Gauge fields, quantum interference, and holonomy transformations

J. Anandan*
Max-Planck-Institut für Physik und Astrophysik, Werner-Heisenberg-Institut für Physik, D-8000 Munich 40, Federal Republic of Germany
and Los Alamos National Laboratory, Los Alamos, New Mexico 87545
(Received 23 April 1985)

The question of which quantities are measured when a gauge field is experimentally detected is investigated. The Josephson effect is considered in connection with this question and a generalization of this effect for non-Abelian gauge theories is obtained. It is shown that the gauge field in an n-dimensional manifold can be reconstructed from the holonomy transformations (parallel-transport operators around loops) for an n-dimensional set of loops. An application is given to Minkowski space-time compactified by the addition of null infinity. An equivalence principle for gauge fields is also formulated.

I. INTRODUCTION: LOCALITY AND GAUGE FIELDS

In order to do physics, it is necessary to communicate information. In the present space-time description of physics, this information is typically a vector at a point in space-time which must then be compared with a similar vector at another point in space-time. This vector may, in classical physics, be the energy momentum or angular momentum of a particle and, in quantum physics, the value of the quantum field or wave function at a given space-time point, which may belong to the spin vector space, color vector space, etc. Therefore, any local physical theory must confront the problem of comparing vectors at different points.

A natural way of making this comparison is to introduce a connection. By this is meant a prescription for parallel transporting a vector v at a point P to another v' at a point Q along a path C joining P and Q. A comparison between v' and a similar vector u at Q may then be regarded as a comparison between v and u with respect to the path C. If v is a tangent vector or a spinor vector then the connection is the usual gravitational connection. On the other hand, if v belongs to an internal vector space such as the color space then the connection is called a gauge field.[1] But the physical measurements made in an actual experiment are local measurements. So, v and v' have to be determined by separate measurements and this knowledge therefore cannot give information about the connection. Hence, to determine how v is parallel transported along C, it is necessary to compare v' with its parallel transport v'_1 along another path C_1, joining P and Q. Since the vectors v' and v'_1 are at the same point Q, they can be compared by means of a local measurement and this will give information about the connection. But v'_1 can then be obtained from v' by parallel transport along the *closed curve* formed by C and C_1. If this parallel transport is the identity for all closed curves then the curvature of the connection is said to be zero. If not, the connection or gauge field is nontrivial. To summarize, the *locality of the laws of physics* first suggests the introduction of a connection or gauge field and second requires that all the physically meaningful information in the connection is contained in parallel transport along closed curves.

The important connection between the locality of the laws of physics and gauge fields can also be seen from the following consideration. *A priori*, it would have been possible to define gauge fields on momentum space as much as on space-time. The position operator is a derivative operator in momentum space and can be written as a covariant derivative with respect to a connection on momentum space. This may seem especially natural for the harmonic oscillator, whose Hamiltonian contains momentum and position symmetrically. However, such a "momentum-space gauge field" would introduce nonlocal correlations in space-time which, so far, have not been experimentally observed.

An example of communication of information in space-time is given by a particle beam described by a wave function ψ. If P and Q are two points on this beam then *a priori* we have no way of comparing $\psi(P)$ and $\psi(Q)$. However, if the beam is coherently split at P and recombined at Q then the result of the interference contains information about how the wave is propagated along each of the beams. Suppose that we turn on an electromagnetic field, which is the simplest example of a gauge field. Then it is well known that the shifting of the interference fringes is determined by the phase factor

$$F_\gamma=\exp\left[\frac{iq}{\hbar c}\oint_\gamma A_\mu dx^\mu\right], \tag{1.1}$$

where q is the charge of the particle, A_μ is the vector potential, and γ is a closed curve going around the interfering beams. This F_γ is the operator acting on the wave function when it is parallel transported around γ with respect to the electromagnetic connection. Its importance is seen from the fact that, as pointed out by Aharonov and Bohm,[2] it has physical consequences even when the field strength vanishes along the beams.

For this reason, Wu and Yang[3] pointed out that the complete description of electromagnetism is provided by phase factors of the form (1.1). This remark has been subsequently generalized[4] to all gauge fields when it was shown that the "holonomy transformation"

$$F_\gamma = P\left[\exp\left[-ig\oint_\gamma A^j_\mu T^j dx^\mu\right]\right] \tag{1.2}$$

determines the interference pattern for two beams interfering around γ which is a piecewise-differentiable curve that begins and ends at a point 0 (a loop at 0) in the interference region. Also, a gauge potential A^j_μ can be reconstructed from transformations of the form (1.2), with 0 being any fixed point, and it is then unique up to gauge transformation.[5,6] Here, T^i generates the gauge group and P denotes path ordering. When the holonomy group is compact or semisimple, the traces of the operators (1.2) in the fundamental representation are sufficient to reconstruct the gauge potential up to gauge transformations.[6] Using the traces has the advantage that they are gauge invariant, whereas F_γ is gauge covariant.

What is directly observed in the interference of two beams is really not (1.1) or (1.2), but rather a Hermitian operator H_γ. For a compact gauge group, $H_\gamma = \frac{1}{2}(e^{i\phi}F_\gamma + e^{-i\phi}F_\gamma^\dagger)$, where ϕ is the phase difference between the interfering beams in the absence of the gauge field. Then, F_γ can be regarded as unitary and the intensity in the interference region is determined by[4] $(\psi + e^{i\phi}F_\gamma\psi)^\dagger(\psi + e^{i\phi}F_\gamma\psi) = 2\psi^\dagger\psi(1+\psi^\dagger H_\gamma\psi/\psi^\dagger\psi)$, where ψ is the value of the wave function for one of the interfering beams. The space-time variation of the intensity determines the "interference term" $\psi^\dagger H_\gamma \psi/\psi^\dagger\psi$. By doing the experiment for different ψ's, the Hermitian operator H_γ can therefore be determined.

To determine F_γ from H_γ, note first that a basis can be chosen in which H_γ and F_γ are simultaneously diagonal. If $F_\gamma = \mathrm{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n})$ then $H_\gamma = \mathrm{diag}(\cos(\theta_1 + \phi), \cos(\theta_2+\phi), \ldots, \cos(\theta_n+\phi))$ in this basis. Hence, $\theta_r + \phi$ $(r = 1,2,\ldots,n)$ can be determined from H_γ only up to a sign and modulo 2π by this experiment and even then it is necessary to separate θ_r from ϕ. But the resulting ambiguity in F_γ can be eliminated by noting that the sign of the known phase ϕ is fixed by convention and that it is proportional to the mass for a massive particle whereas θ_r is independent of the mass or that θ_r is a type-II phase shift as opposed to ϕ which is a type-I phase shift.[4] This determines θ_r modulo 2π and, hence, F_γ is uniquely determined. Also, as will be seen later, the Josephson effect, when generalized to an arbitrary gauge theory, enables $\sin(\theta_r+\phi)$, $r=1,\ldots,n$, to be determined, in principle, which would be an alternative way of eliminating the above-mentioned ambiguity in determining F_γ from H_γ.

But when a charged matter field ψ is coupled to the electromagnetic field, we also have *gauge-invariant* quantities of the form

$$F_C = \psi^*(x_1^\mu)\exp\left[-i\frac{q}{\hbar c}\int_{C\,x_1}^{x_2} A_\mu dx^\mu\right]\psi(x_2^\mu)\ , \tag{1.3}$$

where the integral is along a path C that joins two points whose coordinates are x_1^μ and x_2^μ. A gauge transformation on A_μ is compensated by the corresponding gauge transformation on ψ so that (1.3) is gauge invariant. This raises the questions as to whether quantities of the form (1.3) are experimentally observable and if so whether it violates the statement of Wu and Yang that quantities of the form (1.1) associated with closed curves contain *all* the *observable* information in the electromagnetic field. In Sec. II we shall answer the first of these questions in the affirmative and the second in the negative by considering the Josephson effect for superconductors. This leads to a generalization of the Josephson effect for non-Abelian gauge theories and possible applications are mentioned. For non-Abelian gauge fields, however, it is pointed out that a measurement of the gauge field, in general, determines operators that are more general than (1.2) [see (2.8)].

Nevertheless, the operators (1.2), for all loops at a point, are sufficient for reconstructing the gauge potential (up to gauge transformations) because of the theorem[6] mentioned above. But they are not all necessary. Indeed, it is shown in Sec. III that in an n-dimensional manifold M, there exists an n-dimensional subset of the infinite-dimensional set of loops at a fixed point such that the holonomy transformations (1.2) for this subset of curves are necessary and sufficient for reconstructing the gauge field. An interesting application of this result is given to Minkowski space-time made compact by adding conformal infinity.

In Sec. IV we make use of this result to formulate a "principle of equivalence" for gauge fields, which is analogous to Trautman's formulation of the principle of equivalence for the gravitational field. It is therefore hoped that this principle would be of help in providing insight into the problem of unifying gravity and gauge fields.

II. THE JOSEPHSON EFFECT AND GAUGE FIELDS

Superconductivity is due to part of the conduction electrons forming pairs called Cooper pairs, which can be regarded as being in a pure quantum-mechanical state that is spread out over the entire superconductor, normally of macroscopic dimensions. The order parameter $\psi(x^\mu)$ may be regarded approximately as the wave function of these Cooper pairs. So, in a Josephson junction, by which is meant a thin normal conductor or an insulator separating two superconductors, the wave function tunnels from both sides of the junction to interfere and give a current. Hence, this current should be periodic in the *gauge-invariant* phase difference

$$\Delta\phi^* = \int_{C\,x_1}^{x_2}\left[\nabla\phi - \frac{2e}{\hbar c}\mathbf{A}\right]\cdot d\mathbf{r}\ , \tag{2.1}$$

where ϕ is the phase of ψ, $2e$ is the charge of the Cooper pair, and the integral is over the shortest path C across the junction in the frame of the junction. Therefore the lowest-order component of the current density is

$$j = j_0 \sin\Delta\phi^*\ , \tag{2.2}$$

which was first predicted by Josephson.[7]

We now rewrite (2.1) as

$$\exp(i\Delta\phi^*)=\frac{\psi^*(x_1^\mu)\exp\left[-\frac{i2e}{\hbar c}\int_{x_2}^{x_1}A_\mu dx^\mu\right]\psi(x_2^\mu)}{|\psi(x_2^\mu)|\,|\psi(x_1^\mu)|}. \tag{2.3}$$

It follows immediately that quantities of the form (1.3) are observable by means of the Josephson effect. But this does not violate the statement mentioned in the previous section that the phase factors (1.1) (which are special cases of (2.3) corresponding to x_1 coinciding with x_2) provide a complete description of the electromagnetic field. This is because, in order to measure the electromagnetic field by means of the Josephson effect, or any other effect that depends on (2.3), it is necessary, in the experiment being performed, to fix physically the relationship between $\psi(x_1)$ and $\psi(x_2)$. This can be done only by parallel transporting $\psi(x_1)$ and $\psi(x_2)$ along curves C_1 and C_2 to a common point x and comparing the results

$$\psi_1=\exp\left[-\frac{i2e}{\hbar c}\int_{C_1 x_1}^{x}A_\mu dx^\mu\right]\psi(x_1^\mu)$$

and

$$\psi_2=\exp\left[-\frac{i2e}{\hbar c}\int_{C_2 x_2}^{x}A_\mu dx^\mu\right]\psi(x_2^\mu).$$

Suppose $\psi_2=\chi\psi_1$, where χ is a complex number to be determined by the experimental conditions. Then, from (2.3),

$$e^{i\Delta\phi^*}=(\chi^*/|\chi|)\exp\left[-\frac{i2e}{\hbar c}\oint_\gamma A_\mu dx^\mu\right],$$

where γ is the loop at x formed from C_1, C, and C_2. Hence, the Josephson current (2.2) depends on the phase factor (1.1) associated with a closed curve and the phase of χ which depends on the experiment being performed.

An example of such an experiment is a superconducting ring interrupted by a Josephson junction and enclosing a magnetic flux. Since, in the interior of the superconductor there is no electromagnetic field and no current, ψ is covariantly constant in directions normal to the four-velocity t^μ of the superconductor inside the ring except inside the junction. Therefore, if the superconductor is nonrotating

$$\psi(x_2)=\exp\left[-\frac{i2e}{\hbar c}\int_{C_1 x_1}^{x_2}A_\mu dx^\mu\right]\psi(x_1),$$

where C_1 is the longer spacelike path perpendicular to t^μ, through the interior of the superconductor, joining the points x_1,x_2 on the two sides of the Josephson junction. Therefore, this is a special case of the above analysis corresponding to $\chi=1$ and $x_2=x$. Hence,

$$\exp(i\Delta\phi^*)=\exp\left[-\frac{i2e}{\hbar c}\oint_\gamma A_\mu dx^\mu\right]=\exp\left[-\frac{i2e}{\hbar c}F\right],$$

where F is the magnetic flux enclosed by the ring, and a Josephson current, given by (2.2), will flow through the ring. It should be noted that the γ here is spacelike, unlike the γ for which (1.1) is determined by the interference of two coherent beams, which is made up of timelike curves. But if the superconducting ring has an angular velocity Ω, then we cannot find a closed curve around the ring which is perpendicular to t^μ everywhere, and it follows that[8] $\chi=\exp(i2m\Omega A/\hbar)$ in the nonrelativistic limit, where $2m$ is the mass of the Cooper pair and A is the area enclosed by the ring.

The above discussion of the Josephson effect suggests a generalization of this effect for non-Abelian gauge fields. Suppose that ψ is a matter field coupled to a gauge field A_μ^i. Consider a "Josephson junction" by which we mean a potential barrier that enables ψ from both sides to tunnel through and interfere. If x_1 and x_2 are the extreme points of a shortest path C through the junction, then it is shown in the Appendix that the component along the path C of the current density at any point x on this path is

$$j^k(x)=i\,\mathcal{J}_0[\psi^\dagger(x_2)U^k(x)\psi(x_1)-\psi^\dagger(x_1)U^k(x)^\dagger\psi(x_2)], \tag{2.4}$$

where

$$U^k(x)=P\left[\exp\left[-ig\int_x^{x_2}A_\mu^jT^jdx^\mu\right]\right]\times T^kP\left[\exp\left[-ig\int_{x_1}^{x}A_\mu^jT^jdx^\mu\right]\right], \tag{2.5}$$

and T^j generate the relevant representation of the Lie algebra of the gauge group. For the superconducting case, the gauge field is the electromagnetic field and gT^j are replaced by $2e$. Then (2.4) and (2.5) give (2.2) with (2.1), for this special case. Thus (2.4) is a generalization of the Josephson equation (2.2). It must also be supplemented by the equation of motion for the ψ field.

There are two possible applications of this effect in the very big and very small scales. According to the inflationary-universe scenarios,[9] based on grand unified theories, there was a spontaneous symmetry breaking in the early Universe with the formation of bubbles in which the Higgs field had different expectation values, in general. The Higgs field is like the order parameter in a superconducting system which is believed to obey the nonlinear Ginzburg-Landau equation. The energy of the space between the bubbles is that of the original "false vacuum" which is greater than the energy of the true vacuum inside any of the bubbles. So, when two such bubbles come sufficiently close to each other, by which we mean that the shortest distance between them is $\lesssim\hbar/\sqrt{2mV}$ where V is the height of the potential barrier (the difference between the energies of the false and true vacua), the Higgs field ψ should tunnel through this barrier producing the current density predicted by (2.4). It follows that the total current (integrated current density) will depend on the gauge field present. Another possible application is in nuclei. If quarks pair to form the analogs of Cooper pairs, then we may expect tunneling phenomena, similar to the superconducting type, in nuclei. Also, if absolute confinement is not assumed in the bag model of hadrons, then there would be tunneling between two bags giving rise to a generalized Josephson current.

It is known that the magnetic flux enclosed by a superconducting ring is quantized. We consider now the analog of this for a non-Abelian gauge theory. Suppose that we have an n-dimensional representation $D(G)$ of a gauge group G. Let $\psi_1,\psi_2,\ldots,\psi_m$ $(m\le n)$ be a set of fields on which $D(G)$ acts such that the only element of $D(G)$ that leaves them invariant is the identity apart from points where one or more of these fields vanish. For the ground state, far away from such points, there would exist a hypersurface orthogonal timelike vector field t^μ such that $\psi_1,\psi_2,\ldots,\psi_m$ are covariantly constant in the directions perpendicular to t^μ. Then if γ is a closed curve, the points of which are at such far distances, and which is orthogonal to t^μ everywhere, then

$$P\left[\exp\left[-ig\oint_\gamma A^j_\mu T^j dx^\mu\right]\right]=1\ . \tag{2.6}$$

Hence, γ can enclose a vortex line containing a gauge field flux which is quantized according to (2.6). A particular case of this, for an SO(3) gauge theory, for which $n=3$ and $m=2$, has been given by Nielsen and Olsen.[10] If the gravitational field is non-negligible, and ψ has intrinsic spin, then we must generalize (2.6) by adding the gravitational connection to the gauge field connection in (2.6). The bubbles in the early Universe, mentioned above, are usually assumed to be simply connected. But it is *a priori* possible for them to be nonsimply connected. For instance, it can be in the form of a ring enclosing a gauge field which is subject to the condition (2.6).

For completeness, we also consider the non-Abelian analog of a nonrotating superconducting ring with a Josephson junction. Then in (2.4),

$$\psi(x_2)=P\left[\exp\left[-ig\int_{C_1 x_1}^{x_2} A^j_\mu T^j dx^\mu\right]\right]\psi(x_1)\ ,$$

where C_1 is a spacelike path through the ring. Hence, on defining

$$\psi(x)=P\left[\exp\left[-ig\int_{x_1}^{x} A^j_\mu T^j dx^\mu\right]\right]\psi(x_1)\ ,$$

$$j^k(x)=i \oint_0\left\{\psi^\dagger(x)P\left[\exp\left[-ig\oint_\gamma A^j_\mu T^j dx^\mu\right]\right]T^j\psi(x) -\psi^\dagger(x)T^jP\left[\exp\left[-ig\oint_\gamma A^j_\mu T^j dx^\mu\right]\right]\psi(x)\right\}\ , \tag{2.7}$$

where γ is the loop at x formed from C and C_1.

Thus, in this special case, the current (2.7) measures

$$P\left[\exp\left[-ig\oint_\gamma A^j_\mu T^j dx^\mu\right]\right]$$

associated with a closed curve. Another case of an experiment in which this quantity is measured, discussed in Sec. I, is the interference of two coherent beams in the presence of a gauge field.[4] In general, however, the above considerations suggest that an experiment to measure a gauge field would measure an operator of the form

$$\chi_1 P\left[\exp\left[-ig\int_{x_2}^{x_1} A^j_\mu T^j dx^\mu\right]\right] \times\chi_2 P\left[\exp\left[-ig\int_{x_3}^{x_2} A^j_\mu T^j dx^\mu\right]\right]\cdots \times\chi_n P\left[\exp\left[-ig\int_{x_1}^{x_n} A^j_\mu T^j dx^\mu\right]\right]\chi_{n+1}+\text{H.c.}\ , \tag{2.8}$$

where the operators $\chi_1,\ldots,\chi_{n+1}$ are determined by the experimental conditions. These need not commute with the parallel-transport operators which depend on the gauge potential. This is an essential difference between the non-Abelian and the Abelian gauge fields. Nevertheless, (2.8) is also associated with a closed curve that begins and ends at x_1, similar to the Abelian case. Since the apparatus responsible for the operators $\chi_1,\ldots,\chi_{n+1}$ is made up of matter fields that interact with the same gauge field A^j_μ, it is reasonable to suppose that $\chi_1,\ldots,\chi_{n+1}$ transform covariantly under a gauge transformation (i.e., undergo similarity gauge transformations). Then (2.8) also would transform covariantly. Otherwise gauge invariance would be broken. As seen above, it is possible to have special gedanken experiments in which $\chi_1,\ldots,\chi_{n+1}$ are each equal to the identity. Then (2.8) is

$$P\left[\exp\left[-ig\oint_\gamma A^j_\mu T^j dx^\mu\right]\right]+\text{H.c.}\ ,$$

where γ is a loop at x_1. As mentioned in Sec. I, from these quantities A^i_μ can be reconstructed up to gauge transformations and we can then determine (2.8) for any given experimental situation.

III. GAUGE FIELD AS A MAP ON A FINITE-DIMENSIONAL SET OF LOOPS

As already remarked, the set of all loops at a point 0 in an n-dimensional manifold M form an infinite-dimensional space. It is convenient to have a finite-dimensional subspace of these loops so that the specification of the holonomy transformations (parallel-transport operators) for these loops is sufficient to determine the gauge field up to gauge transformations. It was shown previously that there exists a $2n$-dimensional set of loops so that specifying the holonomy transformations for these curves is sufficient to reconstruct the gauge field.[6] But these holonomy transformations had to satisfy certain compatibility conditions which implies that there is redundancy in the chosen set of curves. We shall show now that there exists an n-dimensional set of loops at 0 so that specifying the gauge group elements for these curves in a differentiable manner is sufficient to reconstruct the corresponding gauge potential, which is then unique up to gauge transformations.

Let $(x^\mu,\ \mu=1,2,\ldots,n)$ be a coordinate system on M. If M does not admit a coordinate patch that covers all of M then the construction to be given now may be done locally, or it can be easily extended to apply to the entire manifold. Let 0 be a fixed point in M and K a set of differentiable curves through 0 which vary differentiably so that for every $P\in M$, $P\neq 0$, there is a unique curve belonging to K which passes through P. By an rth coordinate line we mean a line for which the coordinates $x^1, \ldots, x^n$, with the exception of x^r, have fixed values, while x^r varies over the real line. Let Σ be a hypersurface containing 0 such that every coordinate line meets Σ at exactly one point (Fig. 1).

Let $P(x^\mu)$ be an arbitrary point in M. Suppose that the rth coordinate line through P meets Σ at P_r. Let l, l_r be the unique elements of K that pass through P and P_r, respectively, and γ_r the closed curve at 0 formed from l, the rth coordinate line through P and l_r in the obvious way. Then γ_r $(r=1,\ldots,n)$ may be regarded as a function of M and hence $\{\gamma_r\}$ form an n-dimensional subspace of the infinite-dimensional set of loops at 0. Assign an element $\underline{g}(\gamma_r(x^\mu)) \equiv \underline{g}_r(x^\mu)$ of the gauge group G to each $\gamma_r(x^\mu)$ such that, for each r, $\underline{g}_r$ is a differentiable map of M into G, and when γ_r encloses zero area $\underline{g}(\gamma_r)$ is the identity. We now show that from $\{\underline{g}(\gamma_r(x^\mu))\}$ the vector potential can be determined in a particular gauge. This implies that *no* further restriction needs to be placed on $\underline{g}_r$ for it to describe a connection.

Suppose $P'(x^0,\ldots,x^r+\delta x^r,\ldots,x^n)$ is a point neighboring P and $l' \in K$ passes through P'. Then $\delta \underline{g}_r \equiv \{\underline{g}(\gamma_r(P'))\}^{-1}\underline{g}(\gamma_r(P))$ is the holonomy transformation for the closed curve formed from l, the rth coordinate line segment connecting P and P', and l'. Choose a gauge so that the parallel transport along any portion of each element of K is the identity. Then $\delta \underline{g}_r = 1 - igA_r^j T^j \delta x^r$ so that $\underline{g}_r^{-1}\partial \underline{g}_r/\partial x^r = igA_r^j T^j$ (no summation over r, g is the coupling constant), from which A_r^j can be determined. By varying r, the vector potential $A_\mu^j(x^\alpha)$, $\mu=1,\ldots,n$, can be determined in this gauge, which is called a radial gauge. Also, $A'^j_\mu(x^\alpha)$ is any other vector potential having the same $g_r(\gamma_r)$ as the holonomy transformations if and only if A'^j_μ and A^j_μ are related by the local gauge transformation

$$U(x^\mu) = P\left[\exp\left[-ig\int_{l\,0}^{P} T^j A'^j_\mu \, dx^\mu\right]\right].$$

This can be proved by writing $\underline{g}_r(x+dx)^{-1}\underline{g}_r(x)$ in terms of the connections A and A' which then give the stated gauge transformation between them and conversely this gauge transformation immediately implies that A and A' have the same $\underline{g}_r$'s. Also $\underline{g}_r$ satisfies the Yang-Mills equation

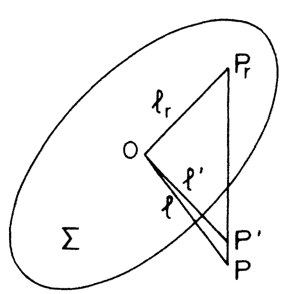


FIG. 1. Reconstruction of the vector potential in a radial gauge from the holonomy transformations for an n-dimensional space of loops in an n-dimensional manifold.

$$\partial_s\partial^s a^r - \partial_s\partial^r a^s + 2[a^s,\partial_s a^r] - [a^r,\partial_s a^s] - [a_s,\partial^r a^s] + [a_s,[a^s,a^r]] = T^i j^{ir},$$

where $a_r = \underline{g}_r^{-1}\partial_r \underline{g}_r$, and ∂_r denotes $\partial/\partial x^r$.

An example of M is the four-dimensional Minkowski space-time and A^j_μ is then the usual gauge potential in the radial gauge. Another example of M is the conformal completion of Minkowski space-time into a compact manifold[11] by adding a boundary called the conformal infinity $\mathscr{I}$, which is suitable for studying massless fields such as gauge fields. Then every asymptotically timelike line passes through two points I^+, I^- on $\mathscr{I}$ called future and past infinity, respectively. Every asymptotically spacelike line originates and terminates at a point I^0, on $\mathscr{I}$, called spatial infinity. The hypersurface $\mathscr{I}$ contains two hypersurfaces called the future null cone $\mathscr{I}^+$ which has I^+ as the vertex and the past null cone $\mathscr{I}^-$ which has I^- as the vertex. Every null line originates at a point on $\mathscr{I}^-$ and terminates at a point on $\mathscr{I}^+$.

The point I^+ may be taken as the point 0 in the construction given earlier and $\mathscr{I}^+$ the hypersurface Σ. The congruence K can be taken to be a set of parallel timelike lines and the generators of the cone $\mathscr{I}^+$, all of which pass through I^+. The coordinate system can be chosen so that one set of coordinate lines are the same as the timelike lines of K and the other coordinate lines are spacelike lines which pass through I^0. It follows that the gauge field can be constructed from the holonomy transformations associated with triangles that have two fixed vertices I^+, I^0, and the third vertex varying over the entire Minkowski space-time. The gauge in which A^j_μ $(\mu=0,1,2,3)$ is constructed is then called a temporal gauge since $A^j_0=0$ in this gauge in the above coordinate system. Thus, the temporal gauge is a special case of a radial gauge in this conformally completed Minkowski space-time. Also, it is interesting that, in this case, the basis loops are triangles with two points fixed, which is the simplest possible basis. It should be noted that a finite-dimensional basis of loops for the conformally completed Minkowski space-time has also been given by Kozameh and Newman.[12]

IV. AN EQUIVALENCE PRINCIPLE FOR GAUGE FIELDS

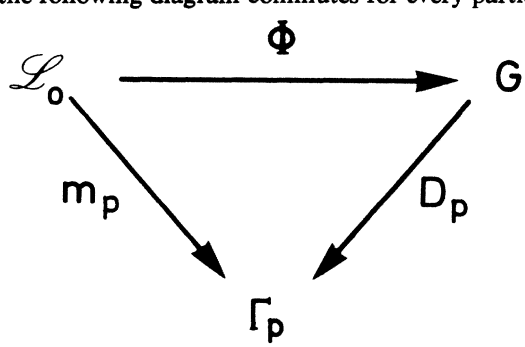
It may appear at first sight that it is not possible to have an equivalence principle for gauge fields because different particles, in general, have different motions in a gauge field. For instance, in an electromagnetic field, which is the simplest gauge field, particles with different charge-to-mass ratios have different classical trajectories. But by considering the *quantum-mechanical motions* of the particles, it is possible to formulate a principle of equivalence for gauge fields, which is analogous to Trautman's[13] formulation of the principle of equivalence for the gravitational field. The latter principle states that the same affine connection on space-time is determined by the classical motions of different test particles in a gravitational field.

But while classical motions are given by trajectories in space-time, quantum-mechanical motions are given by the propagation of De Broglie waves in space-time. We saw in Sec. II that when a gauge field is measured by a quantum-mechanical wave function, we determine a map from a set of loops into a representation of the gauge group, in an approximate sense as explained below.

To make this idea mathematically precise, we define by L_0 the set of all piecewise-differentiable closed curves in a manifold M beginning and ending at 0, which we called the loops at 0. Then L_0 has a differentiable structure obtained from M, which makes it an infinite-dimensional manifold.[14] Let $\mathscr{L}_0$ be the quotient set of L_0 by the equivalence relation that identifies any two loops whose images differ by the image of a loop that encloses no area, i.e., a path section which doubles back on itself. Then any function of $\mathscr{L}_0$ can be regarded as a function on L_0 by making this function constant on equivalent loops of L_0. By a map on $\mathscr{L}_0$ being "differentiable" we shall mean that it is differentiable as a map on L_0. Clearly $\mathscr{L}_0$ is a group with the product defined by the composition of loops. Then by using a particle p which probes the gauge field by means of an appropriate experiment, such as the interference of two coherent beams[4] or the generalized Josephson effect discussed in Sec. II, we can determine, in principle, a homomorphism m_p from $\mathscr{L}_0$ onto Γ_p, up to conjugacy, where Γ_p is a group of linear transformations.

Owing to the quantum uncertainty principle, this map can be determined operationally only approximately: The particle p cannot distinguish clearly between two loops whose images are separated by shortest distances of the order of the Compton wavelength $\hbar/mc$ of p. This uncertainty may be decreased by increasing the mass m, which, however, increases the Schwarszchild radius $2Gm/c^2$ of the particle. Since the uncertainty must also exceed the Schwarszchild radius, there is then a minimum to this uncertainty which can be shown to be the Planck length $(G\hbar/c^3)^{1/2}\sim 10^{-33}$ cm. This is like the determination of the geometry of space-time by a quantum-mechanical probe which can at best be determined to an uncertainty of the order of the Planck length.[15] If we do not take a strictly operational approach then the existence of m_p as a precise map may be postulated and should satisfy the principle of equivalence for gauge fields which we now state as follows.

There exists a finite-dimensional group G and differentiable homomorphisms $\Phi:\mathscr{L}_0\rightarrow G$ and $D_p:G\rightarrow\Gamma_p$ such that the following diagram commutes for every particle p:



Even though the group Γ_p and the maps m_p, D_p depend on the particle p, G and Φ are independent of the particle and determine the geometry of the gauge field in the spirit of the usual principle of equivalence. Example: For an electromagnetic field, Γ_p consists of just the identity if p is neutral and Γ_p is a nontrivial representation of U(1) or T(1) (translational group in one dimension) if p is charged. But G and Φ are the same in both cases. If G is chosen to be U(1), then the Γ_p's are restricted to be images of representations D_p of U(1) implying that charge is quantized. Conversely, the empirical fact that the charge is quantized, which is reflected in $\{\Gamma_p\}$, justifies the choice of $G=\text{U}(1)$ for electromagnetism. Thus the Aharonov-Bohm type of experiments or the Josephson effect which give the map m_p, together with charge quantization, imply that electromagnetism is a U(1) gauge field, according to the above principle of equivalence. This is because, using the theorem mentioned in Sec. I, for an arbitrary Lie group G, the gauge potential can be constructed from Φ and it is unique up to gauge transformations.[5,6] But Sec. III implies that the potential can be constructed from the restriction of Φ to a finite-dimensional set of loops. Hence, in the above equivalence principle L_0 may be replaced by this finite-dimensional basis of loops. Then Φ, m_p, and D_p (whose domain is the range of Φ) should obey only the restriction that they are differentiable and take the identity into the identity.

The above principle of equivalence will, of course, apply to the gravitational connection, in which case $G=\text{SL}(2,C)$. An interesting problem is whether it can be extended to provide a complete description of the gravitational field including the metric, in which case we would have a unified description of gravity and gauge fields.

Note added in proof. I have recently become aware of the work of R. Giles [Phys. Rev. D **24**, 2160 (1981)] and of L. Gross [J. Funct. Anal. **63**, 1 (1985)] on gauge field and holonomy which are related to the present work.

ACKNOWLEDGMENTS

This paper is based on a seminar given by the author at the University of California, Berkeley and he wishes to acknowledge financial support at this time from the National Science Foundation under Grant No. PHY 79-13146. He also thanks I. M. Singer for a useful suggestion.

APPENDIX: GENERALIZATION OF THE JOSEPHSON EFFECT FOR NON-ABELIAN GAUGE THEORIES

Consider a matter field ψ minimaly coupled to a gauge field A_μ^j and obeying a wave equation analogous to the Ginzburg-Landau equation obeyed by the order parameter in superconductivity. Suppose that there is a Josephson junction (a barrier that enables ψ from both sides to tunnel through it), in the region $x_1<x<x_2$, in a Cartesian coordinate system.

In the case of superconductivity, the current through the Josephson junction can be obtained from the boundary conditions for the Ginzburg-Landau equation that were found from the microscopic theory by De Gennes.[16] These conditions may be stated, in a slightly more general form, as

$$\frac{\partial\psi}{\partial x}(x_1)+\frac{i2e}{\hbar c}A_x(x_1)\psi(x_1) =\alpha\left[\beta\exp\left[-\frac{i2e}{\hbar c}\int_{x_2}^{x_1}A_x dx\right]\psi(x_2)-\psi(x_1)\right],$$

$$\frac{\partial\psi}{\partial x}(x_2)+\frac{i2e}{\hbar c}A_x(x_2)\psi(x_2) =\alpha\left[\psi(x_2)-\beta\exp\left[-\frac{i2e}{\hbar c}\int_{x_1}^{x_2}A_x dx\right]\psi(x_1)\right],\tag{A1}$$

where α and β are real positive numbers, and we have put in the electromagnetic phase factors in the right-hand side of these two equations to make them explicitly gauge covariant. The simplest generalization of (A1) for an arbitrary gauge theory is

$$\frac{\partial\psi}{\partial x}(x_1)+igA_x^j(x_1)T^j\psi(x_1) =\alpha\left\{\beta P\left[\exp\left[-ig\int_{x_2}^{x_1}A_x^jT^jdx\right]\right]\psi(x_2)-\psi(x_1)\right\},$$

$$\frac{\partial\psi}{\partial x}(x_2)+igA_x^j(x_2)T^j\psi(x_2) =\alpha\left\{\psi(x_2)-\beta P\left[\exp\left[-ig\int_{x_1}^{x_2}A_x^jT^jdx\right]\right]\psi(x_1)\right\}.\tag{A2}$$

If ψ is a scalar field then its current in the x direction is

$$j^k(x)=\frac{-i\hbar}{2m}\left[\psi(x)^\dagger T^k\left[\frac{\partial\psi}{\partial x}(x)+igA_x^j(x)T^j\psi(x)\right] -\left[\frac{\partial\psi}{\partial x}(x)+igA^j(x)T^j\psi(x)\right]^\dagger T^k\psi(x)\right],\tag{A3}$$

where T^j are the (Hermitian) generators of the representation of the gauge group. Hence, on using (A2),

$$j^k(x_1)=-\frac{i\hbar\alpha\beta}{2m}\left\{\psi^\dagger(x_1)T^kP\left[\exp\left[-ig\int_{x_2}^{x_1}A_x^jT^jdx\right]\right]\psi(x_2)-\psi^\dagger(x_2)P\left[\exp\left[-ig\int_{x_1}^{x_2}A_x^jT^jdx\right]\right]T^k\psi(x_1)\right\},$$

$$j^k(x_2)=-\frac{i\hbar\alpha\beta}{2m}\left\{\psi^\dagger(x_1)P\left[\exp\left[-ig\int_{x_2}^{x_1}A_x^jT^jdx\right]\right]T^k\psi(x_2)-\psi^\dagger(x_2)T^kP\left[\exp\left[-ig\int_{x_1}^{x_2}A_x^jT^jdx\right]\right]\psi(x_1)\right\}.\tag{A4}$$

To obtain the current at any point $x\in[x_1,x_2]$, we try the ansatz

$$j^k(x)=-\frac{i\hbar\alpha\beta}{2m}\left\{\psi^\dagger(x_1)P\left[\exp\left[-ig\int_{x}^{x_1}A_x^jT^jdx\right]\right]T^kP\left[\exp\left[-ig\int_{x_2}^{x}A_x^jT^jdx\right]\right]\psi(x_2) -\psi^\dagger(x_2)P\left[\exp\left[-ig\int_{x}^{x_2}A_x^jT^jdx\right]\right]T^kP\left[\exp\left[-ig\int_{x_1}^{x}A_x^jT^jdx\right]\right]\psi(x_1)\right\},\quad x_1\le x\le x_2\,,\tag{A5}$$

which contains the equations (A4) as special cases. We now show that the current (A5) is covariantly conserved, in the stationary situation in which a gauge can be chosen such that $A_0^j=0$ and $\partial j_0^k/\partial t=0$. Then,

$$\frac{\partial j^k}{\partial x}(x)=-\frac{\hbar\alpha\beta g}{2m}A_x^l\left\{\psi^\dagger(x_1)P\left[\exp\left[-ig\int_x^{x_1}A_x^jdx\right]\right] \times(T^kT^l-T^lT^k)P\left[\exp\left[-ig\int_{x_2}^{x}A_x^jT^jdx\right]\right]\psi(x_2)\right\}+\text{H.c.}=gC^{klm}A_x^lj^m$$

on using the Lie-algebra relations $[T^l,T^k]=ic^{lkm}T^m$. Hence, the gauge-covariant divergence of the current

$$\frac{\partial j^{k\mu}}{\partial x^\mu}+gC^{lkm}A_\mu^lj^{m\mu}=0\,.$$

Since (A5) is covariantly conserved and is in agreement with the currents at $x=x_1$ and $x=x_2$, it follows that (A5) is the current at any $x\in[x_1,x_2]$.

---

*Present address: Department of Physics and Astronomy, University of South Carolina, Columbia, SC 29208.

[1]C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954).
[2]Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).
[3]T. T. Wu and C. N. Yang, Phys. Rev. D **12**, 3843 (1975).
[4]J. Anandan, Nuovo Cimento **53A**, 221 (1979).
[5]J. Anandan, Int. J. Theor. Phys. **19**, 537 (1980), Appendix.
[6]J. Anandan, in *Conference on Differential Geometric Methods in Theoretical Physics, Trieste, 1981*, edited by G. Denardo and H. D. Doebner (World Scientific, Singapore, 1983).

[7]B. D. Josephson, Phys. Lett. **1**, 251 (1962); Rev. Mod. Phys. **36**, 216 (1964).

[8]J. E. Zimmermann and J. E. Mercerau, Phys. Rev. Lett. **14**, 887 (1965); J. Anandan, Phys. Rev. D **24**, 338 (1981).

[9]A. H. Guth, Phys. Rev. D **23**, 347 (1981).

[10]H. B. Nielsen and P. Olsen, Nucl. Phys. **B61**, 45 (1973).

[11]R. Penrose, in *Relativity, Groups and Topology,* edited by B. DeWitt and C. DeWitt (Gordon and Breach, New York, 1964).

[12]C. N. Kozameh and E. T. Newman, Phys. Rev. D **31**, 801 (1985).

[13]A. Trautman, Usp. Fiz. Nauk. **89**, 3 (1966) [Sov. Phys. Usp. **9**, 319 (1966)].

[14]See, for example, J. Eells, in *Symposium Internacional de Topologia Algebraica, Mexico, 1956* (La Universidad Nacional Autonoma de Mexico y la Unesco, 1958).

[15]J. Anandan, in *Quantum Optics, Experimental Gravitation and Measurement Theory,* edited by P. Meystre and M. O. Scully (Plenum, New York, 1983).

[16]P. G. DeGennes, Phys. Lett. **5**, 22 (1963); Rev. Mod. Phys. **36**, 225 (1964).