(2+1)-dimensional spacetimes containing closed timelike curves

Matthew P. Headrick and J. Richard Gott III
Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544
(Received 10 June 1994)

We investigate the global geometries of (2+1)-dimensional spacetimes as characterized by the transformations undergone by tangent spaces upon parallel transport around closed curves. We critically discuss the use of the term "total energy-momentum" as a label for such parallel-transport transformations, pointing out several problems with it. We then investigate parallel-transport transformations in the known (2+1)-dimensional spacetimes containing closed timelike curves (CTC's), and introduce a few new such spacetimes. Using the more specific concept of the holonomy of a closed curve, applicable in simply connected spacetimes, we emphasize that Gott's two-particle CTC-containing spacetime does *not* have a tachyonic geometry. Finally, we prove the following modified version of Kabat's conjecture: if a CTC is deformable to spacelike or null infinity while remaining a CTC, then its parallel-transport transformation cannot be a rotation; therefore its holonomy, if defined, cannot be a rotation other than through a multiple of 2π.

PACS number(s): 04.20.Gz, 04.20.Jb

I. INTRODUCTION

This paper concerns time travel in (2+1)-dimensional spacetimes governed by an analogue of Einsteinian gravity [1,2]. Such spacetimes serve simultaneously as toy models whose underlying structure is very similar to (3+1)-dimensional gravity but with far simpler dynamics, and as a handy notation for investigating a limited kind of cosmic string dynamics. The second use for (2+1)-dimensional spacetimes may sound less significant than the first, but in fact it adds considerable physical relevance to the work. For instance, when Gott discovered [3] that in 2+1 dimensions surprisingly simple situations could produce closed timelike curves (CTC's) and therefore support time travel, physicists would have paid little attention had there not been the possibility that such situations could be physically realized by cosmic strings in 3+1 dimensions.

In (2+1)-dimensional spacetimes the connection between physics and geometry provided by Einstein's equations can be written in an almost trivial way as a relation between the physical matter fields and the operation of parallel transport of vectors around closed curves. This paper is primarily about the use of the geometrical tool of parallel transport for understanding the structure of spacetimes, especially spacetimes containing closed timelike curves. In Sec. II we explain the relation between the physical matter fields and the parallel transport operation, and discuss the merits and problems of a proposed definitional association between this geometrical tool and the physical notion of total momentum. One of the main purposes of this paper is to show that, despite superficial similarities, Gott's CTC-containing solution in fact has little in common with the solution for a spacetime containing a single tachyon, which some investigators find unphysical [4]. In order to do that we explicitly construct the tachyon solution in this section. Finally, following Carroll *et al.* [5], we use parallel transport to define the more specific concept of the holonomy of a closed curve in simply connected spacetimes. In Sec. III we review the known (2+1)-dimensional spacetimes containing CTC's and introduce a few new ones. While doing so we find out what we can about the parallel transport operation and the holonomy in such spacetimes, and their relation to the existence of CTC's. In Sec. IV we prove the following theorem, which can be seen as a modified version of Kabat's conjecture: if a CTC can be deformed to spacelike or null infinity while remaining a CTC, then the result of parallel transport around it cannot be a rotation. Finally, in Sec. V we comment briefly on the implications of (2+1)-dimensional spacetimes containing CTC's to time travel in the real (3+1)-dimensional world.

II. THE PARALLEL-TRANSPORT TRANSFORMATION AND THE HOLONOMY

The most striking feature of gravity in 2+1 dimensions is the absence of vacuum degrees of freedom of the curvature. Unlike in 3+1 dimensions, the Riemann tensor $R_{\mu\sigma\rho\nu}$ is a unique function of the Ricci tensor $R_{\mu\sigma}$. By Einstein's equations the Ricci tensor is in turn a function of the energy-momentum tensor $T_{\mu\nu}$.

The geometry of a spacetime may be characterized by the Lorentz transformations suffered by the tangent spaces of points upon parallel transport around closed curves. These Lorentz transformations are the main object of study of this section and are important for the entire paper. For brevity, they will be called parallel-transport (PT) transformations. A PT transformation is a Lorentz transformation that acts on the tangent space of a given point, called the base point, and is defined for a given curve that starts and ends at that point. We can use PT transformations to express the relation between the local geometry of the spacetime and the energy-momentum tensor.

0556-2821/94/50(12)/7244(16)/$06.00

The local geometry at a point of spacetime may be characterized by the PT transformations for infinitesimal closed planar curves within a neighborhood of that point. In general, an infinitesimal loop c can be characterized by the antisymmetric tensor

$$f^{\rho\sigma} = \oint_c x^\rho dx^\sigma , \tag{1}$$

while the resulting infinitesimal transformation $\Lambda_\mu{}^\nu(c)$ can be expressed as the two-form

$$\omega_{\mu\nu} \equiv \Lambda_{\mu\nu} - g_{\mu\nu} , \tag{2}$$

hence the fact that the Riemann tensor is of rank four:

$$\omega_{\mu\nu} = \tfrac{1}{2} R_{\mu\sigma\rho\nu} f^{\rho\sigma} . \tag{3}$$

In three dimensions, however, both $f^{\rho\sigma}$ and $\omega_{\mu\nu}$ may be expressed as single vectors:

$$n_\nu \equiv \tfrac{1}{2}\sqrt{g}\,\varepsilon_{\nu\sigma\rho} f^{\rho\sigma} \tag{4}$$

is the normal to the loop with a length equal to the area it encloses, while

$$a^\rho \equiv \frac{1}{2\sqrt{g}} \varepsilon^{\rho\nu\mu}\omega_{\mu\nu} \tag{5}$$

is the axis of the transformation with a length equal to its angle or boost parameter. In 2+1 dimensions, thanks to Einstein's equations and the identity

$$R^{\mu\nu} - \tfrac{1}{2} R g^{\mu\nu} = \frac{1}{4g}\varepsilon^{\beta\gamma\mu}\varepsilon^{\rho\sigma\nu} R_{\beta\gamma\rho\sigma} , \tag{6}$$

we then have the following remarkably simple relation:

$$a^\mu = 8\pi T^{\mu\nu} n_\nu . \tag{7}$$

The quantity $T^{\mu\nu}n_\nu$ itself can in general be interpreted as the total energy-momentum of the matter "encircled" by the infinitesimal loop c represented by n_ν, that is, passing through the infinitesimal surface with boundary c.

Consider, for example, a spacetime containing static matter: $T^{\mu\nu} = 0$ except $T^{00} = \rho(x^1, x^2)$. We then have $a^0 = 8\pi\rho n_0$, $a^1 = a^2 = 0$; in other words, parallel transport induces a rotation about an angle of $8\pi\rho$ times the area of the projection of the loop onto a constant-time surface, or 8π times the amount of matter passing through the loop. This describes the geometry of the spacetime: constant-time surfaces have Gaussian curvature $8\pi\rho$, while surfaces orthogonal to them are flat. Similarly, a spacetime containing transcendent tachyonic matter moving in the x^1 direction will have $T^{\mu\nu} = 0$ except $T^{11} = p(x^0, x^2)$, so that $a^1 = 8\pi p n_1$, $a^0 = a^2 = 0$. (Note that this energy-momentum tensor satisfies the weak and the strong energy conditions but not the dominant one.) Again, the geometry is flat in surfaces containing the direction of motion of the matter. Parallel transport around an infinitesimal loop that encircles some of the matter results in a boost through a boost parameter of 8π times the amount of matter encircled. A third simple case is a spacetime filled with photonic matter (matter moving at the speed of light) moving in the x^1 direction. Here the energy-momentum tensor is given by $T^{00} = T^{01} = T^{10} = T^{11} = \rho$, with all other components zero. We then have $a^0 = a^1 = 8\pi\rho(n_0 + n_1)$, $a^2 = 0$, which represents a null transformation of parameter $8\pi\rho(n_0 + n_1)$. The flat planes are again those containing the direction of motion of the matter $dx^0 = dx^1$, $dx^2 = 0$. Since these are also the planes orthogonal to that direction, it is instead the planes generated by the vectors $b^0 = -b^1 = 1$, $b^2 = 0$ and $c^0 = c^1 = 0$, $c^2 = 1$ that have "Gaussian" curvature $-\frac{1}{2}R = 16\pi\rho$.

In these three special cases, the vectors a^μ representing the PT transformations are parallel for all infinitesimal loops, so a^μ for any infinitesimal loop can itself be parallel transported around any other loop and return to its starting value. Suppose we have a finite loop C with base point p, bounding a surface made up of many infinitesimal surfaces, each of which is bounded by the infinitesimal loop c_i. Then the PT transformation $\Lambda(C)$ can be represented by the vector

$$A^\mu = \sum_i a_i^\mu , \tag{8}$$

where the vectors a_i^μ are all parallel transported to the point p before being added up. (It is crucial that the result of this parallel transport is path independent.) Since A^μ is not an infinitesimal vector we cannot simply use the representation

$$\Lambda_\mu{}^\nu(C) = \delta_\mu^\nu + \Omega_\mu{}^\nu \tag{9}$$

appropriate for infinitesimal Lorentz transformations, where

$$\Omega_{\mu\nu} \equiv \sqrt{g}\,\varepsilon_{\mu\nu\rho}A^\rho = \sum_i \sqrt{g}\,\varepsilon_{\mu\nu\rho}a_i^\rho = \sum_i (\omega_{\mu\nu})_i . \tag{10}$$

Instead, we interpret $\Omega_\mu{}^\nu$ as an element of the Lie algebra of SO(2,1), the Lie group of Lorentz transformations, and find

$$\Lambda(C) = \exp(\Omega) . \tag{11}$$

Therefore in the above examples it makes sense to sum the vectors $a_i^\mu = 8\pi T^{\mu\nu}(n_\nu)_i$ over the surface to obtain, on the one hand, the vector A^μ representing the PT transformation for the boundary of the surface, and on the other hand 8π times the total amount of matter passing through the surface. (Again, the "total amount of matter passing through the surface" is only defined because the vectors $T^{\mu\nu}n_\nu$ can be parallel transported in a path-independent way.)

To illustrate how this works let us assume that in the first example given above, where $T^{00} = \rho(x^1, x^2)$, ρ vanishes outside of some bounded region of the constant-time surfaces, where this region represents an extended massive object. The PT transformation for any closed curve that encircles this object is clearly a rotation through an angle equal to the integral over the object of the Gaussian curvature. Since the Gaussian curvature is $8\pi\rho$, this PT transformation is represented by the vector $A^0 = 8\pi M$, $A^1 = A^2 = 0$, where M is the total mass of the object. This vector is also 8π times the energy-momentum vector we would obtain by parallel transporting all the local energy-momentum vectors to the base point of the curve and summing them up. There are three cases for the geometry of the constant-time surfaces: if $0 < M < \frac{1}{4}$, the surface can be described as a cone with vertex angle $2\arcsin(1-4M)$ and a rounded vertex; if $M = \frac{1}{4}$ it can be described as a half-infinite cylinder with a round cap on the end; and if $M = \frac{1}{2}$, it is a round surface with topology S^2. These are the only possible values of M, assuming that the spacetime satisfies the weak energy condition $\rho \geq 0$, since there are no other non-negatively curved surfaces without boundary. In the case $0 < M < \frac{1}{4}$, the cone can be mapped onto a plane with a wedge of angle $8\pi M$ removed and the two boundary lines identified. Exterior to the object this mapping will be an isometry. Inside the object it cannot be. It is important to note that the only constraint on the mapping is that it must be an isometry outside the object, where spacetime is flat. This means that the point mapped to the vertex of the wedge can be chosen completely arbitrarily, as long as it is within the object; *this point in no way represents the "center of mass" of the object.* The notion that the vertex of the wedge represents the center of mass [4,6] is an extrapolation from the case of a single point particle, where the vertex coincides with the only location in which there is mass and is therefore a reasonable place to locate the "center of mass." For extended or multiple sources, however, this reasoning falls through and in general one cannot rigorously define a "center of mass" or "center of mass velocity."

Returning to the discussion of the mapping from the conical surface to the plane with a wedge removed, one can extend this mapping from the constant-time surface to the whole spacetime. Under this mapping, the spacetime exterior to the object will look like Minkowski space with a wedge removed and the two faces of the wedge identified. In this representation the $8\pi M$ rotation experienced by a vector being parallel transported around the object is concentrated at the wedge crossing. This representation will be used throughout the paper for massive objects within which the energy-momentum tensor can be put in the form $T^{00} = \rho$.

We can construct a similar coordinate system for a spacetime containing a transcendent tachyonic object "moving" in the x^1 direction, with $T^{11} = p(x^0, x^2)$. The PT transformation for a curve that encircles this tachyon will be a boost about the x^1 axis through a boost parameter of $8\pi P$, where P is the total pressure of the tachyon measured by parallel transporting to the base point of the curve all the local energy-momentum vectors $T^{\mu\nu}n_\nu$, where n_ν is the normal to the surface bounded by the curve. (It should be noted that because the energy-momentum tensor in these examples is divergence-free with the given coordinate systems, and not merely covariant divergence-free, all surfaces bounded by a given curve are equivalent.) In this case, there are no geometrical constraints on P, although the weak energy condition requires $P \geq 0$. The surfaces x^1 = const are flat outside of the region containing the tachyons; they can be mapped onto (1+1)-dimensional Minkowski space either with a wedge removed from one of the two regions that is timelike related to the origin or, equivalently, with a doubly covered wedge of the same size in one of the two regions that is spacelike related to the origin. The size of the wedge is measured by the parameter of the boost required to take one edge of the wedge to the other edge; that parameter will of course be $8\pi P$. This mapping can be extended uniformly to the whole spacetime, and will be an isometry outside the tachyon. It concentrates the PT transformation for curves that encircle the tachyon at the wedge crossing. Figure 1 shows a spacetime containing a tachyon with $P \approx 0.03$, in the representation where the missing wedge is to the future of the tachyon. The two half-planes bounding the wedge are identified. Also shown are a closed curve encircling the tachyon and the axis vector of its PT transformation.

We can construct a similar representation of the spacetime containing photonic matter moving in the x^1 direction $[T^{00} = T^{01} = T^{10} = T^{11} = \rho(x^1 - x^0, x^2)]$, confined to a bounded region of the $x^0 + x^1$ =const surfaces. The same considerations about PT transformations apply here. The representation of this spacetime in Minkowski space has no missing or doubly covered wedges, but the coordinate system is discontinu-

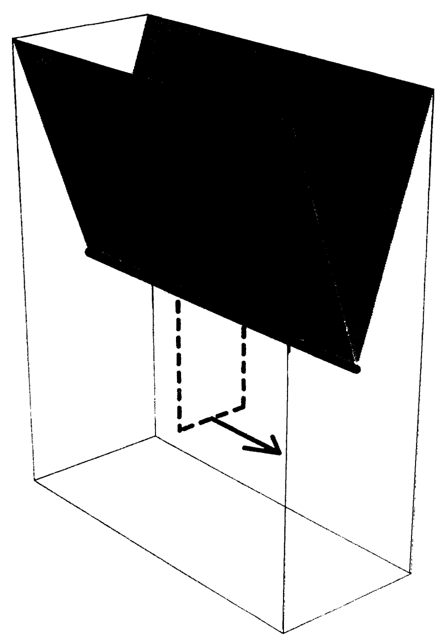

FIG. 1. Part of a (2+1)-dimensional spacetime containing a single transcendent tachyon of total pressure $P \approx 0.03$; the thick solid line is the world line of the tachyon. The vertical coordinate represents time and the horizontal ones space. Outside of the tachyon the metric is the Lorentz metric. The two half planes are identified with each other and the space between them is not part of the spacetime. Also shown is a loop encircling the tachyon (dashed line) and the axis of the PT transformation for that loop; note that it is parallel with the tachyon's world line.

ous on a null half-plane ($x^0 = x^1$ and either $x^2 > 0$ or $x^2 < 0$—we will choose the former). Each point in the real spacetime mapped into this half-plane is mapped to two copies, one of which will be continuous with the mapping of the point's past light cone and the other continuous with the mapping of the point's future light cone. The transformation bringing the former copy to the latter copy is the null transformation represented by A^μ, where $A^0 = A^1 = 8\pi M$, $A^2 = 0$. This is also, of course, the PT transformation for closed curves that encircle the photonic object.

Thus in cases where the total energy-momentum is well defined because it can be calculated uniquely, there is an extremely simple relation between that energy-momentum and the PT transformation for loops that encircle all the matter in the spacetime. This relation has inspired some authors, most notably Deser *et al.* [2,4] and Carroll *et al.* [5,7] to *define* the "total energy-momentum vector" of any (2+1)-dimensional spacetime to be $1/8\pi$ times the vector A^μ representing the PT transformation for a loop that encircles all the matter in the spacetime. Normally in general relativity the notion of total energy-momentum is not defined, precisely because the curvature of spacetime usually makes it impossible to integrate a rank-two tensor over a hypersurface in a covariant way to obtain a vector. Furthermore, due to the exchange of energy and momentum between the matter fields represented by $T_{\mu\nu}$ and the gravitational field, it is impossible to construct any (globally) conserved quantity from the $T_{\mu\nu}$ field, conservation generally being considered a necessary characteristic of any good notion of total energy or momentum. It is possible in asymptotically flat spacetimes to define a conserved total energy-momentum vector P^μ representing not only the matter field $T_{\mu\nu}$ but also the energy and momentum of the gravitational field. P^μ will be Lorentz covariant with respect to coordinate transformations that reduce to Lorentz transformations at spacelike infinity [8]. However, in 2+1 dimensions only empty spacetime is asymptotically flat, provided that the weak energy condition is satisfied.

The new parallel-transport based definition of total energy-momentum in 2+1 dimensions attempts to overcome this difficulty, but it is problematic in several respects. First, there is the problem of choosing the loop. Normally, we would like quantities that relate to an entire system not to be defined with reference to arbitrary points within the system. If one chooses to call such points "observers" (see [5]), this means that global quantities should not be observer dependent. For instance the special relativistic total energy-momentum is equally well defined at any point in Minkowski spacetime or, if one prefers, on the world line of the center of mass, which is uniquely defined. The normal (3+1)-dimensional general-relativistic total energy-momentum is a vector at spacelike infinity, which is also well defined in asymptotically flat spacetimes. As emphasized earlier, however, the PT transformation is a Lorentz transformation acting on the tangent space of the base point of a given closed curve (and can be represented as a vector in that tangent space). In order to be said to "encircle" an entire system, we could require a curve to pass only through empty spacetime, and to be the boundary of an orientable surface through which all the matter in the spacetime crosses in the same direction exactly once. For tachyons, this requirement is problematic since one cannot really say in which direction tachyonic matter is "moving"—in a system with two separate tachyons, it would be hard to say whether they passed through a given surface in the same direction or in opposite directions. Even if one barred tachyonic matter by imposing the dominant energy condition, there are still many possible curves with many possible base points to choose from that satisfy the above requirement. One might hope that, as in the three examples described above, every closed curve through a given base point would give the same "energy-momentum" vector, and that the resulting vector field would be a parallel vector field. Then each "observer" (that is, each point) would observe a well-defined total energy-momentum, and observers at different points would at least agree parallel-transportwise. But this is not true if there is more than one separate object in the spacetime, since through any given base point there will be curves *not* deformable to each other in empty spacetime, each of which satisfactorily encircles all the matter. The PT transformations for such curves will in general disagree, and are not even necessarily similar. (We will say that two linear transformations T_1 and T_2, on the same or on different spaces, are *similar* if there is a linear mapping S such that $T_1 = ST_2S^{-1}$.) For instance, Fig. 2 is a diagram of a spacetime containing three objects in relative motion, with removed wedges shaded. The Lorentz transformations associated with crossing the wedges of particles 1, 2, and 3 (in the diagram's counterclockwise direction) are Λ_1, Λ_2, and Λ_3, respectively. Since the objects are in relative motion, these Lorentz transformations do not commute. Figure 2(a) shows a closed curve, with base point marked by the cross, whose PT transformation is $\Lambda_2\Lambda_3\Lambda_1$. Figure 2(b) shows another closed curve in the same spacetime that also encircles all three objects, but in a different order. (Note that in the three-dimensional spacetime, this curve does not intersect itself, and is the boundary of a surface through which all three particles pass exactly once in the same direction.) The PT transformation for this curve is $\Lambda_3\Lambda_2\Lambda_1$, which is not similar to $\Lambda_2\Lambda_3\Lambda_1$.

One might be able to remove the ambiguity in choice of curve by requiring that the closed curve not only encircle all the matter in the spacetime, but also be deformable to spacelike infinity through empty spacetime. The purpose of moving the curve to spacelike infinity would be to ensure that, no matter how the particles interacted or moved, one could always translate the curve forward in time (in order to demonstrate conservation of every and momentum, for instance). In Sec. IV, we will make use of the notion of a curve at infinity, and in particular study its PT transformation. The definition of "deformable to spacelike infinity" can be made rigorous (see Sec. IV), but it remains to be proved that in a general spacetime such a curve is unique (up to deformation through empty spacetime). For instance, it is unique in spacetimes containing noninteracting nontachyonic objects whose wedges do not intersect, since those wedges can be put in a circular ordering that must be followed by all curves that are deformable to spacelike infinity. Even if it is unique in a general spacetime, the vector representing its PT transformation is not a vector in the tangent space of any point, or a vector field defined at all points, but rather something like a different vector for each spacelike direction. The "energy-momentum vector" is then essentially a parallel vector field at spacelike infinity. To be fair, the standard (3+1)-dimensional general-relativistic total energy-momentum is also a parallel vector field at infinity. However, in that case the structure of spacelike infinity is well understood, because the spacetime is required to be asymptotically flat. In the general (2+1)-dimensional case the structure of spacelike infinity is not at all well understood, and until it is this definition of the total energy-momentum vector cannot stand on its own. If this definition can be made to make sense, then such energy-momentum vectors for distant subsystems can be combined to obtain the energy-momentum vector for the entire system. Because of the non-Abelian nature of the group SO(2,1), however, this combination will not be linear, and the problem then arises that the energy-momentum vector will not be additive. The usefulness of the fact that this definition gives us a vector becomes questionable if we cannot add these vectors together.

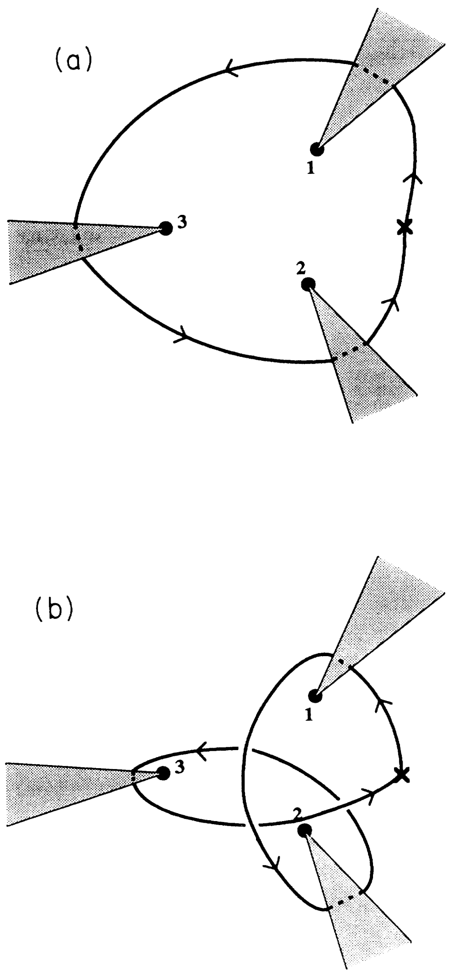

FIG. 2. Diagrammatic representation, in which the time coordinate has been suppressed, of a (2+1)-dimensional spacetime containing three masses in relative motion. The masses are represented by the large dots; the shaded areas are not part of the spacetime, and the two rays bounding each shaded area are identified with each other. Different nonintersecting loops in this spacetime, with the same base point marked by the cross, are shown in (a) and (b). As explained in the text, the PT transformations for these two loops are different, although both loops satisfy the condition of being the edge of a surface that the world lines of all the masses cross through.

This brings us to an even more problematic aspect of the PT-transformation-based definition of the "total energy-momentum vector:" it is not analogous to any notion of energy-momentum in physical (3+1)-dimensional gravity. It is based on the fact that Lorentz transformations in 2+1 dimensions can be represented by vectors and on the relation $a^\mu = 8\pi T_{\mu\nu} n_\nu$. Since neither of these conditions holds in 3+1 dimensions, there is no analogous notion for physical systems. The physical (3+1)-dimensional definition of the total energy-momentum vector has no analogues in 2+1 dimensions, so there can be no relevant notion of total energy-momentum in 2+1 dimensions. PT transformations in 2+1 dimensions have direct analogues in 3+1 dimensions, and can therefore help us understand the geometrical properties of physical systems, particularly those containing cosmic strings. Whether or not it is well defined, this "total energy-momentum" in 2+1 dimensions cannot help us understand anything about energy or momentum in (3+1)-dimensional systems and is therefore unproductive.

We will therefore not adopt the terminology of "total energy-momentum" to refer to PT transformations. Nonetheless PT transformations are clearly useful tools for understanding the geometrical properties of spacetimes, and much of this paper is devoted to their study. There is, however, another concept, that we shall refer to as the "holonomy" of a closed curve, that is similar to the PT transformation but even more useful because it allows us to see distinctions that the PT transformation misses. One of the main purposes of this paper is to emphasize that a spacetime containing a Gott time machine is *not* similar to a spacetime containing a tachyon. The idea that they are similar is due to a too-trustful attitude toward PT transformations. The PT transformation for a loop around a Gott pair of particles can in fact be the same as the PT transformation for a loop around a tachyon. But those PT transformations do not tell the whole story about the exterior geometries of the spacetimes. As Carroll *et al.* point out, "parallel transport of a spinor around a single tachyonic particle is not equivalent to parallel transport around the Gott two-particle system, although parallel transport of an SO(2,1) vector does not distinguish between the two cases" [5]. This is because the spinor, unlike the vector, is affected by the rotation through a full 2π induced by parallel transport around a Gott pair, but not induced by parallel transport around a tachyon. We know that parallel transport around a loop that encircles a Gott pair induces a rotation through 2π because we can contract the loop continuously to a vanishingly small loop, which obviously has a PT transformation at the identity. By measuring the PT transformation of each intermediate loop in the contraction we obtain a path C_1 in SO(2,1) from the PT transformation of the original loop to the identity. Similarly, by contracting a loop with the same PT transformation but encircling a tachyon, we obtain another path, C_2, connecting the same points in SO(2,1). However, it can be shown (see Sec. III) that these two paths are not homotopic to each other [5] thus exhibiting the multiple connectedness of SO(2,1). The fundamental group of SO(2,1) is isomorphic to the integers, and its generator is what we have been informally calling a "2π rotation"—a path from the identity to the identity consisting of rotations of all angles from 0 to 2π. The path $C_2C_1^{-1}$ from the identity to the identity is homotopic to this generator; therefore, while the PT transformation for a loop encircling a tachyon is a boost, for a loop encircling a Gott pair we might say it is a 2π rotation plus a boost.

A natural and systematic language in which to express such distinctions is provided by the concept of the holonomy of a loop [5], which we shall now define. We remind the reader that the universal covering space of SO(2,1), denoted here $\overline{\text{SO}}(2,1)$, is the set of equivalence classes under homotopy of paths in SO(2,1) that end on the identity. The *holonomy* of a loop that is contractible to zero is defined to be the element of $\overline{\text{SO}}(2,1)$ that is the path in SO(2,1) made up of the PT transformations of the intermediate loops during the contraction. A problem arises if the loop can be contracted to zero in two different ways that give homotopically distinct paths in SO(2,1); this could only happen if the two contractions were not homotopic to each other, which could only happen if the spacetime had a nontrivial second homotopy group π_2. Therefore in order for both the existence and the uniqueness of a loop's holonomy to be assured, both the first and the second homotopy groups of the spacetime must be trivial.

Unlike SO(2,1), $\overline{\text{SO}}(2,1)$ contains elements that are not the exponential of any element of its Lie algebra [which is of course the same as the Lie algebra of SO(2,1), namely, the space of two-forms]. Thus while the exponential map for SO(2,1) is onto but not one to one, the exponential map for $\overline{\text{SO}}(2,1)$ is one to one but not onto. This indicates another failure of the parallel transport-based definition of the total energy-momentum vector: some closed curves, such as those surrounding a Gott pair, have holonomies that are not representable by vectors because they are not the exponential of any element of the Lie algebra. This forces us either to project the holonomy down to SO(2,1) to obtain the PT transformation and then find the vector representing that transformation, which entails losing information about the geometry of the spacetime without gaining much in return, or to say that spacetimes like the Gott solution have an undefined total energy-momentum, which seems like a shame considering that they contain nothing but ordinary matter—and which indicates a failure of the definition, not something wrong with the spacetime.

Carroll *et al.* [5] give an elegant geometrical description of SO(2,1) and $\overline{\text{SO}}(2,1)$ in terms of anti–de Sitter space and its universal covering space, respectively. Picking an arbitrary point in anti–de Sitter space to represent the identity element of SO(2,1), the tangent space can be identified with the vector representation of the Lie algebra of SO(2,1). Under this identification, the exponential map in the two spaces can be identified, giving a one-to-one mapping between SO(2,1) and anti–de Sitter space. This mapping can be extended to a one-to-one mapping between $\overline{\text{SO}}(2,1)$ and universal anti–de Sitter space. (For the exact mapping, see [5].) Inextendible geodesics through the origin in universal anti–de Sitter space are mapped to paths in $\overline{\text{SO}}(2,1)$ of the form $\exp(\lambda\Omega)$ ($\lambda \in \mathbb{R}$), where Ω is an element of the Lie algebra of $\overline{\text{SO}}(2,1)$. The Penrose diagram of universal anti–de Sitter space is shown in Fig. 3. (Again, see [5] for the exact form of the conformal coordinates.) The angular coordinate is suppressed in this diagram. The

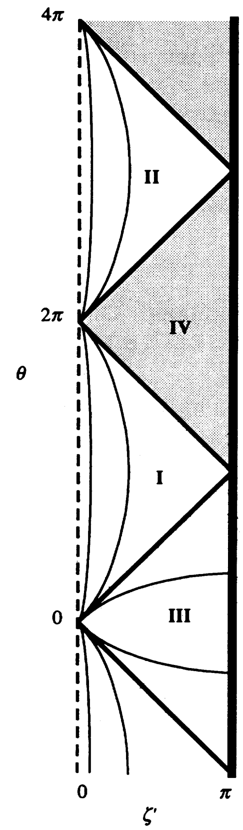

FIG. 3. Part of the Penrose diagram of (2+1)-dimensional universal anti–de Sitter space, representing the Lie group $\overline{\text{SO}}(2,1)$. θ is the time coordinate, ranging from $-\infty$ to $+\infty$; ζ' is the radial coordinate, ranging from 0 to π. The angular coordinate is suppressed. The dashed line $\zeta' = 0$ is the coordinate singularity at the origin of polar coordinates. The heavy solid line $\zeta' = \pi$ represents spacelike and null infinity. The identity of $\overline{\text{SO}}(2,1)$ is represented by the point $\theta = \zeta' = 0$. Three geodesics through the origin, two timelike and one spacelike, are shown. The shaded regions are points not reachable by geodesics through the origin, i.e., points with no inverse images under the exponential map. The solid lines separating the different labeled regions represent null surfaces.

origin is at $\theta = 0$, $\zeta' = 0$. The line $\zeta' = 0$ is a coordinate singularity at the origin of the polar coordinates; points on it represent the pure rotations covering the element of SO(2,1)

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} . \qquad (12)$$

The line $\zeta' = \pi$ represents spacelike and null infinity. Future and past timelike infinity are disjoint points and not represented on the diagram. A few geodesics through the origin are shown. Region I on the diagram contains the rotations with $0 < \theta < 2\pi$. The holonomy for loops around a single massive particle in an open spacetime is in this region, or at the point $\zeta' = 0$, $\theta = 2\pi$. Region II contains the rotations with $2\pi < \theta < 4\pi$. Region III contains the boosts; the holonomy for loops around a single tachyon lies in this region. Region IV and the other shaded region are not accessible from the origin by geodesics, although they cover the same region of anti–de Sitter space as region III. Carroll *et al.* [5] showed that the holonomy for a loop around a Gott pair lies in region IV. Thus, while the PT transformation for such a loop is a boost, the holonomy is not. We will refer to this diagram in the next section to identify holonomies in various specific spacetimes.

III. EXACT SOLUTIONS CONTAINING CLOSED TIMELIKE CURVES

As in 3+1 dimensions, it is not hard to invent (2+1)-dimensional spacetimes that contain CTC's. The easiest kind involve no matter; these are Minkowski space with points identified under the action of a discrete group of isometries. Although such examples may be considered artificial, they are nonetheless helpful in trying to understand the geometries of CTC-containing spacetimes. Here we will consider groups of isometries that are isomorphic to the integers; that is, we will constrain the solutions to be merely singly periodic, rather than doubly or triply periodic. Spacetimes constructed by this method can be classified according to the generator of the group of isometries. Since an isometry of Minkowski space can be decomposed into a Lorentz transformation about some axis plus a translation parallel to that axis, there are nine possibilities: timelike translation; rotation; rotation with timelike translation; boost; spacelike translation; boost with spacelike translation; null translation; null transformation; and null transformation with null translation. Of these possibilities, the rotation and the spacelike translation definitely do not produce spacetimes containing CTC's, since a point will be identified with other points that are spacelike related to it. Similarly, the null translation, null transformation, and combination will identify points only with other points that are null or spacelike related to them, not timelike related. This leaves the timelike translation, which produces probably the simplest example of a spacetime containing CTC's, the rotation with timelike translation, in which the rotation makes it more difficult for CTC's to form but they do so anyway, the boost, producing Misner space, which is well known for containing CTC's, and the boost with spacelike translation, producing Grant space, the object of some recent investigation [9,10], in which the spacelike translation makes it more difficult for the CTC's of Misner space to form, but they form anyway. We will consider each of these spacetimes in turn.

The simplest is the one in which points of Minkowski space are identified under a timelike translation, so that in the rest frame of that translation the time coordinate is cyclic. Every point of the resulting space is then obviously timelike separated from itself. Furthermore, any closed curve in this spacetime has a trivial PT transformation. (Note that, since all of these spacetimes are multiply connected, loops in them do not have well-defined holonomies.)

Less trivial is the example of Misner space, in which points are identified under the action of a boost. The displacement between identified points is now timelike if and only if the points are spacelike related to the boost axis. So while CTC's do not go through every point in the spacetime, there are CTC's arbitrarily far from the axis. The axis itself cannot be included in the spacetime, which is therefore multiply connected; the CTC's are among the loops not homotopic to zero. Clearly the PT transformation for any loop will be the boost that brings the base point of the curve to its identified point in Minkowski space. Since this is a power of the boost used to construct the spacetime, in some sense we can count the number of times the loop "crossed" the identification.

More complicated than Misner space is Grant space, in which the identification is performed under the action of a boost plus a spacelike translation parallel to the axis of the boost. Again, the points threaded by CTC's are the ones that are spacelike related to the axis of the boost. However, due to the additional spacelike displacement, CTC's passing through points near the axis must loop "around" the identification many times. This is in contrast with Misner space, which has no intrinsic distance scale so that all points spacelike related to the axis are essentially equivalent. Although the axis can now be included in the spacetime, the spacetime is still multiply connected because the axis itself now has the topology S^1, and is not homotopic to zero. Thus the spacetime has topology $\mathbb{R}^2 \times S^1$, so its fundamental group is isomorphic to the integers. CTC's passing close to the axis are in elements of the fundamental group corresponding to large integers. The PT transformation for any loop is the boost used in the construction of the spacetime, to the power of the integer corresponding to its fundamental group element; loops homotopic to zero will be in the identity of the fundamental group and will have the identity as their PT transformation.

The last CTC-containing spacetime constructed by identifying points in Minkowski space is in some ways a counterpart to Grant space. The isometry is a rotation through an angle θ plus a timelike translation chosen parallel to the axis of rotation. All the points identified with a given point p in Minkowski space lie on a helix that winds around the axis of rotation. Except for a finite number, all of these points will be in either the future or the past light cone of p, because their spatial separation from p (in this reference frame) is bounded while their temporal separation goes to plus or minus infinity. So every point in this spacetime lies on a CTC. The PT transformation for a CTC that joins p to its nth identified point q in Minkowski space (where we order the identified points in the obvious way) is a rotation of angle $n\theta$. We will now show that, if $n\theta$ is not a multiple of 2π, then this CTC will not be deformable to spacelike infinity. We first note that a rotation of $n\theta$ in the spacetime about the spacetime's axis, brings p to a point directly to the past of q. Now in a deformation to spacelike infinity we can assume without loss of generality that p (and hence q) moves out on a spacelike geodesic orthogonal to the axis, movements in time and tangentially around the axis are irrelevant because q will move in exactly the same way, leaving the interval between them unchanged. So if p is not directly to the past of q, that is, if $n\theta$ is not a multiple of 2π, then a movement out to spacelike infinity will increase their spatial separation without bound, while their temporal separation remains fixed. This spacetime therefore has the curious (and telling) property that, while there are CTC's arbitrarily far from the axis, a CTC cannot be deformed to spacelike infinity if its PT transformation is a rotation (the only possibilities being rotations and the identity). If θ is a rational number times 2π, there is a "single" CTC at spacelike infinity, not counting repeated traversals of the same curve; that CTC is the one for the smallest positive n such that $n\theta$ is a multiple of 2π. If on the other hand θ is not a rational number times 2π, then there will be no CTC's with trivial PT transformation, and the spacetime is very odd indeed, there are CTC's at all spatial distances, but for each CTC there is a distance past which none of the other CTC's are homotopic to it; thus there is no "single" CTC at spacelike infinity. One lesson we can draw is that the existence of CTC's, and particularly CTC's at infinity, is linked not to the PT transformation of a curve that loops around the spacetime once, the "total energy-momentum" of the spacetime, but rather to the PT transformations of the CTC's themselves. This example is a good illustration of the techniques we shall use in the next section to prove the main result of this paper.

While such model spacetimes are interesting to play with, they are artificial from a physical viewpoint in the sense that their interesting geometrical properties are built directly into them, rather than being caused by physical matter and interactions. Therefore it is interesting to try to find CTC-containing (2+1)-dimensional spacetimes containing only what would be considered "physical" matter [that is, matter somehow analogous to physical (3+1)-dimensional matter] and whose geometry arises from the gravitational effects of that matter rather than, say, from a manipulated topology.

The first example of such a (2+1)-dimensional spacetime was discovered [2] as soon as (2+1)-dimensional gravity first came under intensive investigation. The solution contains a single point mass with spin. Since then similar solutions have been found for extended, rapidly rotating particles, although these solutions are under some dispute [11,12]. The geometry of a spacetime containing one stationary such source, of mass M and angular momentum J, is similar to that of a spacetime containing a single stationary spinless mass; external to the source it may be described by Minkowski space with a wedge of angle $8\pi M$ removed and the two faces of the wedge identified. Unlike in the case of a spinless source, however, the transformation P bringing a point x on one face to its identified point $P(x)$ is not simply the rotation $R_{8\pi M}$, but that rotation plus a translation in time of $8\pi J$: $P(x) = R_{8\pi M}(x) + 8\pi J\hat{t}$. Thus a future-directed timelike curve that circles the particle in the correct direction gains a constant jump backward in the time coordinate regardless of how near or far the curve is from the source. For a point particle one can approach the particle arbitrarily closely, making the trip around its circumference arbitrarily short, so there are clearly CTC's close to the particle. (Note that, while it helps, a nonzero mass is not necessary for the solution to contain CTC's.) In fact, CTC's pass through every point in this spacetime (contrary to a claim in [11]). From a point far away from the particle, it is not necessary to circle the particle at that distance; instead the curve can first head straight for the particle, circle closely around it many times in order to go backward in the time coordinate an arbitrary amount, and have enough time to return to the base point. However, there are no CTC's of large spatial size; if the entire CTC is at outside of a distance $r_0 = 4J/(1-4M)$ from the source, the time necessary to get once around the source is larger than the jump backward in the time coordinate gained. Thus the spacetime does not contain CTC's at arbitrary spatial distances, let alone CTC's that can be deformed to spacelike infinity. The holonomy for a curve that encircles the source is simply a rotation of angle $8\pi M$, just as for a spinless particle of mass M.

Gott subsequently showed [3] that there are CTC's in a spacetime containing two massive particles moving past each other at high velocity. If the particles are of equal mass M and moving in opposite directions each with speed $\beta < 1$ relative to the "laboratory" reference frame, then it is necessary and sufficient that

$$\beta > \cos(4\pi M) \tag{13}$$

for there to exist CTC's in the system. These CTC's encircle the two particles in the direction opposite to their relative motion, just as the CTC's in the spinning particle solution encircle the particle in the direction opposite to its spin. Figure 4 is a to-scale picture of such a spacetime and one of its CTC's. This CTC is geodesic except at the two points marked by dots, where it crosses the wedges of this representation. Also shown is the axis of the PT transformation of this curve, more about which shortly.

In the "laboratory" reference frame of the Gott spacetime, the frame related to each particle rest frame by a boost of magnitude χ, the wedges that are cut out of Minkowski spacetime are warped in such a way that, by crossing them in the direction opposite to the direction of motion of the particle in the given reference frame, one jumps backward in the time coordinate, as one does when crossing the wedge of a spinning particle; one also jumps forward in the direction in which one was traveling. Unlike in the spinning-particle case, however, the wedge for each individual particle is scale invariant, so the magnitude of the jump backward in the time coordinate is not constant but proportional to the distance from the particle. We can therefore enlarge a given CTC equally in all directions. The only distance that does not scale with the size of the CTC is the spatial separation between the two particles, which stays constant and therefore diminishes in relative magnitude as we enlarge the CTC. Since this spacelike interval hinders the timelike curve from returning to its base point, its reduction in relative magnitude will not endanger the existence of the CTC at large size. This can be shown rigorously by writing out the coordinates of a particular CTC and showing that it can be deformed to arbitrarily large size. While there is thus in some sense a "CTC at infinity" in this spacetime, Cutler showed [13] that not every point in the spacetime is the base point of a CTC. In fact, there are complete, edgeless, achronal surfaces not asymptotically null in the spacetime, both to the future and to the past of the CTC-containing region and the spacetime has a Cauchy horizon. So the "CTC at infinity" is somehow limited to a subregion of spacelike infinity.

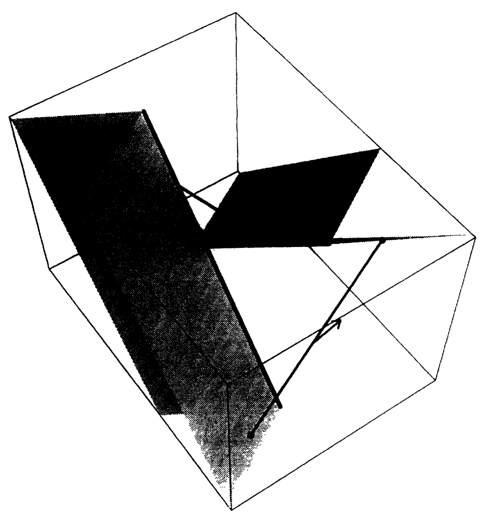

FIG. 4. A Gott two-particle spacetime. The view is from above, that is, from the future looking toward the past. Again, the vertical direction represents the time coordinate and the horizontal directions the space coordinates. The thick solid lines represent the world lines of the particles. The two half-planes coming off of each world line are identified with each other. Thus the two large dots shown on the two half-planes coming off of each world line represent the same point in the spacetime. The two wedges between these half-planes do not represent parts of the spacetime. Outside the particles, the spacetime is flat, and the metric in the coordinate system shown is the Lorentz metric. The two line segments shown are both timelike and together form a CTC. The arrow represents the axis of the PT transformation for this CTC with base point at the base of the arrow; it is spacelike.

Soon after Gott presented this solution, Deser *et al.* pointed out [4] that the PT transformation for a CTC in this spacetime was a boost, and furthermore that the condition stated above for the existence of CTC's in the spacetime coincided exactly with the condition for the PT transformation to be a boost for loops that encircle the two particles the way the CTC's do. Thus the axis of the PT transformation for the CTC shown in Fig. 4 is spacelike. Carroll *et al.* explained [5] that this meant that the holonomy of the CTC was in region IV of $\overline{SO}(2,1)$ (see Fig. 3). This holonomy is the product of the holonomies for loops that go around each source separately, and those holonomies, as explained in the previous section, are both in region I of $\overline{SO}(2,1)$. Since rotation through a positive angle means moving in the future-directed timelike direction in the universal anti–de Sitter space representation of $\overline{SO}(2,1)$, the holonomy could not be in region III of $\overline{SO}(2,1)$. This is part of the basis for our insistence that, though they can both be encircled by closed paths with boost PT transformations, tachyons and Gott pairs are quite distinct situations, both physically and geometrically. Physically, the Gott pair is made of matter satisfying the weak, strong, and dominant energy conditions, matter as normal as one could hope to find in (2+1)-dimensional systems. Geometrically, as Grant conjectured [9] and Laurence proved [10], the CTC-containing region of the Gott solution is isometric to half of the CTC-containing region of multiply connected Grant space. (Although the Gott spacetime is simply connected, its CTC-containing region is not.) A tachyon's spacetime can be described as Minkowski space missing a wedge; the CTC-containing section of Grant space is more like just a wedge. Both spacetimes require boost identifications, but there the similarity ends. (In fact Grant space requires an identification by a boost plus a translation, which both the PT transformation and the holonomy miss.)

This apparent similarity between the exterior geometries of a Gott pair of particles and a tachyon is part of the basis for an argument by Deser *et al.* [4] that Gott's (2+1)-dimensional solution is unphysical. It is important to note that the use by Deser *et al.* of the term "cosmon," referring both to particles in 2+1 dimensions and to cosmic strings in 3+1 dimensions, obscures the distinction between (3+1)-dimensional systems, whose physicality we can argue about, and (2+1)-dimensional systems where the adjective "physical" can only be used in a very limited sense. According to their argument, Gott pairs of cosmic strings are supposed to be unphysical because Gott pairs of particles in 2+1 dimensions resemble in a particular way tachyons in 2+1 dimensions, which are presumably objectionable because tachyons in 3+1 dimensions are thought not to exist. (The proper analogue in 3+1 dimensions may be tachyonic strings, which are also thought not to exist.) There may be reasons why Gott pairs of cosmic strings cannot exist in Nature, but they certainly are not the result of any problematic aspects of Gott pairs in 2+1 dimensions. As we have seen, their claim is not true that the exterior wedge identification in the most general case "could be equivalent to a boost rather than to a rotation, in which case CTC's are always present;" the counterexample is particularly relevant: the solution for a single tachyon in 2+1 dimensions contains no CTC's. In addition, as a few examples show, there is nothing objectionable about a holonomy in regions III or IV of $\overline{\mathrm{SO}}(2,1)$. Figure 5 shows the solution for an elastic collision between two masses A and B. (This spacetime is a combination of the decaying-particle solution discovered by Carroll *et al.* [7] and its time-reversed image.) Although this is clearly an acceptable solution, its geometry to the future and past of the collision event resembles that of a tachyon. In particular, a closed curve that passes directly to the future and directly to the past of the collision has a holonomy in region III of $\overline{\mathrm{SO}}(2,1)$. Another example of an unobjectionable spacetime containing a curve with holonomy in region III is provided by any spacetime with two identical masses in relative motion (not necessarily satisfying the criterion for the existence of CTC's) and any curve that encircles the two masses in opposite senses. Finally and most importantly, Carroll *et al.* showed conclusively [5] that one cannot object to the Gott solution on the basis of its holonomies. They proved that in a closed universe it is possible *out of static initial conditions* to produce a Gott pair, that is, a pair of particles encircled by a curve with holonomy in region IV; as proved by 't Hooft [14], this Gott pair will not produce CTC's because the closed spacetime is too small.

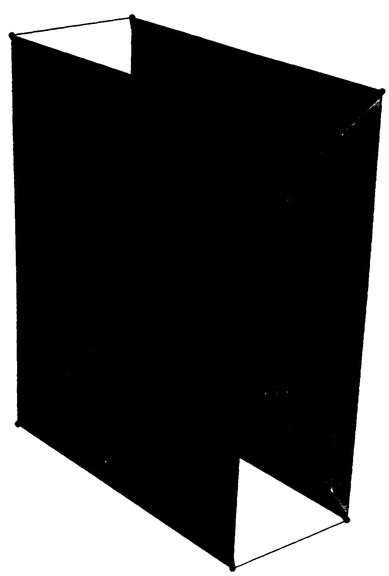

FIG. 5. A (2+1)-dimensional spacetime in which two masses A and B bounce elastically. Again, time is represented by the vertical dimension and the spacetime is locally flat with a Lorentz metric outside of the sources. The planes divide the space of the figure into three disjoint regions. Of these, only the near and far regions are part of the actual spacetime, while the region sandwiched between them is excised. At any time, except the time of the collision, the position of particle A is represented by two identified points, and similarly for particle B; hence each particle has two world lines in this representation.

The other problem Deser *et al.* present with regard to the Gott spacetime is that it contains CTC's at spacelike infinity; this is supposed to be an unacceptable boundary condition. We wonder, however, how they know so much about boundary conditions at spacelike infinity. In our own Universe we do not know what spacelike infinity looks like (if it exists) since we have not seen it yet. We certainly have no way of knowing whether or not there are CTC's there. The working physicist is, of course, free to impose simple and convenient boundary conditions (e.g., asymptotic flatness) on a system in order to isolate and understand the processes occurring within it. But boundary conditions are tools of physicists, and they should not be confused with laws of physics. There may be laws of nature that restrict the possible structures of spacelike infinity, and even that prohibit CTC's there, but in the absence of evidence such laws should not be postulated *ad hoc*.

Returning to the description of CTC-containing exact solutions in 2+1 dimensions, the Gott solution has analogues containing two photonic particles [15] or two tachyons instead of two massive particles. The structure of these solutions is similar to that of the Gott solution; in particular they share the property that all the CTC's may be deformed to spacelike infinity. They also share the property that there are CTC's in the spacetime exactly when the holonomy for loops that encircle the two particles (the same way the CTC does) is in region IV of $\overline{\mathrm{SO}}(2,1)$. [For the two-tachyon solution it is clearly crucial to make the distinction between regions III and IV of $\overline{\mathrm{SO}}(2,1)$.] In the two-tachyon case the condition for the existence of CTC's is

$$\beta < \cosh(4\pi P)\,, \tag{14}$$

where the tachyons travel in opposite directions, each with speed $\beta > 1$ relative to the "laboratory" reference frame, and where P is the total pressure of each tachyon. The condition in the two-photon case is simply

$$4\pi M > 1\,, \tag{15}$$

where M is the energy of each oppositely directed photonic particle in the "laboratory" reference frame.

Kabat showed [6] that there were CTC's in some systems of n spinless particles with equal masses M, moving in the laboratory frame with equal speeds $\tanh\chi$ in a pattern with $2\pi/n$ rotational symmetry. The condition for the existence of CTC's in such a spacetime is

$$\cosh\chi\sin(\pi/n)\sin(4\pi M) + \cos(\pi/n)\cos(4\pi M) > 1\,. \tag{16}$$

(Note that for $n = 2$ this reduces to the condition for the Gott spacetime.) This is also the condition for the holonomy for loops that encircle the particles (the same way the CTC's do) to be in region IV of $\overline{\mathrm{SO}}(2,1)$. Figure 6 shows such a spacetime with $n = 4$, as well as one of the CTC's and the axis of its PT transformation at several base points. As in Fig. 4, this axis is spacelike.

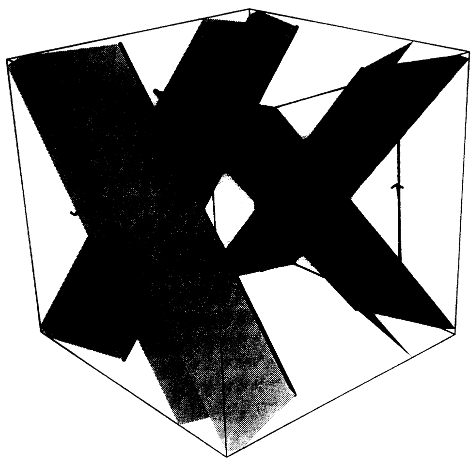

FIG. 6. A Kabat four-particle spacetime. A CTC is shown, along with the axis of the PT transformation for the CTC at three different base points. (One of these axes is mostly obscured by a wedge.) These axes agree when parallel transported along the CTC. As one can see from the picture, they are spacelike. Again, large dots represent points where the CTC crosses the wedges.

Finally, by combining the spinning particle solution with the Gott solution, it is possible to create a borderline solution containing CTC's deformable to spacelike infinity, the holonomy for which is not in region IV of $\overline{\mathrm{SO}}(2,1)$ but rather on the boundary between region I and region IV. We begin with the Gott solution, containing two point particles, each of mass M and each moving relative to the "laboratory" reference frame with speed $\beta = \cos 4\pi M$, and impact parameter $2d$. Although this spacetime contains no CTC's, it is on the edge of satisfying the criterion for their existence. The PT transformation Λ for loops that encircle both strings is a null transformation. We now endow each particle with spin angular momentum J in the same direction as their orbital "angular momentum." For any J CTC's will appear around each particle separately, but these will not go (deform) to infinity. As a simple calculation will show, however, at $J = d\sqrt{1-\beta^2}/4\pi$ closed null geodesics appear, and these go to infinity. For $J > d\sqrt{1-\beta^2}/4\pi$ the spacetime contains CTC's that go to infinity, for which the PT transformation remains the null transformation Λ.

IV. CLOSED TIMELIKE CURVES AT INFINITY

Regardless of their relevance to questions of "physicality" for (2+1)-dimensional spacetimes, parallel-transport transformations are obviously useful tools for understanding the geometrical structure of such spacetimes. For instance, they provided the first evidence of a similarity between the CTC-containing regions of the Gott spacetime and Grant space [9], and helped elucidate the actual isometry relating these two regions [10].

As discussed in Sec. II, one of the main complications is using holonomies to understand spacetimes is the problem of the choosing the loop for which to calculate the holonomy. It seems to us that if one is trying to understand holonomies in CTC-containing spacetimes, as we are here, then it makes the most sense to concentrate on the holonomies for the CTC's themselves. This choice has the two advantages of being unambiguous and being more fruitful than an arbitrary other choice.

Kabat put forward the following conjecture: "Closed timelike curves exist surrounding a system of spinless particles only if their total momentum is spacelike" [6]. While this conjecture is ambiguous as stated due to the use of the phrase "total momentum," we would interpret it as saying that the PT transformation for a CTC surrounding a system of spinless particles is a boost. We would add the stronger conjecture that the holonomy for a CTC surrounding a system of spinless particles is in region IV of $\overline{\mathrm{SO}}(2,1)$.

In all of the known CTC-containing systems of spinless particles the CTC's are all deformable to spacelike infinity. Furthermore, in all the known CTC-containing spacetimes in which the CTC's are deformable to spacelike infinity, including both the multiply connected ones and the matter-containing ones, the PT transformation for those CTC's is either the identity, a null transformation, or a boost. We submit that if the Kabat conjecture is true it is because CTC's in systems of spinless particles are necessarily deformable to spacelike infinity. It would follow immediately from a proof of this statement, together with the theorem we shall prove in this section, that the PT transformation for any CTC in a system of spinless particles is either the identity, a null transformation, or a boost. Perhaps it can be proved that the PT transformation for a CTC in a system of spinless particles cannot be the identity or a null transformation.

This section is principally devoted to proving the following theorem: the PT transformation for a non-self-intersecting CTC at infinity is not a rotation. This theorem is proved in all generality for (2+1)-dimensional spacetimes, making no assumptions about their matter content or topology. We believe that this theorem lies at the heart of the relationship between CTC's and nonrotational PT transformations first noticed by Deser *et al.*

In order to prove the statement rigorously, it is first necessary to give several technical definitions for the notions we have been using loosely up until now. A connected open set is said to be *flat* if it is isometric to some open subset of Minkowski space. This subset and the corresponding isometry will be unique up to isometries of Minkowski space. A set is said to be *locally flat* if every point in it has a flat neighborhood. On an open set this is equivalent to the condition $R_{\mu\nu} = 0$, which is equivalent to the condition $T_{\mu\nu} = 0$. A (locally flat) path $p(\lambda)$ ($\lambda \in I$), where I is an interval of the real line, is said to be *segmentwise flat* if for every compact subinterval I' of I, the image of I' under p has a flat neighborhood. For instance, in a spacetime such as the Gott, Kabat, or tachyon solutions that is constructed from Minkowski space by removing wedges, any path that does not cross any wedges is segmentwise flat. Note that if p intersects itself, that is if $p(\lambda_0) = p(\lambda_1)$ for $\lambda_0 \neq \lambda_1$, then the PT transformation for p restricted to $[\lambda_0, \lambda_1]$ must be trivial for p to be segmentwise flat.

Unless I is itself compact, segmentwise flatness of p does not imply that the entire image of I under p has a flat neighborhood. For instance, let $I = [0,1)$ and suppose that $\lim_{\lambda \to 1} p(\lambda) = p(0)$. If p encircles a point mass then any neighborhood of $p(I)$ will also encircle it and will therefore not be flat. However, given a segmentwise flat path $p(\lambda)$ we can define a path $q(\lambda)$ ($\lambda \in I$) in Minkowski space, unique up to isometries of Minkowski space, such that for any compact subinterval I' of I and flat neighborhood N' of $\{p(\lambda) : \lambda \in I'\}$ there is an open subset M' of Minkowski space and an isometry $P' : N' \to M'$ such that $q(\lambda) = P'(p(\lambda))$ for all $\lambda \in I'$. To construct q, pick a sequence $\{I_n\}$ of compact subintervals of I, starting with I_0, such that $n < m$ implies $I_n \subset I_m$ and $\cup_n I_n = I$. Pick a flat neighborhood N_0 of $\{p(\lambda) : \lambda \in I_0\}$, and an open subset M_0 of Minkowski space isometric to N_0, letting $P_0 : N_0 \to M_0$ be the isometry. Define $q(\lambda)$ to be $P_0(p(\lambda))$ for $\lambda \in I_0$. Now, given a flat neighborhood N_1 of $\{p(\lambda) : \lambda \in I_1\}$, we wish to choose the open subset M_1 of Minkowski space isometric to it in such a way that for all $p(\lambda)$ where $\lambda \in I_0$ the isometry $P_1 : N_1 \to M_1$ matches the isometry P_0, that is, $q(\lambda) \equiv P_0(p(\lambda)) = P_1(p(\lambda))$. This can be done by matching P_1 to P_0 at a single point $p(\lambda_0)$, where $\lambda_0 \in I_0$, and on its tangent space, denoted $T_{p(\lambda_0)}$, thus uniquely determining P_1 and ensuring that it matches P_0 throughout the connected component of $N_0 \cap N_1$ containing $p(\lambda_0)$. This connected component clearly contains $p(\lambda)$ for all $\lambda \in I_0$. We thus define $q(\lambda) = P_1(p(\lambda))$ for $\lambda \in I_1$; this definition will be consistent with the previous one for all $\lambda \in I_0$. Repeating this procedure for each I_n in turn we obtain a definition of q for all $\lambda \in I$. We now show that this construction is unique, that is, for any compact subinterval I' of I, M' and $P' : N' \to M'$ can be chosen so that $P'(p(\lambda)) = q(\lambda)$ for all $\lambda \in I'$. There will be some n such that $I' \subset I_n$. Again, P' can be chosen to match P_n at some point $p(\lambda_0)$, where $\lambda_0 \in I'$, and on its tangent space $T_{p(\lambda_0)}$. With this choice it must match throughout the connected component of $N' \cap N_n$ containing $p(\lambda_0)$, which includes $p(\lambda)$ for all $\lambda \in I'$. Thus the path $q(\lambda)$ ($\lambda \in I$) is well defined and is unique for each choice of P_0.

A segmentwise flat path $p(\lambda)$ ($\lambda \in I$) is said to *go to (future timelike, future null, spacelike, past null, past timelike) infinity* if the corresponding path $q(\lambda)$ in Minkowski space goes to (future timelike, future null, spacelike, past null, past timelike) infinity. A *loop at infinity with property P* is a surface $s(r,\lambda)$ ($r \in [0,+\infty)$, $\lambda \in [0,1]$) such that (1) for each $r_0 \in [0,+\infty)$ $s(r_0,\lambda)$ ($\lambda \in [0,1]$) is a loop with property P and (2) for each $\lambda_0 \in [0,1]$, $s(r,\lambda_0)$ ($r \in [0,+\infty)$) is a segmentwise flat path that goes to spacelike or null infinity. This last requirement automatically implies that the image of s is a locally flat subset of the spacetime. We now turn to applying these definitions to CTC-containing spacetimes.

Theorem. The parallel-transport transformation is not a rotation for a CTC $s(r,\lambda)$ ($r \in [0,+\infty)$, $\lambda \in [0,1]$) at infinity such that $s(r_1,\lambda_1) = s(r_2,\lambda_2)$ implies $r_1 = r_2$ and either $\lambda_1 = \lambda_2$ or $|\lambda_1 - \lambda_2| = 1$.

Proof. The general outline of the proof is as follows. First, we construct a surface $q(r,\lambda)$ ($r \in [0,+\infty)$, $\lambda \in [0,1]$) in Minkowski space, locally isometric to $s(r,\lambda)$. The method here is similar to that used earlier to construct a path in Minkowski space locally isometric to a given segmentwise flat path. Since Minkowski space does not contain any CTC's the mapping from s to q will be discontinuous: for all $r \in [0,+\infty)$, $s(r,0) = s(r,1)$ whereas $q(r,0) \neq q(r,1)$. The composition of the isometry from the tangent space $T_{q(r,1)}$ of $q(r,1)$ to $T_{s(r,1)} = T_{s(r,0)}$ with the isometry from $T_{s(r,0)}$ to $T_{q(r,0)}$ will then induce an isometry P_r on Minkowski space whose homogeneous component is similar to the PT transformation for the CTC $s(r,\lambda)$ ($\lambda \in [0,1]$). These isometries P_r will be shown to be identical for all values of r. Finally, we will show that the homogeneous component of an isometry P of Minkowski space cannot be a rotation if the displacement undergone under action by P is timelike for a sequence of points approaching spacelike or null infinity, as it is for the points $q(r,1)$.

We begin by constructing the surface $q(r,\lambda)$, $r \in [0,r_0]$, for some $r_0 > 0$. Let $S(r_0) = \{s(r,\lambda) : r \in [0,r_0], \lambda \in [0,1]\}$. For each $\lambda_0 \in [0,1]$, the path $s(r,\lambda_0)$ is segmentwise flat, so the curve $\{s(r,\lambda_0) : r \in [0,r_0]\}$ has a flat neighborhood $N_0(r_0,\lambda_0)$ [with $N_0(r_0,0) = N_0(r_1,1)$]. For each $\lambda_0 \in (0,1)$, there is an open interval $I(\lambda_0) \subset (0,1)$ such that

$$R(r_0,\lambda_0) \equiv \{s(r,\lambda) : \lambda \in I(\lambda_0), r \in [0,r_0]\} \tag{17}$$

is contained in $N_0(r_0,\lambda_0)$. [This is because a sequence of points in $S(r_0) - N_0(r_0,\lambda_0)$ with values of λ converging to λ_0 must have a limit point in $S(r_0) - N_0(r_0,\lambda_0)$, which is impossible because $N_0(r_0,\lambda_0)$ contains all points in $S(r_0)$ with $\lambda = \lambda_0$.] Similarly, for $\lambda_0 = 0$ and 1 there is a set $I(0) = I(1) = [0,a) \cup (b,1] (0 < a < b < 1)$ such that

$$\begin{aligned} R(r_0,0) &\equiv R(r_0,1) \\ &\equiv \{s(r,\lambda) : \lambda \in I(0), r \in [0,r_0]\} \end{aligned} \tag{18}$$

is contained in $N_0(r_0,0)$. We now define $N_1(r_0,\lambda_0)$ to be the connected component of $R(r_0,\lambda_0)$ in $N_0(r_0,\lambda_0) - (S(r_0) - R(r_0,\lambda_0))$. $N_1(r_0,\lambda_0)$ is a flat neighborhood of $\{s(r,\lambda_0) : r \in [0,r_0]\}$ that intersects $S(r_0)$ exactly in $R(r_0,\lambda_0)$. The collection of these $N_1(r_0,\lambda_0)$ for all $\lambda_0 \in [0,1]$ is a covering of $S(r_0)$, from which we may choose a finite subcover by the compactness of $S(r_0)$. This subcover will necessarily include $N_1(r_0,0) = N_1(r_0,1)$, since it is the only N_1 that contains $\{s(r,0) : r \in [0,r_0]\}$. Of this finite collection choose a subcollection such that no neighborhood, when restricted to $S(r_0)$, is contained in another neighborhood, and such that each point of $S(r_0)$ is covered by at least one and at most two neighborhoods. [If a point of $S(r_0)$ is covered by three different neighborhoods, one of these is a subset when restricted to $S(r_0)$ of the other two and may therefore be discarded.] Let the values of λ_0 for the neighborhoods of this subcollection be $\lambda_1, \ldots, \lambda_n$, where $0 = \lambda_1 < \lambda_2 \cdots < \lambda_n = 1$ (we include both $\lambda_0 = 0$ and $\lambda_0 = 1$, although their neighborhoods are the same). These neighborhoods have the property that $R(r_0,\lambda_i)$ intersects $R(r_0,\lambda_j)$ if and only if $|i-j| \leq 1$ or $\geq n-2$.

Using these flat neighborhoods $N_1(r_0,\lambda_i)$ we now construct the surface $q(r,\lambda)$ in Minkowski space. Let M_1 be an open set in Minkowski space isometric to $N_1(r_0,\lambda_1)$ and let $P_1 : N_1(r_0,\lambda_1) \to M_1$ be the isometry. For $r \in [0,r_0]$ and $\lambda \in [0,a)$ define $q(r,\lambda) = P_1(s(r,\lambda))$. (Note: We could define q for values of λ throughout $I(\lambda_1)$, but for now we choose to let q be undefined for $\lambda \in (b,1]$.) The isometry P_2 from $N_1(r_0,\lambda_2)$ to Minkowski space is uniquely determined by specifying it for a single point and the tangent space at that point, since $N_1(r_0,\lambda_2)$ is a connected open set. We specify this isometry at some point $s(r',\lambda')$, where $r' \in [0,r_0]$ and $\lambda' \in [0,a) \cap I(\lambda_2)$, to match the isometry P_1 at that point. P_1 and P_2 will then match throughout the connected component of $N_1(r_0,\lambda_1) \cap N_1(r_0,\lambda_2)$ containing $s(r',\lambda')$, which includes all of $\{s(r,\lambda) : r \in [0,r_0], \lambda \in [0,a) \cap I(\lambda_2)\}$. Hence we may extend the definition of q to include values of λ in $I(\lambda_2)$, in a way that is consistent with the old definition: for $r \in [0,r_0]$ and $\lambda \in I(\lambda_2)$ set $q(r,\lambda) = P_2(s(r,\lambda))$. [Since $N_1(r_0,\lambda_1) \cap N_1(r_0,\lambda_2)$ may have other connected components, and P_1 and P_2 do not necessarily match on these, we have *not* constructed an isometry between $N_1(r_0,\lambda_1) \cup N_1(r_0,\lambda_2)$ and an open set in Minkowski space.] We now repeat this procedure successively for each interval $I(\lambda_i)$, obtaining a definition of $q(r,\lambda)$ for all $\lambda \in [0,1]$ and $r \in [0,r_0]$. The last application of the procedure uses the isometry P_n on $N_1(r_0,\lambda_n)$ to define $q(r,\lambda)$ for $\lambda \in (b,1]$ rather than for all $\lambda \in I(\lambda_n)$—for $\lambda \in [0,a)$, q is already defined using the isometry P_1. Thus although $s(r,0) = s(r,1)$ for all r, $q(r,0) \neq q(r,1)$ where q is defined. For any $r \in [0,r_0]$ and $\lambda' \in I(\lambda_i) \cap I(\lambda_{i+1})$, $q(r,\lambda)$ is defined twice, once as $P_i(s(r,\lambda))$ and once as $P_{i+1}(s(r,\lambda))$, but by construction the two definitions agree.

We now wish to extend the surface $q(r,\lambda)$ so that it is defined for all $r \in [0,+\infty)$. For a given $r_1 > r_0$ we can obtain another surface $q'(r,\lambda)(r \in [0,r_1], \lambda \in [0,1])$ using the same construction. Clearly, we can set $q'(0,0) = q(0,0)$, as well as set equal the two induced mappings from $T_{s(0,0)}$ to $T_{q(0,0)} = T_{q'(0,0)}$. We now wish to show that they then match for all $r \in [0,r_0]$ and $\lambda \in [0,1]$. To do this we show that they match on the curve $r = 0$, $\lambda \in [0,1]$. The construction of q gives for each r and λ a mapping from $T_{s(r,\lambda)}$ to $T_{q(r,\lambda)}$, and similarly for q'. These two mappings will also be shown to match for $r = 0$, $\lambda \in [0,1]$. Since for any $\lambda_0 \in [0,1]$ the curve $s(r,\lambda_0)$ $(r \in [0,+\infty))$ is segmentwise flat, the curves $q(r,\lambda_0)$ $(r \in [0,r_0])$ and $q'(r,\lambda_0)$ $(r \in [0,r_1])$ in Minkowski space corresponding to the curve $s(r,\lambda_0)$ are uniquely determined by specifying $q(0,\lambda_0)$ and $q'(0,\lambda_0)$ and the mappings from $T_{s(0,\lambda_0)}$ to $T_{q(0,\lambda_0)}$ and to $T_{q'(0,\lambda_0)}$, respectively. It follows that q and q' match for all $r \in [0,r_0]$ and $\lambda \in [0,1]$.

The construction of q provides the following for each $\lambda_0 \in [0,1]$: a neighborhood $N(\lambda_0)$ of $s(0,\lambda_0)$; a neighborhood $M(\lambda_0)$ of $q(0,\lambda_0)$; an isometry $P_{\lambda_0} : N(\lambda_0) \to M(\lambda_0)$ between them; and an interval $I(\lambda_0) \subset [0,1]$, that is open as a subset of [0,1], containing λ_0 and such that for all $\lambda \in I(\lambda_0)$ $s(0,\lambda) \in N(\lambda_0)$ and $q(0,\lambda) = P_{\lambda_0}(s(0,\lambda))$. The construction of q' similarly provides $N'(\lambda_0) \ni s(0,\lambda_0)$, $M'(\lambda_0) \ni q'(0,\lambda_0)$, $P'_{\lambda_0} : N'(\lambda_0) \to M'(\lambda_0)$, and $I'(\lambda_0) \ni \lambda_0$, with the same properties. We define $N''(\lambda_0)$ to be the connected component of $N(\lambda_0) \cap N'(\lambda_0)$ containing $s(0,\lambda_0)$, and set $I''(\lambda_0) = I(\lambda_0) \cap I'(\lambda_0)$. $I''(\lambda_0)$ is thus an interval containing λ_0, open as a subset of [0,1], such that for all $\lambda \in I''(\lambda_0)$ $s(0,\lambda) \in N''(\lambda_0)$, $q(0,\lambda) = P_{\lambda_0}(s(0,\lambda))$, and $q'(0,\lambda) = P'_{\lambda_0}(s(0,\lambda))$. By assumption and since $N''(0)$ is connected, $P_0 = P'_0$ on $N''(0)$. Since [0,1] is compact, we can choose a finite set $\{\lambda_k\}$ such that $\cup_k I''(\lambda_k) = [0,1]$. Now for each i such that $I''(\lambda_i)$ intersects $I''(0)$, $P_{\lambda_i} = P'_{\lambda_i}$ on $N''(\lambda_i)$. This is because $N''(\lambda_i)$ is connected, and for any $\lambda \in I''(0) \cap I''(\lambda_i)$ $P_{\lambda_i}(s(0\lambda)) = P'_{\lambda_i}(s(0,\lambda))$ [since

$$\begin{aligned} P_{\lambda_i}(s(0,\lambda)) = q(0,\lambda) &= P_0(s(0,\lambda)) \\ &= P'_0(s(0,\lambda)) \\ &= q'(0,\lambda) \\ &= P'_{\lambda_i}(s(0,\lambda))] \end{aligned}$$

and P_{λ_i} induces the same mapping from $T_{s(0,\lambda)}$ to $T_{q(0,\lambda)} = T_{q'(0,\lambda)}$ as P'_{λ_i} (since it induces the same mapping as P_0, which induces the same mapping as P'_0, which induces the same mapping as P'_{λ_i}). By the same reasoning, for any λ_j such that $I''(\lambda_j)$ intersects one of these $I''(\lambda_i)$, $P_{\lambda_j} = P'_{\lambda_j}$ on $N''(\lambda_j)$, and so on until we know that for all k $P_{\lambda_k} = P'_{\lambda_k}$ on $N''(\lambda_k)$, which implies that for all $\lambda \in [0,1]$ $q(0,\lambda) = P_{\lambda_k}(s(0,\lambda)) = P'_{\lambda_k}(s(0,\lambda)) = q'(0,\lambda)$ for some k, and similarly for the mappings from $T_{s(0,\lambda)}$ to $T_{q(0,\lambda)} = T'_{q'(0,\lambda)}$.

We have thus shown that the surface $q(r,\lambda)$ $(r \in [0,+\infty)$, $\lambda \in [0,1])$ in Minkowski space, locally isometric to the surface $s(r,\lambda)$, exists and is unique up to isometries of Minkowski space. The path $s(r,0) = s(r,1)(r \in [0,+\infty))$, being segmentwise flat, is locally isometric to a set of paths in Minkowski space that are related to each other by isometries of Minkowski space; since both $q(r,0)$ and $q(r,1)(r \in [0,+\infty))$ are locally isometric to $s(r,0)$, they are related by an isometry of Minkowski space. This isometry relates not only the points $q(r,0)$ and $q(r,1)$ but also their tangent spaces $T_{q(r,0)}$ and $T_{q(r,1)}$ via $T_{s(r,0)} = T_{s(r,1)}$. Denote by $L_0 : T_{s(r_0,0)} \to T_{q(r_0,0)}$ and $L_1 : T_{s(r_0,0)} \to T_{q(r_0,1)}$ the tangent space mappings defined by the local isometry between the surface $s(r,\lambda)$ and $q(r,\lambda)$, for some $r_0 \in [0,+\infty)$. Parallel transport (along any path) defines a mapping $L_p : T_{q(r_0,0)} \to T_{q(r_0,1)}$, and parallel transport along the CTC $s(r_0,\lambda)$ $(\lambda \in [0,1])$ defines the Lorentz transformation $\Lambda_s : T_{s(r_0,0)} \to T_{s(r_0,0)}$. The local isometry between the surfaces $s(r,\lambda)$ and $q(r,\lambda)$ ensures that parallel transport will give the same result when done on either surface; that is, $\Lambda_s = L_1^{-1} L_p L_0$. Now, the homogeneous component of the Minkowski space isometry P bringing $q(r_0,1)$ to $q(r_0,0)$ is $\Lambda_q = L_P L_0 L_1^{-1} : T_{q(r_0,1)} \to T_{q(r_0,1)}$. Thus the PT transformation for any of the CTC's in the surface $s(r,\lambda)$ is similar to the homogeneous component of P: $\Lambda_s = L_1^{-1} \Lambda_q L_1$.

It remains to be shown that the homogeneous component of P cannot be a rotation. Though the easiest part of the proof, it is really the crux of the problem. First we note that, for all $r \in [0,+\infty)$, $q(r,1) - P(q(r,1)) = q(r,1) - q(r,0)$ is a timelike vector, since there is a future-directed timelike curve $q(r, \lambda)$ ($\lambda \in [0,1]$) joining $q(r,0)$ to $q(r,1)$. Now, assuming that the homogeneous component of P is a rotation, there is a Lorentzian coordinate system in which $P(x) = R_\theta(x)+\tau\hat{t}$. For a point a distance ρ from the t axis, the displacement $P(x)-x$ is timelike if and only if $2\rho\sin(\theta/2) < |\tau|$. Since $q(r,1)$ goes to spacelike or null infinity as r goes to infinity, its distance ρ from the t axis goes to infinity in any (Lorentzian) coordinate system, so the displacement under action by P cannot be timelike for large r, in contradiction with what we have proven. Q.E.D.

V. CLOSED TIMELIKE CURVES GENERATED BY COSMIC STRINGS IN 3+1 DIMENSIONS

As noted in the Introduction, one of the principle reasons for investigating gravity in 2+1 dimensions is that the results in this arena are directly relevant to understanding (3+1)-dimensional spacetimes containing infinitely long, straight, parallel cosmic strings.

For instance, in the first example of a (2+1)-dimensional solution to Einstein's equations described in Sec. II, the only nonvanishing component of the energy-momentum tensor is $T_{00} \equiv \rho$, implying $R = -16\pi\rho$. We convert this solution to a (3+1)-dimensional solution by adding a third independent spatial coordinate, and setting $g_{33} = 1$ and $g_{\mu 3} = g_{3\mu} = 0$ for $\mu \neq 3$. All the components of the connection and the Riemann and Ricci tensors with any index equal to 3 will then vanish, while their other components, along with the curvature scalar, will remain unchanged. From Einstein's equations it follows that $T_{33} = (1/16\pi)R$, $T_{\mu 3} = T_{3\mu} = 0$ for $\mu \neq 3$, and the remaining components of the energy-momentum tensor are unchanged. We find that $T_{33} = -\rho = -T_{00}$. Since this is the energy-momentum tensor for the "matter" that makes up cosmic strings we see that the solution for any number of infinite, straight, parallel, and stationary strings is equivalent to (2+1)-dimensional solutions of this form [16]. Note that in this system of units ($G = c = 1$), mass is dimensionless in 2+1 dimensions but has the same units as length and time in 3+1 dimensions. Hence the mass of a particle in 2+1 dimensions corresponds to the mass per unit length of a cosmic string in 3+1 dimensions [16].

Similarly, (2+1)-dimensional solutions containing particles in relative motion can be converted into (3+1)-dimensional solutions containing moving infinite straight parallel cosmic strings. In particular Gott's two-particle spacetime described in Sec. III can be converted into a (3+1)-dimensional spacetime containing two parallel infinite strings passing each other at very high speed, indeed, it was in this form that the solution was originally discovered [3]. Since the particles' mass M is replaced by the strings' mass per unit length μ, the criterion for the existence of CTC's becomes

$$\beta > \cos 4\pi\mu\,, \tag{19}$$

where β is the speed of the strings.

Although two straight infinite strings must be parallel in order to be equivalent to (2+1)-dimensional particles, it should be noted that many two-string spacetimes admit a coordinate system in which the strings are parallel, including all the two-string spacetimes that contain CTC's. The set of spacetimes containing two nonintersecting infinite straight strings moving at constant velocity can be divided into four classes: (1) those spacetimes that can be reduced by a change of coordinates to two stationary and parallel strings; (2) those that can be reduced to two stationary but skew strings; (3) those that can be reduced to two strings arbitrarily close to being parallel and moving arbitrarily slowly (these spacetimes are on the borderline between the two previous cases); and (4) those spacetimes that can be reduced to two moving, parallel strings. All of the CTC-containing spacetimes are in the fourth class. The four different classes can be constructed [3] starting with the two strings at rest, parallel to the z axis in the laboratory coordinate system, and displaced from each other in the y direction. The first string is rotated about the y axis by an angle ϕ and the second string by an angle $-\phi$. This provides a static spacetime with the strings skewed. The first string can then be boosted by to a γ factor of γ_F in the x direction, while the second string is boosted to the same γ factor in the opposite direction The four cases are then (1) $\gamma_F = \cos\phi = 1$: strings parallel and at rest; (2) $\gamma_f\cos\phi < 1$: a boost in the z direction with speed $\gamma_F\beta_F\tan\phi$ will make the strings motionless but skew; (3) $\gamma_F\cos\phi = 1$: there is no frame that minimizes the string velocities, but as the solution is boosted in the z direction with arbitrarily high boost parameters the strings move slower and slower, becoming closer and closer to being parallel and approaching arbitrarily close to case (1); (4) $\gamma_F\cos\phi > 1$: in the frame that equalizes and minimizes the strings' speeds β_s we have $\gamma_s = \gamma_F\cos\phi$ [3] and the strings are parallel. For this construction the criterion for the existence of CTC's is $\gamma_F\cos\phi\sin(4\pi\mu) > 1$, which is satisfied only by the case (4) solutions.

One of the major issues for any CTC-containing solution is its applicability to time-machine technology. Since a cosmic string has no ends, it is either infinite or in a loop. If we are not lucky enough to live in a universe that contains two infinite cosmic strings that pass each other at very high speed, the question arises as to whether it is possible to make a time machine based on the same effect by manipulating finite lengths of cosmic string. Tipler [17] and Hawking [18] showed that, under certain conditions including the weak energy condition, CTC formation cannot originate within a finite singularity-free region of spacetime if the Cauchy horizon extends to future null infinity [17] or is compactly generated [18]. Thus, with the weak energy condition it would seem that a successful attempt to manufacture CTC's within a finite region of space will be accompanied by the creation of a singularity within that region. This does not immediately imply, however, that with sufficiently advanced technology one could not make a time machine. There is no reason to suspect that spacetime singularities could not in principle be created through deliberate human action. Furthermore, there is the possibility of realizing a solution in which the Tipler-Hawking theorems do not apply because the Cauchy horizon is neither compactly generated nor reaches future null infinity. This possibility is exemplified by the maximally extended Kerr solution, which obeys the weak energy condition and contains CTC's, and in which the Cauchy horizon is completely enclosed by the finite event horizon.

To illustrate this possibility more concretely we will describe an approximation to the Gott solution created with a single finite loop of string. Such a loop could in principle be manufactured by heating a long volume of space to the grand unified theory (GUT) temperature, supercooling it, and appropriately manipulating the Higgs field to nucleate the string loop [19]. Garfinkle and Vachaspati [20] gave a simple example of cosmic string loop motion, neglecting gravitational radiation, in which the loop is in the shape of a rectangle at all times, and such that the center of the rectangle is fixed. As a function of the time t, the length of the rectangle is $(2L-2t)\cos\alpha$, while its height is $2t\sin\alpha$, where $2L\cos\alpha$ is the initial length of the loop and $2L\sin\alpha$ the final length. The horizontal sides of the rectangle thus move at a speed of $\beta = \sin\alpha$, the vertical sides at a speed of $\beta = \cos\alpha$, and the corners at the speed light, as kinks in cosmic strings must do. (This is why they are able to neglect gravitational radiation. Because the straight segments of string moving uniformly emit no gravitational radiation, the only gravitational radiation emitted comes from the four corners or kinks which move at the speed of light and accelerate the string segments they pass through. The kinks thus act like four point sources of gravitational radiation. If the mass per unit length μ in the strings is many orders of magnitude less than one Planck mass per Planck length, which we expect, then the fractional dynamical effects of this radiation will be small, of order 50μ [21].)

In this solution the vertical segments of the rectangle collide, but they can be made to avoid each other by a small distance by introducing additional sides so that the entire loop no longer remains in one plane [19]. By making α very small, the vertical segments can be made to pass each other at very high speed. For instance, if $\alpha = 10^{-5}$, then the initial (horizontal) segments would each have a speed of $\beta = 10^{-5}$; as the long loop collapses these segments will be converted by the motion of the kinks into short vertical segments each with speed $\beta = 1-5\times10^{-11}$ and a γ factor of 10^5. If the mass per unit length of the cosmic string is 10^{-6}, then the condition for the existence of CTC's in the infinite-string solution is satisfied. (Gott showed [3] that Hawking's proof that a circular loop of string will collapse to form a black hole reveals that such a loop will also reach a γ factor of this order of magnitude.) While in this case the string is finite, if the length of the vertical segments is extremely large compared to their impact parameter, then it might be natural to suppose that the spacetime near the segments but far from their endpoints will resemble the infinite-string solution and that CTC's might be present there. Unlike in the Gott solution, however, CTC's would not be present infinitely far from the sources; since the passing strings carry some angular momentum, very far away the spacetime would resemble an exterior Kerr solution. Gott showed [3] that in the finite loop case the condition for the existence of CTC's is the same as the condition for the existence of an event horizon under Thorne's hoop conjecture. Thus there is the danger that a black hole could form, possibly either disrupting the CTC's or cloaking them behind the event horizon [3]. Alternatively, the Cauchy horizon might end on singularities that are naked and thus escape out to future null infinity, thus revealing the CTC's to the surrounding space; but that would of course require violating the cosmic censorship conjecture.

The other main line of attack on time machines is in the quantum regime [9,10,18,20,22–24]. In particular, Hawking [18] has argued that instabilities leading to divergences in quantum fields at the Cauchy horizon would in general prevent the appearance of CTC's in spacetimes that would otherwise classically evolve them. In the Gott solution, however, it appears [9,10,23] that these instabilities are mild, particularly compared to those in the wormhole solutions; unlike the wormhole solutions, the Gott solution has no closed null geodesic, or fountain, where the instabilities pile up. On the other hand, in the Gott spacetime stronger instabilities may occur within the CTC-containing region than on the Cauchy horizon itself [9,10]. In any case, there is some question as to how to do quantum mechanics in regions containing CTC's (see for example [24]). Since these quantum-mechanical divergences appear to become significant within a Planck-scale distance from the Cauchy horizon, solving these problems would appear to necessitate a quantum theory of gravity; indeed, one of the main motivations for studying CTC-containing spacetimes is to gain insight into the constraints on a quantum theory of gravity imposed by such extreme possibilities.

The other principle reason for studying CTC-containing spacetimes is that they may allow us to get a physical handle on some of the fascinating philosophical issues surrounding time travel [25]. What about the concern that time travel incurs "causal paradoxes" [18,26]? Like all dynamical physical theories (with the possible exception of the many-worlds interpretation of quantum mechanics; see [27] for an account of time travel within this theory), general relativity requires a model of the world that includes a single history, one spacetime with world lines in it. Many supposed paradoxes involving time travel rely upon the possibility of changing history, a notion that is ruled out *a priori* by the conceptual foundations of general relativity. The principle of self-consistency that scientists normally use is quite applicable to time travel [28] and requires no special augmentation. Although in our everyday lives we feel that by our actions we affect the future course of events, in a very strict sense we never actually change the future from one future to another, because there will in the end be only a single future; similarly, without any special restrictions on his free will, a time traveler does not actually change the past, although his actions might well have played an active role in historical developments. Since no reasonable concept of free will can include the freedom to do things that are logically impossible, it would seem that the philosophical issues surrounding free will are not relevant one way or the other to the study of CTC-containing spacetimes.

Finally, we would like to point out that, with any general-relativistic time machine, events prior to the construction of the time machine, i.e., prior to the Cauchy horizon, will be inaccessible to time travelers. Since we have constructed no time machine yet, "the fact that we have not been invaded by hordes of tourists from the future" [18] does not constitute evidence for or against the possibility of future time-machine technology.

ACKNOWLEDGMENTS

We would like to thank David Spergel for his comments about finite loops of cosmic string and Demetrios Christodoulou for his comments regarding the inapplicability of such concepts as "total momentum" and "center-of-mass velocity" in spacetimes that are not asymptotically Minkowskian. This research was supported by NSF Grant No. AST 90-20506 and NASA Grant No. NAGW 765.

[1] A. Staruszkiewicz, Acta Phys. Polon. **24**, 735 (1963); S. Giddings, J. Abbot, and K. Kuchar, Gen. Relativ. Gravit. **16**, 751 (1984); J. R. Gott and M. Alpert, *ibid.* **16**, 243 (1984).
[2] S. Deser, R. Jackiw, and G. 't Hooft, Ann. Phys. (N.Y.) **152**, 220 (1984).
[3] J. R. Gott, Phys. Rev. Lett. **66**, 1126 (1991).
[4] S. Deser, R. Jackiw, and G. 't Hooft, Phys. Rev. Lett. **68**, 267 (1992).
[5] S. M. Carroll, E. Farhi, A. H. Guth, and K. D. Olum, Phys. Rev. D **50**, 6190 (1994).
[6] D. N. Kabat, Phys. Rev. D **46**, 2720 (1992).
[7] S. M. Carroll, E. Farhi, and A. H. Guth, Phys. Rev. Lett. **68**, 263 (1992); **68**, 3368(E) (1992).
[8] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New Nork, 1972), pp. 165–171; C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, New York, 1973), pp. 448–483.
[9] J. E. E. Grant, Phys. Rev. D **47**, 2388 (1993).
[10] D. Laurence, report, 1994 (unpublished).
[11] B. Jensen and H. H. Soleng, Phys. Rev. D **45**, 3528 (1992).
[12] H. H. Soleng, Phys. Rev. D **49**, 1124 (1994); N. J. Cornish and N. E. Frankel, Class. Quantum Grav. **11**, 723 (1992).
[13] C. Cutler, Phys. Rev. D **45**, 487 (1992).
[14] G. 't Hooft, Class. Quantum Grav. **9**, 1335 (1992); **10**, 1023 (1993).
[15] S. Deser, Class. Quantum Grav. **10**, S67 (1993).
[16] J. R. Gott, Astrophys. J. **288**, 422 (1985).
[17] F. J. Tipler, Phys. Rev. Lett. **37**, 879 (1976); Ann. Phys. (N.Y.) **108**, 1 (1977).
[18] S. W. Hawking, Phys. Rev. D **46**, 603 (1992).
[19] D. Spergel (private communication).
[20] D. Garfinkle and T. Vachaspati, Phys. Rev. D **36**, 2229 (1987).
[21] R. S. Scherrer, J. M. Quashnock, D. N. Spergel, and W. H. Press, Phys. Rev. D **42**, 1908 (1990).
[22] S. W. Kim and K. S. Thorne, Phys. Rev. D **43**, 3929 (1991); D. S. Goldwirth, M. J. Perry, T. Piran, and K. S. Thorne, *ibid.* **49**, 3951 (1994); D. G. Boulware, *ibid.* **46**, 4421 (1992); H. D. Politzer, *ibid.* **46**, 4470 (1992); M. Visser, *ibid.* **47**, 554 (1993); H. D. Politzer, *ibid.* **49**, 3981 (1994); M. Visser, *ibid.* **49**, 3963 (1994); J. B. Hartle, *ibid.* **49**, 6543 (1994).
[23] D. G. Boulware (unpublished).
[24] J. L. Friedman, N. J. Papastamatiou, and J. Z. Simon, Phys. Rev. D **46**, 4456 (1992).
[25] B. Allen and J. Z. Simon, Nature (London) **357**, 19 (1992); P. J. Nahin, *Time Machines: Time Travel in Physics, Metaphysics, and Science Fiction* (American Institute of Physics, New York, 1993).
[26] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, England, 1973), p. 189.
[27] D. Deutsch and M. Lockwood, Sci. Am. **270** (3) (1994).
[28] J. L. Friedman, M. S. Morris, I. D. Novikov, F. Echeverria, G. Klinkhammer, K. S. Thorne, and U. Yurtsever, Phys. Rev. D **42**, 1915 (1990); M. S. Morris, K. S. Thorne, and U. Yurtsever, Phys. Rev. Lett. **61**, 1446 (1988).

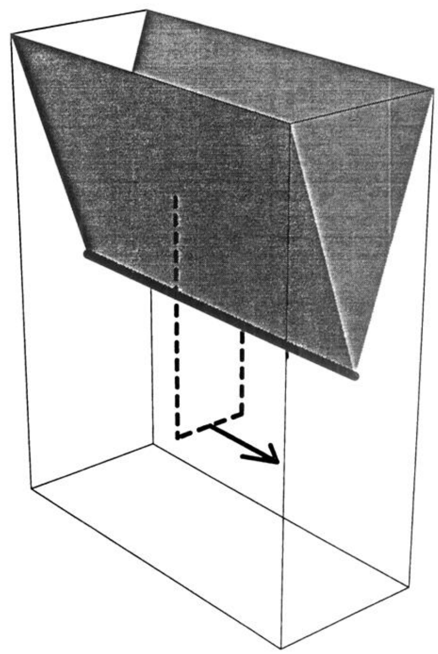

FIG. 1. Part of a (2+1)-dimensional spacetime containing a single transcendent tachyon of total pressure $P \approx 0.03$; the thick solid line is the world line of the tachyon. The vertical coordinate represents time and the horizontal ones space. Outside of the tachyon the metric is the Lorentz metric. The two half planes are identified with each other and the space between them is not part of the spacetime. Also shown is a loop encircling the tachyon (dashed line) and the axis of the PT transformation for that loop; note that it is parallel with the tachyon's world line.

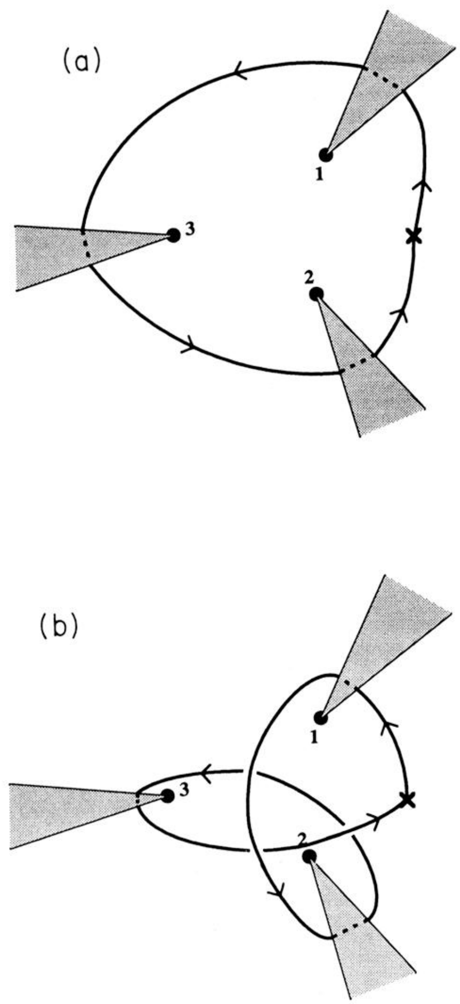

FIG. 2. Diagrammatic representation, in which the time coordinate has been suppressed, of a (2+1)-dimensional spacetime containing three masses in relative motion. The masses are represented by the large dots; the shaded areas are not part of the spacetime, and the two rays bounding each shaded area are identified with each other. Different nonintersecting loops in this spacetime, with the same base point marked by the cross, are shown in (a) and (b). As explained in the text, the PT transformations for these two loops are different, although both loops satisfy the condition of being the edge of a surface that the world lines of all the masses cross through.

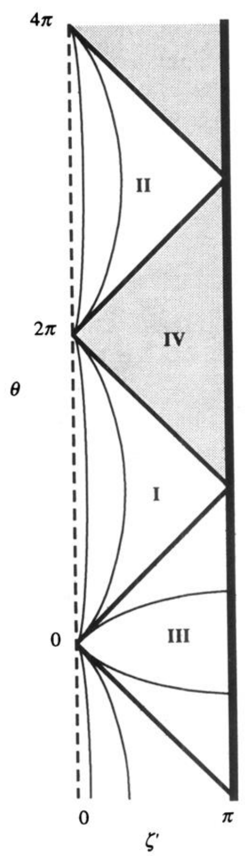

FIG. 3. Part of the Penrose diagram of (2+1)-dimensional universal anti–de Sitter space, representing the Lie group $\overline{\mathrm{SO}}(2,1)$. θ is the time coordinate, ranging from $-\infty$ to $+\infty$; ζ' is the radial coordinate, ranging from 0 to π. The angular coordinate is suppressed. The dashed line $\zeta' = 0$ is the coordinate singularity at the origin of polar coordinates. The heavy solid line $\zeta' = \pi$ represents spacelike and null infinity. The identity of $\overline{\mathrm{SO}}(2,1)$ is represented by the point $\theta = \zeta' = 0$. Three geodesics through the origin, two timelike and one spacelike, are shown. The shaded regions are points not reachable by geodesics through the origin, i.e., points with no inverse images under the exponential map. The solid lines separating the different labeled regions represent null surfaces.

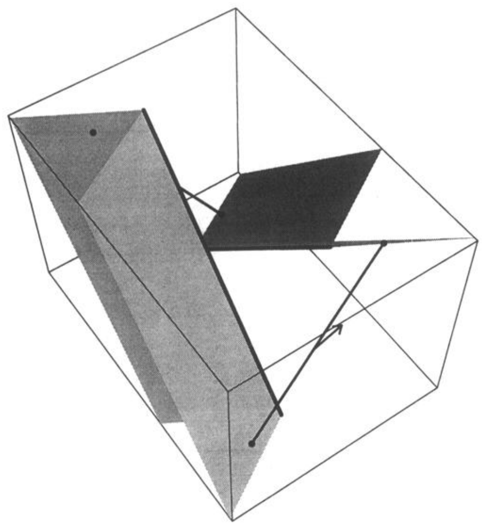

FIG. 4. A Gott two-particle spacetime. The view is from above, that is, from the future looking toward the past. Again, the vertical direction represents the time coordinate and the horizontal directions the space coordinates. The thick solid lines represent the world lines of the particles. The two half-planes coming off of each world line are identified with each other. Thus the two large dots shown on the two half-planes coming off of each world line represent the same point in the spacetime. The two wedges between these half-planes do not represent parts of the spacetime. Outside the particles, the spacetime is flat, and the metric in the coordinate system shown is the Lorentz metric. The two line segments shown are both timelike and together form a CTC. The arrow represents the axis of the PT transformation for this CTC with base point at the base of the arrow; it is spacelike.

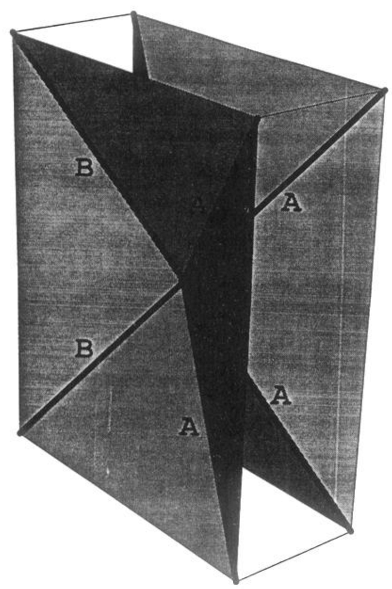

FIG. 5. A (2+1)-dimensional spacetime in which two masses A and B bounce elastically. Again, time is represented by the vertical dimension and the spacetime is locally flat with a Lorentz metric outside of the sources. The planes divide the space of the figure into three disjoint regions. Of these, only the near and far regions are part of the actual spacetime, while the region sandwiched between them is excised. At any time, except the time of the collision, the position of particle A is represented by two identified points, and similarly for particle B; hence each particle has two world lines in this representation.

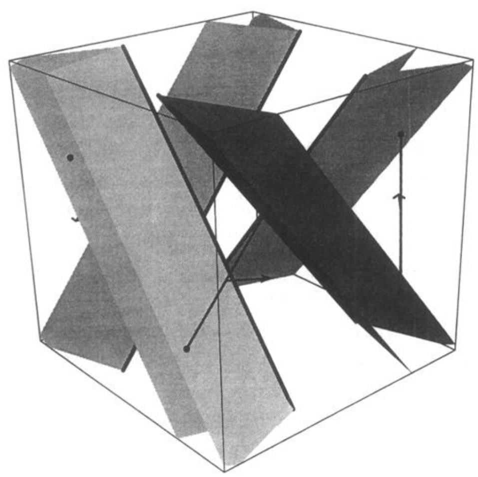

FIG. 6. A Kabat four-particle spacetime. A CTC is shown, along with the axis of the PT transformation for the CTC at three different base points. (One of these axes is mostly obscured by a wedge.) These axes agree when parallel transported along the CTC. As one can see from the picture, they are spacelike. Again, large dots represent points where the CTC crosses the wedges.