then find the quantization condition altered to read

$$\omega' T + \gamma_i = 2\pi p$$

or

$$\omega' = p\omega - \gamma_i\omega/2\pi. \tag{19}$$

In accordance with the correspondence principle, Eq. (19) represents the energy splittings for small p. In a more general framework one would construct an effective Lagrangian for the slow variables and treat this fully quantum mechanically. The phase we have found adiabatically represents a term in the Lagrangian *linear in* $\dot{\theta}$, where θ is the coordinate of nuclear rotation. Such a term contributes nothing to the classical equations of motion (in line with its origin as a pure phase) but does change the quantization condition. The rotational energy is altered from $n^2/2I$, n = integer, to $E_n = (n-\gamma/2\pi)^2/2I$. This agrees with Eq. (19) for large quantum numbers, viz.,

$$E_{n+p_1}(\gamma_1) - E_{n+p_2}(\gamma_2) \simeq \frac{n}{I}\left(p_1 - p_2 - \frac{\gamma_1}{2\pi} + \frac{\gamma_2}{2\pi}\right) = \omega(p_1-p_2) - \omega\left(\frac{\gamma_1}{2\pi} - \frac{\gamma_2}{2\pi}\right). \tag{19a}$$

An example of this general framework is the phenomenon of Λ doubling.[4]

Mechanical analogs.—Simple mechanical analogs exist for many of the systems discussed above. The point is that the mechanical equation $\ddot{x} = A\dot{x}$, for A anti-Hermitian, becomes the Schrödinger equation for $\psi = \dot{x}$.

In our study of mechanical analogs, we have uncovered other phenomena which may be interpreted as noncompact gauge fields. Consider a planar harmonic oscillator in a magnetic field perpendicular to the plane:

$$\ddot{\vec{x}} + B(t)\dot{\vec{x}} + \mu(t)\vec{x} = 0, \tag{20}$$

$$B(t) = \gamma(t)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{21}$$

With γ and μ slowly varying, we have the approximate solution

$$\vec{x}(t) = \begin{pmatrix} 1 \\ i \end{pmatrix} a(t) \exp i\left\{\int_{t_0}^{t} \omega(\tau)\, d\tau\right\}, \tag{22}$$

with

$$\omega^2 + \gamma\omega - \mu = 0, \tag{23}$$

$$\dot{a}/a = -\dot{\omega}/(2\omega+\gamma) \tag{24}$$

(the induced electric field has been ignored).

The second equation indicates that in response to an infinitely slow cyclic variation of the parameters in the (γ,μ) plane the amplitude a gets multiplied by a nontrivial path-dependent factor. Interestingly, amplification occurs despite the arbitrary slow variation and so the relevant gauge group is $GL(1,R)$, a noncompact group. More explicitly, the factor is given by

$$\exp\left(\oint da/a\right) = \exp\left(\oint A\right) = \exp\left(\int F\right),$$

with

$$A=\left(\pm\frac{1}{2(\gamma^2+4\mu)^{1/2}}-\frac{\gamma}{2(\gamma^2+4\mu)}\right)d\gamma-\frac{1}{\gamma^2+4\mu}d\mu \tag{25}$$

and the field strength

$$F=\mp(\gamma^2+4\mu)^{-3/2}. \tag{26}$$

If either γ or μ is constant the area enclosed by the closed path in the (γ,μ) plane collapses and there is no amplification. Also, note that if $\mu=\text{const}$ our system conserves $\dot{\vec{x}}^2+\mu\vec{x}^2=\dot{a}^2+a^2(\omega^2+\mu)$. In the adiabatic limit, $a^2(\omega^2+\mu)=\text{const}$, in contrast to the standard adiabatic theorem $a^2(\omega+\frac{1}{2}\gamma)=\text{const}$ for the case $\gamma=\text{const}$.[5]

This material is based upon research supported in part by the National Science Foundation under Grant No. PHY77-27084, supplemented by funds from the National Aeronautics and Space Administration. We thank W. Kohn and R. Schrieffer for helpful comments.

[a]On leave from University of Washington, Seattle, Wash. 98195. Present address: Institute for Advanced Study, Princeton, N. J. 08540.

[1]M. Born and V. Fock, Z. Phys. **51**, 165 (1928); L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955), p. 290.

[2]M. V. Berry, to be published.

[3]B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).

[4]Λ doubling is discussed somewhat in the spirit of this note by G. Wick, Phys. Rev. **73**, 51 (1948).

[5]Taking the induced electric field into account abolishes the amplification in this particular example. Nevertheless, the phenomenon discussed does arise for differential equations describing physical systems, in particular, feedback circuits with effectively imaginary resistance or variable-length pendula acted upon by Coriolis forces.