Global Topology and Local Violation of Discrete Symmetries

J. Anandan*
Raman Research Institute, C.V. Raman Avenue, Bangalore 560 080, India
and Department of Physics and Astronomy, University of South Carolina, Columbia, South Carolina 29208
(Received 14 March 1997; revised manuscript received 17 December 1997)

Cosmological models that are locally consistent with general relativity and the standard model are constructed, in which an object transported around the Universe undergoes P, C, and CP transformations. This leads to an enlargement of the gauge groups of electroweak and strong interactions to include antiunitary transformations. Gedanken experiments in these cosmological models show that if all interactions obey Einstein causality then P, C, and CP cannot be violated. But another model in which the sign of the charge may be reversed is allowed. Some possible implications to modifications of the standard model are considered. [S0031-9007(98)06512-0]

PACS numbers: 98.80.Hw, 02.40.–k, 04.20.Gz, 11.30.Er

The great success of the standard model has provided hardly any experimental motivation to modify it at present. I consider here some interesting physical consequences of generalizing the gravitational, electromagnetic, weak, and strong fields, by modifying the global topology of an appropriate Kaluza-Klein space-time. Such topologies need to be studied also if in quantum gravity the Feynman amplitudes for all possible topologies are summed. These generalizations are locally consistent with the *causal* dynamics of the standard model and general relativity. And yet, it will be shown by gedanken experiments around global circuits in space-time that all but one of these models are incompatible with the observed violations of parity (P), charge conjugation (C), and CP symmetries. This leads to the consideration of some possible modifications of the standard model.

Consider first a nonorientable space [1]. An example is obtained by identifying a pair of opposite faces of a rectangular box (Fig. 1) continuously so that A, B, C, D are identified with A', B', C', D', respectively. All sections parallel to $ABA'B'$ have the topology of the Mobius strip M^2. Now let AB, BC become infinite in length while keeping $L = AB'$ large but finite. The Cartesian product of this space with the real line R is a nonorientable manifold $M^4 = M^2 \times R^2$. This amounts to the identification $(0, y, z, t) \leftrightarrow (L, y, -z, t)$. We may take this multiply connected manifold [2] endowed with a flat Minkowskian metric to be a space-time in the absence of matter.

In the presence of matter, we may consider the Einstein–de Sitter or the Friedmann-Robertson-Walker cosmological model with zero spatial curvature, with metric [3]

$$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu$$
$$= -c^2 dt^2 + a^2(t)(dx^2 + dy^2 + dz^2) \qquad (1)$$

and energy-momentum tensor

$$T^{\mu\nu} = (\rho + P) u^\mu u^\nu + P g^{\mu\nu}, \qquad (2)$$

where the density ρ and the pressure P are constant in each hypersurface orthogonal to u^μ. Then (1) and (2) satisfy the Einstein's field equations for appropriate choices of $a(t)$, $\rho(t)$, and $P(t)$ with $u^\mu = \delta_0^\mu$. For example, for a pressure-free universe (galaxies idealized as grains of dust with their random velocities neglected) $P = 0$ and then $a(t) = At^{2/3}$, whereas for radiation $P = \frac{1}{3}\rho$ and then $a(t) = Bt^{1/2}$, where A and B are constants [3], assuming zero cosmological constant. All of the astrophysical evidence we have at present are consistent with (1). Again, this cosmology may be made nonorientable by the identification described above (Fig. 1) to obtain a space-time, denoted S_1.

Note that when the triad $OXYZ$ is taken to O', identified with O, its z axis has reversed direction compared to an identical triad which was left at O, while the X and Y axes remain the same (Fig. 1). So, the triad has changed handedness around this closed curve. A left-handed glove taken around any such closed curve, denoted Γ, will return as a right-handed glove. Another interesting aspect of this and other space-times discussed here is that they allow for some locally conserved quantities to be globally nonconserved. For example, the momentum $\mathbf{p}$ of a free particle moving in the space $M^3 = M^2 \times R$ described in Fig. 1 is *locally* conserved, meaning that in any orientable neighborhood containing the particle $\mathbf{p}$ is conserved. But if it goes around Γ then its momentum component p_z

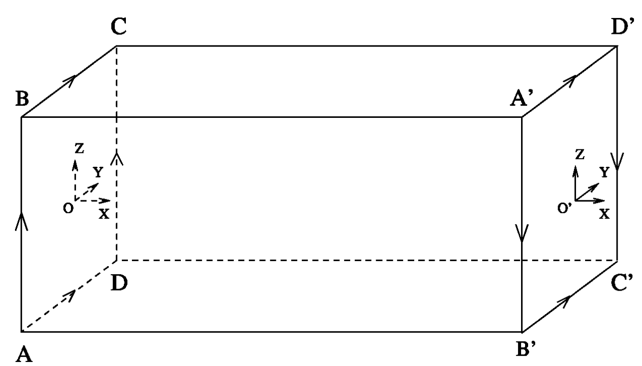

FIG. 1. A nonorientable space with zero spatial curvature, obtained by identifying the ends $ABCD$ and $A'B'C'D'$ so that each primed letter represents the same location as the unprimed one.

would reverse. Similarly, the angular momentum $\mathbf{J}$ of a torque-free gyroscope would be locally conserved. Yet if it goes around Γ then the components J_x, J_y of $\mathbf{J}$ would reverse. Therefore, p_z, J_x, J_y are not globally conserved. Also, the helicity $\mathbf{J} \cdot \mathbf{p}/|\mathbf{p}|$ would reverse so that a left-handed neutrino would return as a right-handed neutrino.

To see how this is possible, note first that for every Killing field ξ_a^μ, the law $\nabla_\nu T^{\mu\nu} = 0$ implies, via Killing's equation, the local conservation law $\nabla_\mu j_a^\mu = 0$ where $j_a^\mu = T^{\mu\nu}\xi_{a\nu}$. The locally conserved momentum and angular momentum components correspond to the independent translational and rotational Killing fields of M^3. On defining the "charges" inside an oriented manifold V with boundary ∂V by $Q_a = \int_V \sqrt{-g}\, j_a^0\, d^3x$, from Gauss' theorem,

$$\frac{dQ_a}{dt} = -\int_{\partial V} \sqrt{-g}\, j_a^i\, dS_i\,. \qquad (3)$$

So, Q_a is conserved iff the flux that is the right-hand side (RHS) of (3) vanishes. If V is taken to be the interior of the box in Fig. 1, then as a particle goes out of V through the end $A'B'C'D'$, because of the identification, it simultaneously comes into V through the opposite end $ABCD$ in M^3. Since the identification reverses the z direction, the contribution that this simultaneous exit and entry the particle makes to the RHS of (3) vanishes for $Q_a = p_x, p_y, J_z$ but not for $Q_a = p_z, J_x, J_y$. Even the definition of p_z, J_x, J_y depends on the chosen neighborhoods; the above argument shows that for the chosen maximal neighborhood they are not conserved, unlike p_x, p_y, and J_z. This is because the Killing fields corresponding to p_z, J_x, J_y cannot be globally defined because of the reversal of z direction in the identification.

The local U(1) gauge symmetry of electromagnetism implies that the U(1) group acts locally at each point in space-time. This naturally leads to the five dimensional Kaluza-Klein (KK) geometry that is obtained from space-time by replacing each point by a circle on which the U(1) group acts locally, and conversely. The electromagnetic field provides a (1-1) correspondence between neighboring circles, called a connection. As we go around a closed space-time curve, denoted γ, beginning and ending at a point o, this correspondence leads to the rotation of the circle at o, called the holonomy transformation of the electromagnetic connection. When a wave function is taken around this curve it is acted upon by this rotation and acquires the phase factor $\exp(-ie \oint_\gamma A_\mu\, dx^\mu)$, which can be experimentally observed, where A_μ is the electromagnetic potential. If e is the smallest unit of charge, then these phase factors for different closed curves γ define the electromagnetic field [4]. Then, for a given γ and electromagnetic field, these phase factors may be used to define the various charges that replace e.

The electromagnetic field may now be generalized by allowing the identification of the U(1) circles at o to be in the opposite sense so that γ is the projection of a Klein bottle in the KK space-time. In particular, if (x, y, z, t, ϕ) are the coordinates of the KK space-time, where ϕ is the angular variable in the fifth dimension, consider the slab of space-time $0 \leq x \leq L$ with the identification of its ends by the homeomorphism $(0, y, z, t, \phi) \leftrightarrow (L, y, z, t, -\phi)$. Its projection on the usual space-time may be endowed with the metric (1). In this new KK space-time, denoted S_2, each two dimensional surface of constant y, z, t is a Klein bottle K^2. So, S_2 is topologically $K^2 \times R^3$.

Such a generalization amounts to enlarging the electromagnetic gauge group U(1) to O(2) that is generated by the rotations SO(2) $\cong$ U(1) of the circle and the reflection E about a diameter. Because the above circle at o now undergoes a rotation and E under the holonomy transformation associated with γ. Although O(2) is non-Abelian, because it is one dimensional, the gauge field is still Abelian.

Suppose two observers start from the same point on γ and go around γ and meet. Each would then claim, with equal justification, that the charges of all the particles in the other observer have changed sign. So, it is not possible to determine unambiguously whether the sign of two charges at distinct points are the same, because it is necessary to bring these charges to the same space-time point in order to compare them and the result would depend on the paths they take. Only the absolute value of the ratio of the charges would be meaningful. Also, charge is locally conserved because of the U(1) symmetry but is not globally conserved. This is because if two charges at the same space-time point are taken along different paths and brought together again their sum may change. This is similar to the global nonconservation of momentum and angular momentum mentioned above, and can be understood in the same way by means of Gauss' theorem. Again this is due to the Killing field in the fifth dimension that generates the U(1) symmetry not being globally defined. But there is local O(2) gauge symmetry, which includes the usual local U(1) gauge symmetry of electromagnetism.

Also, suppose a charged particle wave function is split into two wave functions that are made to interfere around a closed curve whose generalized electromagnetic holonomy transformation is an improper O(2) transformation. The superposed wave function then has the form

$$\psi(x^\mu, \phi) = \exp(ie\phi)\psi_1(x^\mu) + \exp(-ie\phi)\psi_2(x^\mu)\,, \quad (4)$$

in a local gauge. Now, $\psi^*\psi$ has a nontrivial ϕ dependence that makes it spontaneously break the O(2) symmetry down to the discrete group consisting of an appropriate E and the identity. The charge operator $Q = i\frac{\partial}{\partial\phi}$. Since ψ is a superposition of opposite charges, it violates the "charge superselection rule."

This shows that the often made claim that the local U(1) gauge symmetry implies the charge superselection rule is incorrect, because the O(2) gauge symmetry here contains U(1). When Aharonov and Susskind [5] refuted this claim, they showed how a subsystem may be in a superposition of charge eigenstates, while the entire system does not violate the charge superselection rule. An example is the BCS ground state of a superconductor in which the

Cooper pairs are in a superposition of different charge eigenstates, thereby breaking the electromagnetic U(1) gauge symmetry spontaneously, while the entire superconductor may have a well defined charge and thus obey the charge superselection rule. But in the present case the entire system may be in superposition of charge eigenstates, and is therefore a stronger violation of this rule.

In the usual electromagnetic theory, Q commutes with the *interactions* so that the eigenstates of Q form a "preferred basis" in which the density matrix is diagonal. This gives an effective charge superselection rule. In the present more general electromagnetic theory, because the time evolution may contain E which does not commute with Q that generates the electromagnetic U(1), it is "easy" to produce a superposition of opposite charged states, as in the above example. When such a superposition interacts with an apparatus, the apparatus wave function intensity also gets modulated correspondingly in the fifth dimension. *This would make the fifth dimension observable,* like the other four dimensions.

This construction may be extended to the standard model for which the gauge group is $G = \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)$. The C transform of a spinor ψ is $\psi^C = i\gamma^2\psi^*$, where the $*$ denotes complex conjugation or Hermitian conjugation in quantum field theory. Therefore, as $\psi \rightarrow g\psi$ under $g \epsilon G$, $\psi^C \rightarrow g^*\psi^C$. In the above construction each Klein bottle may be replaced by a generalized Klein bottle that is closed by means of the automorphism α of G that is the complex conjugation $\alpha(g) = g^*$ for every $g \epsilon G$, and the automorphism β of the spinor Lorentz group Λ defined by $\beta(S) = i\gamma^2 S^*(i\gamma^2)^{-1} = -\gamma^2 S^* \gamma^2$ for every $S \epsilon \Lambda$. If each fiber is a homogeneous space G/H, where H is a subgroup of G such that $H^* = H$, then the new KK space-time, denoted S_3, is obtained by the identification $(0, y, z, t, Hg) \leftrightarrow (L, y, z, t, Hg^*)$. This performs a C transformation on all the quantum numbers coupled to the gauge fields of G. Since the operational meaning of a particle is contained in all its interactions, a particle taken around γ in S_3 would become its antiparticle, i.e., it would undergo a C transformation. For example, when taken around γ a neutrino will return as an antineutrino.

This leads to a generalization of the standard model in which the gauge group G is enlarged to a group $\tilde{G}$ that is generated by G acting on itself on the right and the automorphism α. Then $\tilde{G} = \mathrm{O}(2) \times \tilde{\mathrm{SU}}(2) \times \tilde{\mathrm{SU}}(3)$, where $\tilde{\mathrm{SU}}(n)$ denotes the group of unitary and antiunitary transformations on an n dimensional complex vector space that have determinant 1. Even when H is trivial so that $G/H = G$, S_3 is not a principal fiber bundle, because the "twist" in the generalized Klein bottle prevents the definition of the right action of G everywhere. But S_3 is an associated bundle of a principal fiber bundle with structure group $\tilde{G}$ over the usual space-time as the common base manifold. Again by superposing C eigenstates with distinct eigenvalues the extra dimensions become observable, as in the case of S_2 above.

A CP transformation may also be implemented physically by identifying the opposite faces of the slab according to $(0, y, z, t, Hg) \leftrightarrow (L, y, -z, t, Hg^*)$. This new KK space-time will be denoted by S_4. Time reversal may be implemented by the identification $(0, y, z, t) \leftrightarrow (L, y, z, -t)$ so that space-time is time nonorientable. But this would violate causality.

It is not necessary to go all the way around the Universe to obtain the above discrete transformations. Cosmic strings, which are predicted to occur in the early universe, have been characterized by proper orthochronous Poincaré transformations of the affine holonomy group around it [6,7]. These solutions may be generalized to include also discrete transformations of the entire Poincaré group as holonomy, e.g., reversal of the direction along the axis of the cosmic string. The discrete holonomy transformations would require taking out the axis of the cosmic string from space-time or turning it into a singularity. This would constitute a generalization of the gravitational field according to an earlier definition of the gravitational field [7]. A generalized gauge field "flux" may also be introduced into the string by letting the gauge field holonomy around the string to include the new antiunitary transformation α introduced above.

Except for S_2 (and its cosmic string analog) all the space-times discussed above are disallowed by the violation of discrete symmetries in weak interaction [1]. In S_1 consider two small capsules U and U' at the same location, each containing the apparatus for the P violating experiment proposed by Lee and Yang [8], and performed by Wu *et al.* [9]. The magnetic coil, which orients the Co nuclei placed at the center of the coil, is in the x-y plane. When the nuclei undergo β decay, let the intensity distribution of electrons be $f(\theta)$, where θ is the angle between the velocity of the emitted electron and the z axis. Then, $f(\theta) \neq f(180° - \theta)$, which violates P. Suppose now that the two capsules are taken along curves that form a circuit γ such that the handedness changes during continuous transport around γ. Let there be two twins in the capsules performing the two respective experiments. When they meet again and compare their experiments they would find that the currents in the two coils in the X-Y plane are flowing in the same direction. However, the distribution of the outgoing electrons in U' is $f(\theta') = f(180° - \theta)$, which would be in conflict with the distribution $f(\theta)$ obtained in the identical experiment performed in the capsule U. Unlike the "twin paradox" in special relativity (which is not a paradox), here there is perfect symmetry between the two twins: each twin would be justified in saying that it is the other who has undergone a P transformation. But the above contradiction disallows S_1.

Since β decay also violates C, S_3 is also disallowed by the above type of gedanken experiment. Similarly, S_4 is disallowed by doing identical experiments involving kaon decay, which violates CP, in the two capsules. But S_2, with the generalized O(2) electromagnetic gauge field introduced above, is allowed because the charge reversal symmetry (or C restricted to purely electromagnetic phenomena) is an exact symmetry in all known phenomena.

In an expanding universe, there may not be enough time for the capsules to go all the way around the Universe and meet. But in principle, a ring of large number of capsules $\{U_1, U_2, \ldots\}$ may be set up around the Universe. Two identical capsules V_n, V_{n+1} meet midway between two neighboring capsules U_n, U_{n+1} at time $t = -T$; then V_n meets U_n and V_{n+1} meets U_{n+1} at $t = 0 > -T$. Finally, V_n and V_{n+1} meet again at $t = T$ to verify if the relevant experiments in U_n and U_{n+1} gave the same result at $t = 0$. But in each of the above space-times, except S_2, there would then be some n for which the experiments disagree, disallowing this space-time.

It is intriguing that local experiments, combined with the above-mentioned gedanken experiments, should give information about the global topology of space-time. How could a neutron "know" the global topology of space-time so that it can safely decay in a P violating way without leading to the above contradiction if it were taken around the Universe? This suggests that the laws of physics have a global aspect due to the reproducibility of an experiment at different parts of the Universe, such as the experiments in the ring of capsules mentioned above.

We may, however, try to find a direct connection between global topology and local violation of discrete symmetries in terms of actual interactions. It appears that the simplest way of doing this is to suppose that P, C, and CP are not violated by the laws of physics at the most fundamental level, but that these symmetries are broken spontaneously. In the case of P violation, there are then two sets of possible degenerate vacua that are associated with the two possible equivalent orientations. If the boundary conditions in the early universe are such that space is nonorientable then neither vacuum can be chosen all the way around the Universe. So, P is either not spontaneously broken, or broken in orientable domains so that for a pair of neighboring domains different orientations are chosen by P violation. But if the boundary conditions are such that space is orientable then a vacuum with the same orientation may be chosen everywhere so that P is violated in the same manner. It is emphasized that the spontaneous breaking of symmetry may occur even though the interactions are local and causal.

Another possibility is to give up Einstein causality in the fundamental physical laws, which was assumed to obtain the above contradictions. Since this principle is very well confirmed by our experience, it appears that we could give it up only in the early universe when quantum gravitational effects were important. This is reasonable also because quantum gravity requires the quantization of space-time geometry including its causal structure, and is therefore inherently noncausal and perhaps also nonlocal. So, if quantum gravity violates P, C, and CP, then its laws could determine, in a noncausal way, the global topology of the KK space-time to be compatible with these violations, which persisted as the Universe cooled down.

As for the first possibility, left-right symmetric models with spontaneous violation of P have been proposed [10]. But in the above approach, C and CP may also be expected to be violated spontaneously. The second possibility, above, has the advantage that the noncausal nature of quantum gravity could resolve the horizon problem [11], namely, the fact that regions in the early universe which are causally unrelated nevertheless have similar properties, such as temperature and density.

If quantum gravity, which is expected to unify all the interactions, were to violate P, T, C, and CP, then it is not surprising that the electroweak theory obtained as a low energy limit of quantum gravity should also violate these symmetries. It would equally not be surprising if the classical gravitational field, also a low energy limit of quantum gravity, contains a residual violation of P, C, and CP (respectively, CT, PT, and T, assuming CPT symmetry). It is therefore worthwhile to look for experimental evidence of violation of these discrete symmetries in the gravitational interaction [12,13].

I thank Yakir Aharonov, Ralph Howard, Simon Donaldson, and Paule Saksidas for useful discussions. This work was supported by NSF Grant No. PHY-9601280.

*Electronic address: jeeva@sc.edu

[1] Ya. B. Zeldovich and I. D. Novikov, JETP Lett. **6**, 772 (1966). See also D. D. Sokolov, Sov. Phys. Dokl. **195**, 1307 (1970).
[2] Multiply connected universes have been considered by J. R. Gott, Mon. Not. R. Astron. Soc. **193**, 153 (1980).
[3] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W.H. Freeman, New York, 1973).
[4] T. T. Wu and C. N. Yang, Phys. Rev. D **12**, 3843 (1975).
[5] Y. Aharonov and L. Susskind, Phys. Rev. **155**, 1428 (1967). See also E. Lubkin, Ann. Phys. (N.Y.) **56**, 69 (1970).
[6] K. P. Tod, Classical Quantum Gravity **11**, 1331 (1994).
[7] J. Anandan, Phys. Rev. D **53**, 779 (1996).
[8] T. D. Lee and C. N. Yang, Phys. Rev. **104**, 254 (1956).
[9] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. F. Hudson, Phys. Rev. **105**, 1413 (1957).
[10] G. Senjanovic and R. N. Mohapatra, Phys. Rev. D **12**, 1502 (1975); R. N. Mohapatra and G. Senjanovic, Phys. Rev. D **23**, 165 (1981); R. Mohapatra and Andrija Rasin, Phys. Rev. D **54**, 5835 (1996). These models are based on the suggestion of left- right symmetry by J. C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974); R. N. Mohapatra and J. C. Pati, Phys. Rev. D **11**, 566 (1975); **11**, 2558 (1975); which, however, had a slight difference in the masses of left and right Higgs multiplets. This would result in a contradiction if space were nonorientable in an appropriate gedanken experiment involving continuous transport around γ, as discussed in the present Letter.
[11] W. Rindler, Mon. Not. R. Astron. Soc. **116**, 663 (1956); Charles W. Misner, Astrophys. J. **151**, 431 (1968); Phys. Rev. Lett. **22**, 1071 (1969).
[12] N. D. Hari Dass, Phys. Rev. Lett. **36**, 393 (1976); Ann. Phys. (N.Y.) **107**, 337 (1977).
[13] J. Anandan, Phys. Rev. Lett. **48**, 1660 (1982).