Pergamon

Stud. Hist. Phil. Mod. Phys., Vol. 31, No. 2, pp. 227–246, 2000

Printed in Great Britain
1355-2198/00 $ - see front matter

Topology Change and the Unity of Space

Craig Callender*, Robert Weingard✠

Must space be a unity? This question, which exercised Aristotle, Descartes and Kant, is a specific instance of a more general one; namely, can the topology of physical space change with time? In this paper we show how the discussion of the unity of space has been altered but survives in contemporary research in theoretical physics. With a pedagogical review of the role played by the Euler characteristic in the mathematics of relativistic spacetimes, we explain how classical general relativity (modulo considerations about energy conditions) allows virtually unrestrained spatial topology change in four dimensions. We also survey the situation in many other dimensions of interest. However, topology change comes with a cost: a famous theorem by Robert Geroch shows that, for many interesting types of such change, transitions of spatial topology imply the existence of closed timelike curves or temporal non-orientability. Ways of living with this theorem and of evading it are discussed. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Topology Change; General Relativity; Unity of Space; Spacetime; Non-Orientability; Euler Characteristic

1. Descartes and the Unity of Space

In this paper we point out a connection between a traditional philosophical issue about space and a topological invariant of that space, its Euler characteristic. In tracing out that connection we will learn something about topology and the topology of spacetime. Since many philosophers of science are

* Department of Philosophy, Logic and Scientific Method, The London School of Economics, Houghton Street, London WC2A 2AE, U.K. (*e-mail*: c.callender@lse.ac.uk).
✠ Sadly, Professor Weingard (Department of Philosophy, Rutgers University) died on 14 September 1996. The present paper is an updated version of a manuscript originally co-written in 1995.

PII: S1355-2198(00)0003-4

not familiar with this topic, our exposition of topology change is primarily pedagogical in nature. Our hope is that this paper will stimulate philosophers to study the many fascinating issues that connect topology with modern physics.

The philosophical issue is the unity of (three-dimensional physical) space. While usually associated with Kant, Descartes discussed this topic in *The Principles of Philosophy* (1644) and even earlier Aristotle devoted two chapters of *De Caelo* to it. Focusing on Descartes' discussion, the question is whether physical space could occur in two or more 'complete' pieces, such that there are no spatial relations between a point in one of the pieces and a point in any other of the pieces. Each piece seems to be a 'whole' space by itself, in the sense that it is maximal—it cannot be isometrically embedded as a proper part in a connected space of the same dimension. In arguing that there 'cannot be a plurality of worlds' in the *Principles* (Part II, Principle XXII), Descartes is arguing, given his identification of space (extension) with material substance, that space must be a unity.

His idea stems from his assumption that space is infinite, and the thought that, since space is infinite, all the possible places are already part of the world or space. But another, somewhat more general argument for the unity of space is also suggested in the previous principle (XXI). Here Descartes argues that space is 'extended without limit', and he takes this to imply that space is infinite. His argument is

  1. Space is extended without limit
∴ 2. Space is infinite
∴ 3. Space is a unity = There is only one world.

But in fact 1 does not imply 2. Space could be 'extended without limit'—be maximal in the sense mentioned above—and not be infinite. As is well known, three-dimensional spherical space, whose geometry is equivalent to the intrinsic geometry of a three-sphere in four-dimensional Euclidean space, is finite yet unbounded (i.e. without a limit). The more general argument suggested is obtained just by dropping premise 2, that space is infinite, and arguing that

  1. Space is maximal
∴ 3. Space is a unity.

Descartes was arguing that space is necessarily a unity, but it is fair to say that he identified possibility and conceivability. These days we are supposed to know better and we will talk only of conceivability, leaving the questions of possibility and necessity to metaphysics proper.

With that said, however, it is easy to give a counterexample to the claim '1 implies 3', in the sense that it is conceivable that space be maximal and yet not be a unity. We simply make use of two ideas that are familiar from general relativity: i) it is conceivable that space be spherical, and ii) it is conceivable that

Fig. 1. *Space dividing.*

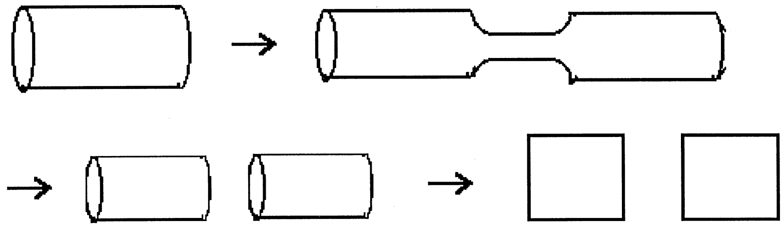

Fig. 2. *An infinite cylinder divides into two infinite planes.*

the geometry of space is a function of time.[1] Putting these together, we can conceive of space not being a unity because we can conceive of spherical space dividing into two (or more) spherical pieces. We can visualise this using a two-dimensional analogue, where we use the surface of a sphere to represent two-dimensional spherical space. The process of space dividing would then be pictured as in Fig. 1.

While we are here picturing this taking place in a higher-dimensional embedding space, this is obviously not required in general. The whole process can be conceived of in an intrinsic way, from the point of view of creatures (two-dimensional ones in our analogue, us in the case of three-dimensional space) that live in the space. From our three-dimensional point of view we can conceive of what it would be like, what we would perceive and measure, if space were spherical and underwent the process of division. At no time must we think in terms of a higher-dimensional space containing physical space.

So much for Descartes' more general argument. But what about the more specific one for the unity of space, based on space being infinite? Here again we can conceive of space dividing into two pieces, but in this case the pieces are each infinite in every direction. A two-dimensional analogue would be an infinite cylinder dividing into two infinite planes (Fig. 2).

Again we are here using the intrinsic geometry of surfaces to represent the geometry of a two-dimensional space. In three dimensions, this corresponds to the division of $E^2 \times S^1$ into two copies of E^3 (E^n is n-dimensional Euclidean space; S^n is n-dimensional spherical space).

This seems to answer Descartes' (and Kant's) question. So is that the end of the discussion? No. Today we do not think of three-dimensional physical space

[1] This example is discussed in R. Weingard (1976).

as an independent structure, but rather as an aspect of four-dimensional relativistic spacetime. There is no such thing as space, simpliciter, but only space relative to a frame of reference. Relative to different frames, space can have very different properties, as in de Sitter spacetime, in which whether space is finite or infinite depends on the frame to which we are referring. Spacetime is the fundamental structure, then, not space (and time), and that suggests a modern form of Descartes' question. Must four-dimensional spacetime be a unity? In fact, many contemporary philosophers, sparked by Quinton (1962), discussed this question. And at least according to one prominent metaphysician, if it is possible for spacetime to have been different, then it is not a unity. But we do not want to pursue these ideas here. It turns out that the question of the unity of space, and related questions, are quite interesting when applied to space as an aspect of spacetime. In fact, because of the constraints imposed by the nature of relativistic spacetimes, the discussion of these questions becomes much richer than in the case of space, simpliciter.

In particular, if we pose the issue of the unity of space in terms of our examples of dividing space, then this question is a special case of the more general question of *whether the topology of space can change*. This is more general because if space divides it certainly changes its topology (connectedness is a topological invariant), but topology change does not imply division. For example, changing from being simply connected to being multiply connected, say from S^3 to $S^1 \times S^1 \times S^1$, does not imply division. The more general question is interesting for a number of different reasons in physics. For example, consider topological particles ('geons'), particles composed of space, in the sense that the presence of such a particle is nothing more than space (or spacetime) having certain topological properties. The particle might be a 'handle' in space, and then the question of whether the number of these particles can change is a question of whether the topology of space can change. Another example where this question arises concerns the issue of whether we can build time machines. If 'building a time machine' means creating a closed timelike curve—say, by creating a wormhole—then one way this might work is by changing the topology of space within a finite region, for this may yield closed timelike curves. Finally, another area where this question is interesting is quantum gravity. Many approaches to the subject find topology change desirable—indeed, inescapable. String theory, Kaluza–Klein theory, topological quantum field theory, Euclidean quantum gravity and others all make use of topology change.

Consequently, the more general question of whether the topology of space can change becomes especially interesting in the context of space as an aspect of spacetime. To appreciate this we need some facts about relativistic spacetimes.

2. Manifolds and Lorentz Metrics

We will take an n-dimensional relativistic spacetime to be a smooth (differentiable) manifold equipped with a semi-Riemannian metric g with a Lorentz signature. That is, at any point p, if $g(A, A) > 0$ for vector A, there exist vectors $B_i, i = 1 \ldots n-1$, at p such that $\{A, B_1, B_2, B_3 \ldots\}$ are mutually orthogonal and $g(B, B) < 0$ for each i. This metric determines the familiar light cone structure of a spacetime and the manifold is a spacetime only with respect to such a metric. An important question for us is under what conditions does a smooth manifold admit an everywhere defined Lorentz metric? We will always assume our smooth manifolds have the required topological structure (paracompact, Hausdorff and differentiable). Since every differentiable manifold admits a globally defined positive definite Riemannian metric, our question becomes: what else is needed in addition to a positive definite Riemannian metric in order to obtain an everywhere well-defined Lorentz metric? The answer is well known. Given a positive definite Riemannian metric g^*, the manifold admits a Lorentz metric iff it admits an everywhere non-vanishing direction field V such that each point is assigned one of a pair $(A, -A)$, where A is a non-zero vector. Given the existence of the direction field V, a Lorentz metric g is then defined, for all vectors B, C by

$$g(B, C) = g^*(B, C) - 2g^*(A, B)g^*(A, C)/g^*(A, A) \tag{1}$$

(see Hawking and Ellis, 1971, p. 39). We see that $g(A, A) > 0$, so A is timelike, and if $g^*(A, B) = 0 (A \perp B)$, $g(B, B) < 0$ and B is spacelike.

The question becomes, when does a manifold admit an everywhere non-vanishing direction field? *When the manifold is not compact, the answer is always; when it is compact it admits such a field iff its Euler characteristic is zero* (Hawking and Ellis, 1971, pp. 40, 52). It is this latter result concerning compact manifolds that will be crucial in our discussion of topology change, but first we must get acquainted with the Euler characteristic.

3. The Euler Characteristic

In the first instance, the Euler characteristic is a property of polyhedra.[2] In three dimensions a polyhedron has two-dimensional faces (or sides), one-dimensional edges that bound the faces, and zero-dimensional vertices that bound edges. We can think of an edge as a one-dimensional face and a vertex as a zero-dimensional face. In higher dimensions, a polyhedron can also have n-dimensional faces which are bounded by $(n-1)$-dimensional faces, $(n-1)$-dimensional faces bounded by $(n-2)$-dimensional faces, etc. Let F_n be the number of n-dimensional faces for a given polyhedron in D dimensions. Then the Euler characteristic χ of that polyhedron is given by

$$\chi = \sum_1^n (-1)^n F_n. \tag{2}$$

It is a theorem that for a given compact manifold M, all the polyhedra topologically equivalent (homeomorphic) to M have the same Euler characteristic. Therefore, we define the Euler characteristic of M to be the

[2] Interestingly, Descartes first discovered the Euler characteristic (though it was of course Euler who recognised its significance and developed it).

Euler characteristic of any polyhedron homeomorphic to it. However, note that all of the polyhedra mentioned in the theorem must be 'well-behaved'. Specifically, every n-face has the same number of bounding $(n-1)$ faces in any polyhedron. A few examples are helpful.

A cube has six 2-faces, twelve 1-faces, and eight 0-faces, so $\chi(\text{cube}) = 8-12+6=2$, while a tetrahedron has four 2-faces, six 1-faces, and four 0-faces, and $\chi(\text{tetrahedron}) = 4-6+4=2$. Both are homeomorphic to a sphere, whose Euler characteristic is therefore 2. On the other hand, the rectangular 'donut' (or torus) has sixteen 2-faces, thirty-two 1-faces, and sixteen 0-faces, so $\chi(\text{torus}) = 0$, in agreement with the fact that the torus and sphere are not homeomorphic.

For a higher-dimensional example that will be useful later, consider a hypercube in four-dimensional Euclidean space. It has eight cubical 3-faces, which when they join together form two 2-faces, three 1-faces, and four 0-faces, so $\chi(\text{four-cube}) = (\text{0-faces}) - (\text{1-faces}) + (\text{two-faces}) - (\text{three-faces}) = 8\times 8/4 - 8\times 12/3 + 8\times 6/2 - 8 = -8+8(3-4+2) = 0$. Since the hypercube in four dimensions is homeomorphic to the three-sphere, S^3, $\chi(S^3)=0$. Note that $\chi(\text{square}) = -4+4=0$, so $\chi(S^1)=0$ as well. Thus we have the series: $\chi(S^1)=0$, $\chi(S^2)=2$, $\chi(S^3)=0$. As the reader might care to verify, this generalises to

$$\chi(S^{2n}) = 2, \qquad \chi(S^1) = \chi(S^{2n+1}) = 0, \quad \text{for } n \geqslant 1 \tag{3}$$

which is a fact we will use later.

Quite apart from the question of topology change, however, this fact has interesting consequences for spacetime. It is a standard textbook fact, for example, that de Sitter spacetime can be represented in five-dimensional Minkowski spacetime as a (pseudo-) hyperboloid of revolution—the locus of points an equal spacetime distance from a given point. This is not a compact spacetime. For ease of visualisation, think of the two-dimensional case in three-dimensional Minkowski spacetime. Since it is a locus of points at equal 'distances' from a given point, it is a kind of (hyper)sphere. But four-dimensional de Sitter spacetime does not have the topology of a four-sphere. It is not compact and not simply connected.

Indeed, we know from the above that a compact manifold M admits a Lorentz metric iff $\chi(M)=0$. Consequently, there is no spacetime with the topology of the four-sphere, or more generally, with the topology of any even-dimensional sphere. But of course there are even-dimensional compact spacetimes. We have already seen that the Euler characteristic of the torus equals zero (and it is intuitively obvious, anyway, that T^2 admits an everywhere non-vanishing direction field), and in four dimensions the same is true of $S^3\times S^1$ and $S^1\times S^1\times S^1\times S^1$. This is because

$$\chi(A\times B) = \chi(A)\chi(B) \tag{4}$$

and $\chi(S^1)=\chi(S^3)=0$. However, since S^1 is not simply connected, neither are $S^3\times S^1$ nor $S^1\times S^1\times S^1\times S^1$, and in four dimensions there are no compact simply connected spacetimes. Once above four dimensions, however, there are

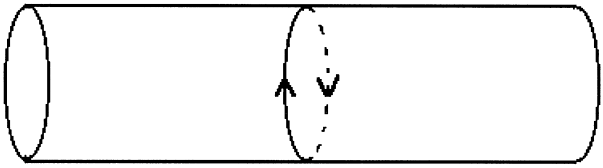

Fig. 3. *Two-dimensional Minkowski spacetime rolled up along its spacelike axis.*

compact simply connected spacetimes at each even dimension $2(2n + 1)$; namely $S^{2n+1} \times S^{2n+1}$, where $n \geqslant 1$. Here we are relying on the facts that the product of simply connected manifolds are simply connected, and that the odd-dimensional spheres do admit a Lorentz metric.

This fact about the odd-dimensional spheres can be seen quite intuitively by looking at the embedding of S^{2n-1} in $2n$-dimensional Euclidean space. Let $x_1 \ldots x_{2n}$ be Cartesian coordinates for such a space with S^{2n-1} centred at the origin $x_i = 0$. Then, as the reader can easily verify, for points $(x_i \ldots x_{2n})$ on the sphere, $V = (x_2, -x_1, x_4, -x_3 \ldots x_{2n}, -x_{2n-1})$ is an everywhere non-vanishing vector field on S^{2n-1}. Using V, we can define a Lorentz metric on S^{2n-1} using the Euclidean metric (we have a vector field rather than just a direction field because spheres are orientable).

We can say something about spacetimes allowing backward time travel as well. It is not hard to find spacetimes containing closed timelike curves, even flat spacetimes such as two-dimensional Minkowski spacetime rolled up around its spacelike axis, pictured in Fig. 3.

But this is not simply connected. Furthermore, there is a spacetime with the same metric that does not have closed timelike curves: two-dimensional Minkowski spacetime. And that is generally true for spacetimes whose closed timelike curves depend on the non-simple connectedness of the spacetime. Is there a four-dimensional simply connected spacetime that contains closed timelike curves? It follows that there are from the fact that S^3 is simply connected (Geroch, 1967, p. 782). Namely, since S^3 admits a globally defined Lorentz metric, the simple theorem that a compact spacetime contains at least one closed timelike curve (Hawking and Ellis, 1971, p. 189) implies that a spacetime with the topology of S^3 contains a closed timelike curve. A four-dimensional spacetime with the topology $S^3 \times R$ is simply connected (since both S^3 and R are connected), and if R is spacelike, it thus contains closed timelike curves, e.g. Taub–NUT–Misner spacetime. Gödel's famous spacetime contains closed timelike curves (through every point), is simply connected, and is not even compact, being topologically Euclidean!

4. Homology

We can gain further insight into the Euler characteristic from homology theory. This is usually developed in terms of 'polyhedra' (or simplicial

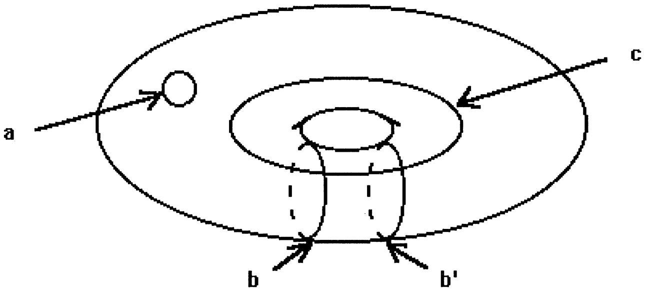

Fig. 4. *The closed curves of a torus.*

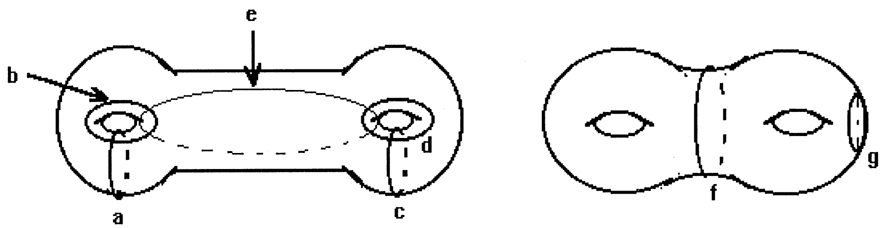

Fig. 5. *The closed curves of a double torus.*

complexes) homeomorphic to a given manifold, but we will work intuitively from the manifold itself. Let us begin by considering curves in two-dimensional compact manifolds. Homology is interested in the number of independent closed curves (curves without boundary) which are not themselves the boundary of a two-dimensional region. The number of such curves is called a manifold's 'Betti number' b_n (of the corresponding dimension) or order of connectivity, and it is a topological invariant of manifolds. When we look at the number of such independent closed curves in two dimensions, we are examining the one-dimensional Betti number b_1. Two such curves are not independent if together they form the boundary of a two-dimensional region. On the sphere, all closed curves bound an area and are homologous (not independent), while on the torus there are three kinds of closed curves (see Fig. 4).

Curves of type a bound an area, while b and c do not. b and b' are not independent while b and c are independent, so the number of independent closed curves that do not bound an area, $b_1 = 2$. Consider now the double torus (Fig. 5). a, b, c, d are four independent closed curves that do not bound areas. e might look like a fifth such curve, but a, e, c together bound an area and are consequently not independent. So here $b_1 = 4$. Notice also that f bounds an area since f and g are not independent and g bounds an area. f is a counterexample to what the earlier examples might suggest, that a curve bounding an area is equivalent to the curve being continuously contractible to a point. To think this is to confuse homology with homotopy.

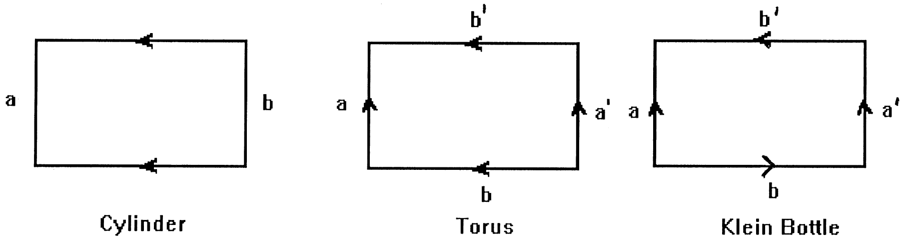

Fig. 6. *The finite cylinder, torus and Klein bottle represented on the plane.*

Next, consider boundary-less surfaces (closed surfaces) which do not bound a volume.[3] Since the sphere and torus are themselves two-dimensional, there is no volume as part of themselves. Therefore, in each case there is one such independent surface and the two-dimensional Betti number $b_2 = 1$. The higher b_n then concern the number of independent closed 'n surfaces' that do not bound an $(n+1)$-dimensional region of the manifold in question. Thus, $b_n = 0, n > 2$ for two-dimensional manifolds. That leaves b_0. Since two points together bound many curves, we can think of them as being non-independent. Thus, in a connected manifold $b_0 = 0$ or 1. It turns out that it is always 1 (Nakahara, 1990).

The Euler–Poincaré theorem then tells us that for a manifold M,

$$\chi(M) = \sum_1^n (-1)^n b_n. \tag{5}$$

For a sphere, $b_0 = 1, b_1 = 0, b_2 = 1$, so

$$\sum_1^n (-1)^n b_n = 1 - 0 + 1 = 2 = \chi(S^2) \tag{6}$$

and for the torus

$$\sum_1^n (-1)^n b_n = 1 - 2 + 1 = 0 = \chi(T^2), \tag{7}$$

agreeing with our earlier result. For the higher spheres S^n note that $b_0 = 1, b_n = 1, b_m = 0, 0 < m < n$, since S^n is simply connected. For S^{2n}

$$\chi(S^{2n}) = b + b = 2, \qquad \chi(S^{2n-1}) = b_0 - b_n = 0, \tag{8}$$

again agreeing with our earlier results.

Lastly, let us compare the finite cylinder, torus and Klein bottle. These can be represented in the plane of the paper using the indicated identifications (Fig. 6). For the cylinder we have $b_0 = b_1 = 1$; closed curves that 'go around' the cylinder do not (by themselves) bound an area, and all such curves are homologous, including the edges. But what about b_2? The cylinder is not a boundary-less surface. The curves a and b together are its boundary. So $b_2 = 0$ and

[3] More precisely, we are interested in closed oriented submanifolds of some manifold M which are not themselves boundaries of M.

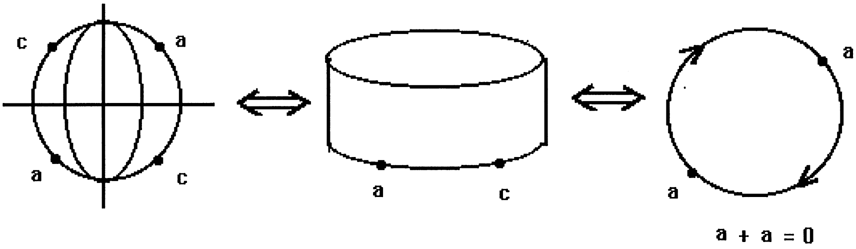

Fig. 7. *The projective plane as a disk.*

χ(cylinder) $= 1 - 1 + 0 = 0$. For the torus, as we have already seen, $\chi(T^2) = 0$. The interesting point here is how the fact that the surface of the torus is boundary-less is represented. In our picture a, b, a', b' are a kind of formal boundary for the torus. However, notice that a is a', b is b' (by the identifications), and that to go around the area of the torus you have to traverse a in one direction and then itself (a') in the other, and similarly for b. The net boundary is zero since its segments cancel. But that is not the case for the Klein bottle. Although compact and without boundary in the topological sense, it has a formal boundary in the sense that two trips around the closed curve b (a and a' cancel) enclose the surface. Consequently $b_2 = 0$ while $b_1 = 1$—not 2 as in the torus. The reason is that a and a' bound an area (the whole surface) but $a = a'$. Closed curves a such that na bound an area (in our formal sense) for some n do not contribute to b_1.

Important examples of this behaviour are the real projective 'planes'. For a given n, RP^n is topologically the n-sphere with anti-podal points identified. The projective plane, RP^n, can be represented as a disc whose opposing boundary points are identified (Fig. 7). Put differently, we get a representation of RP^2 by shrinking b to zero in our picture of the Klein bottle. By reasoning analogous to that of the Klein bottle, $b_0 = 1, b_1 = 0, b_2 = 0$ so $\chi(RP^2) = 1$. Analogously, a solid ball whose opposing boundary points are identified can represent RP^3. The Betti numbers are the same as in RP^2 with the addition of $b_3 = 1$ so $\chi(RP^3) = 0$. As in the case of spheres, this pattern repeats:

$$\chi(RP^{2n}) = 1, \qquad \chi(RP^{2n+1}) = 0. \tag{9}$$

That the Euler characteristic of the odd-dimensional spheres and projective planes equal zero is a special case of the general result that *all* odd-dimensional compact manifolds have $\chi = 0$!

Like the familiar Möbius strip, RP^2 is non-orientable. A right-handed figure can be changed into its mirror image counterpart just by moving it around in RP^2. In Fig. 8, move the figure to the left and it becomes the mirror image counterpart of its earlier self. While all the even-dimensional projective planes are non-orientable, remarkably, the odd-dimensional ones are orientable. In the case of RP^3 this can be easily visualised in terms of the solid ball with opposing boundary points identified.

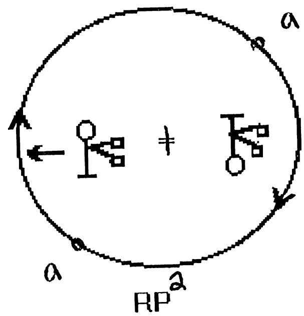

Fig. 8. *The non-orientability of RP^2.*

5. Topology Change in Classical Relativistic Spacetimes

Let us now return to the topology change of physical space within the context of relativistic spacetimes. For purposes of simplicity we restrict ourselves to closed physical space, i.e. compact without boundary. The standard way of setting up the problem is as follows (cf. Borde, 1994; Geroch, 1967; Sorkin, 1986a). Let Σ_1 and Σ_2 be closed $(n-1)$-dimensional manifolds, which we are thinking of as representing space at two different instants (relative to some frame). We want to know if there is an interpolating n-dimensional spacetime whose boundary is the disjoint union of Σ_1 and Σ_2 and with respect to which Σ_1 and Σ_2 are both spacelike.[4] If there is and if Σ_1 is not homeomorphic to Σ_2, then we have a case of topology change. This question is then approached in two steps. The first step is to ignore the question of the Lorentz metric and ask whether, given two $(n-1)$-dimensional closed manifolds Σ_1 and Σ_2, there is an n-dimensional compact manifold M for which Σ_1 and Σ_2 are the disjoint boundary. If such a manifold exists it is a *cobordism* for Σ_1 and Σ_2, and Σ_1 and Σ_2 are *cobordant*. When M exists the second step is to ask whether a Lorentz metric can be put on M with respect to which Σ_1 and Σ_2 are spacelike.

For the first stage we will just quote the results and refer the reader to the literature (e.g., Stong, 1968) . However, it is worth pointing out that there is much of interest here due to the fact that for a given dimension n, the equivalence class of cobordant manifolds has the structure of an abelian group (Milnor and Stasheff, 1974); and composition of cobordisms is not commutative (Baez, 2000). In the case we are most interested in, that of three-dimensional space, the answer is that *any* two closed three-manifolds are cobordant, whether or not they are orientable. This is Lickorish's theorem. However, other dimensions are

[4] Considering spacelike three-spaces that form the boundary of M ensures that we do not obtain trivial cases of topology change resulting from strange choices of three-spaces. For examples of 'poor' choices of three-spaces, see Gowdy (1977). One may also care to examine, as Gowdy does, topology change between null surfaces; however, because this sometimes results in counter-intuitive instances of topology change, we stick with topology change between spacelike surfaces.

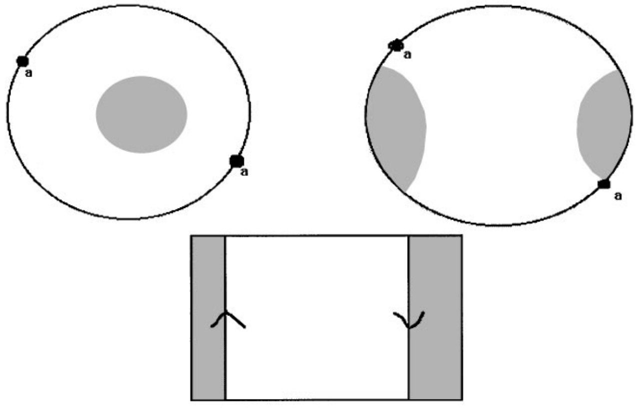

Fig. 9. *Topological equivalence of Möbius strip and projective plane with the interior of a disc removed.*

interesting as well. For unoriented closed manifolds, any two are cobordant in dimensions 1, 3 and 7, while for oriented manifolds with an oriented cobordism, any two are cobordant in dimensions 1, 2, 3, 6 and 7.

As for the second step, we already know the answer. If Σ_1 and Σ_2 are cobordant, then the cobordism M admits a Lorentz metric iff the Euler characteristic of M, $\chi(M)$, is zero. But there is a qualification. We will want there to be an everywhere non-zero vector field on the cobordism M that points outward everywhere on Σ_1 (say), and inward everywhere on Σ_2, so that Σ_1 is the initial time and Σ_2 the final time. *For dimension n of the cobordism, $\chi(M) = 0$ is still sufficient for even n. But for odd n, $\chi(M) = 0$ automatically, though there is an additional selection rule discovered independently by Reinhardt (1963) and Sorkin (1986a):* $\chi(\Sigma_1) = \chi(\Sigma_2)$. Let us look at some examples.

In two dimensions the only compact manifolds that admit a Lorentz metric are the torus, cylinder, Klein bottle and Möbius strip. Except for the Möbius strip, we have already seen that $\chi = 0$. But the homology of the Möbius strip will be the same as the cylinder. Since the torus and the Klein bottle are closed, they represent (in the usual representation), one-dimensional space evolving into itself, while the cylinder (again, as usually represented), is the case of no topology change. The Möbius strip, however, is more interesting. Notice that the Möbius strip is (topologically) equivalent to the projective plane with the interior of a disc removed (Fig. 9). In this form we can easily see that it is an example of one-dimensional closed space evolving into or out of nothing (Sorkin, 1986a) (Fig. 10). Also, we can put a different Lorentz metric on the cylinder so that we get pair annihilation or pair creation of S^1 (Borde, 1994) (Fig. 11).

We can also put the two together to get a Lorentz metric on the torus in which $\phi \to \phi$ by way of two copies of S^1. Thus we *do* have examples of space changing its topology in two-dimensional spacetime, but only in the sense of space appearing or disappearing from nothing.

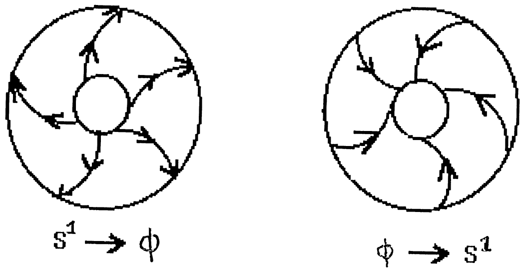

Fig. 10. *Closed space evolving into or out of nothing.*

Fig. 11. *Pair creation/annihiliation of* S^1.

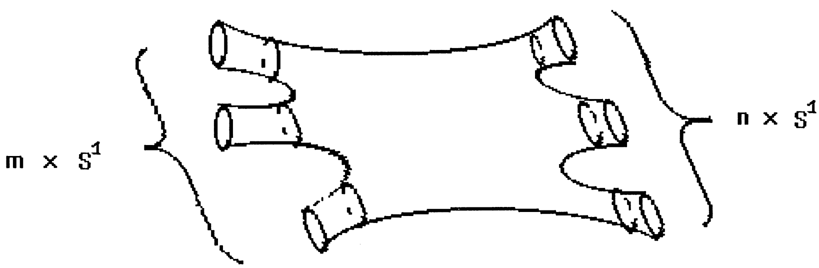

Fig. 12. *Two-dimensional space fusion/division.*

In fact, since S^1 is the only closed one-dimensional manifold ($RP^1 = S^1$), the only case of topology change in the two-dimensional spacetime that does not involve annihilation or creation would be space dividing or fusing as pictured in Fig. 12. To compute its Euler characteristic, notice that the $(m + n)$ circles do bound an area, but $(m + n - 1)$ of them do not. Therefore $b_0 = 1$, as always, $b_1 = m + n - 1$, $b_2 = 0$ and

$$\chi(\text{two-dimensional fusion/division}) = 1 - (m + n - 1) + 1 = 2 - (m + n). \tag{10}$$

For $m + n > 2$, $\chi < 0$ so a Lorentz metric cannot be placed on the manifold. In two-dimensional spacetime closed space cannot divide or fuse.

In three-dimensional spacetimes we can also give a complete account for oriented two-dimensional closed manifolds. Any two closed oriented two-dimensional manifolds are cobordant, and any three-dimensional compact manifold has $\chi = 0$ since it is of odd dimension. Therefore any two-dimensional closed oriented manifolds are *Lorentz cobordant* (Yodzis, 1972), that is, their cobordism admits a Lorentz metric. However, if we demand the additional structure required for regarding one of the two-dimensional spaces as evolving into the other, then the Reinhardt–Sorkin selection rule $\chi(\Sigma_1) = \chi(\Sigma_2)$ must also be satisfied. But every closed oriented two-manifold is topologically equivalent to a sphere with n-handles, where $n \geqslant 0$. In other words, each such manifold is equal to T_n, a torus with n holes in it. For an n-torus, $b_0 = 1, b_1 = 2n, b_2 = 1$, so

$$\chi(T_n) = 1 - 2n + 1 = 2 - 2n, \tag{11}$$

which implies $\chi(T_n) \neq \chi(T_m), m \neq n$. The selection rule forbids any T_n evolving into a $T_m, m \neq n$. Note also that since $\chi(S^2) = 2$, the division or fusion of two-dimensional space—$\chi(S^2 \cup S^2) = \chi(S^2) \cup \chi(S^2)$—has an Euler characteristic equal to four.

Unlike two-dimensional closed manifolds, any two closed three-manifolds are cobordant, either by an oriented or unoriented cobordism. But the dimension of the cobordism is even, so we no longer have the automatic $\chi = 0$ as we had with three-dimensional cobordisms. Nonetheless, we can show in this case that any two closed three-manifolds are Lorentz cobordant. The proof relies on the idea of the connected sum, $M \oplus N$, of two n-dimensional manifolds M and N. To form $M \oplus N$, remove an n-dimensional open ball from both M and N, and then identify the two resulting boundaries. As an example, we can look at the earlier case of the projective plane (RP^2) with a hole in it. This is the direct sum of (RP^2) and a closed disc D^2. Since $\chi(RP^2) = 1$, $\chi(D^2) = 1$, forming the direct sum of these two reduces the Euler characteristic of RP^2 to zero (that of the Möbius strip). That is, $\chi(RP^2 \oplus D^2) = \chi(RP^2) + \chi(D^2) - 2$. And in two dimensions this relationship

$$\chi(M_1 \oplus M_2) = \chi(M_1) + \chi(M_2) - 2 \tag{12}$$

holds quite generally, as is readily seen from polyhedra, since to form $M_1 \oplus M_2$ we join a face of M_1 to that of M_2. Taking the faces to be squares, on joining we lose two faces, four vertices and four edges, so $\chi_1 + \chi_2$ changes by $d\chi = dV - dE + dF = -4 - (-4) - 2 = -2$. Still more generally, this relationship holds for compact manifolds in any *even* dimension. Now, in four dimensions, we have (Geroch, 1967):

$$\begin{aligned} \chi(S^2 \times S^2) &= \chi(S^2)\chi(S^2) = 2 \times 2 = 4, \\ \chi(S^1 \times S^3) &= \chi(S^1)\chi(S^3) = 0 \times 0 = 0, \\ \chi(CP^2) &= 3. \end{aligned} \tag{13}$$

Therefore, if $\chi(M_1) = -|2n|$, we can obtain a compact manifold (with the same boundary) of $\chi = 0$ by successive direct sums with $S^2 \times S^2$. If $\chi(M) = |2n|$, we do the same with $S^1 \times S^3$, while if $\chi(M)$ is odd, we first get $\chi(M + CP^2) =$ even and proceed as before. (Here CP^2 is the projective plane in two complex dimensions. $\chi(RP^4)$ is also odd but RP^4 is not orientable.)

It follows that since any two closed three-manifolds are cobordant, by the above procedure they can be made Lorentz cobordant. And therefore, if Σ_1 and Σ_2 are oriented closed three-manifolds, there exists a compact four-dimensional spacetime for which Σ_1 (say) is the initial time and Σ_2 the final time. From the topological point of view we have (hardly) no restrictions on three-dimensional space changing its topology! And note: this includes a three-dimensional space dividing and fusing as long as the four-dimensional spacetime is connected (Borde, 1994). For instance, an explicit example of a Lorentzian spacetime with a branching off of a bifurcating universe is $CP^2 - 3$ balls. $\chi(CP^2) = 3$, and the Euler characteristic of each 4-dimensional ball is 1, so $\chi(CP^2 - 3 \text{ balls}) = 0$. And it is possible to find a timelike vector field such that it points out of the S^3's that represent the past and into the one that represents the future. Thus the sort of example we used against Descartes of space dividing turns out to be not only conceivable but also (kinematically) physically possible, in some sense.

Is this the end of the story concerning topology change in four-dimensional spacetime, at least as far as pure topology goes? The answer is 'no', due to a famous theorem of Geroch (Geroch, 1967). The theorem asserts that if closed Σ_1 and Σ_2 do not have the same topology then the spacetime M either contains closed timelike curves or is not temporally orientable. Let us discuss this.

Geroch's theorem tells us that topology change comes with a cost. But how high is the cost? Is it so high that it makes it unlikely that our world's spatial topology may change? To adequately answer this question, we would need to investigate the effect 'realistic' energy conditions have on topology change—in particular, we would need to look at a famous theorem of Tipler to the effect that realistic energy conditions preclude topology change. However, the interpretation of these conditions is controversial and there is more than enough to discuss within the realm of pure topology, so let us postpone the discussion of realistic energy conditions to another paper.

Before considering ways of evading Geroch's theorem, let us consider how objectionable admitting closed timelike curves or temporal non-orientability is. We think neither option is as bad as is commonly thought. First, consider temporal orientability. Remember that a spacetime is temporally orientable iff it admits an everywhere continuous and non-vanishing timelike vector field. If a spacetime is not temporally orientable, then it is not possible to divide light cones into past and future in a globally consistent and continuous way. Either the direction of time at a point depends on your worldline through that point or the direction of time changes discontinuously in some spacetime regions. How serious this is depends on your views about the nature and origin of time's (so-called) arrow. Here we want to make just three observations.

First, even a non-temporally orientable spacetime is locally temporally orientable in the sense that about each point there is an open neighbourhood that is temporally orientable (as long as we stay in that neighbourhood). Insofar as the

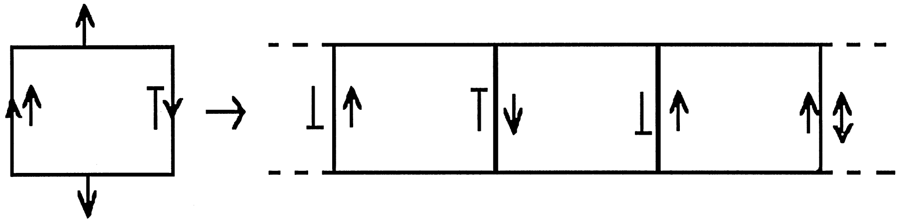

Fig. 13. *Möbius strip spacetime.*

observed universe is temporally oriented, local temporal orientablility might be enough even if spacetime is not temporally orientable.

Second, it is often pointed out that non-temporally oriented spacetimes are not simply connected, and therefore there exists a spacetime with the same metric which is simply connected (the universal covering space of the spacetime) and therefore temporally orientable. The relevance of this remark is that if we had reason to think that spacetime was not temporally orientable, we would know that there is a spacetime that is (in some sense) compatible with the same observational evidence and yet is temporally orientable. So we are never forced by observation (alone) to hold that spacetime is not orientable. This *may* be, but it does not follow that the orientable spacetime will be temporally *oriented*. For example, consider the Möbius strip spacetime pictured in Fig. 13, with asymmetric processes T and $\uparrow$. Because of the indicated identifications, it is not simply connected. The universal covering space is the plane, which we can construct from repeated strips of cut open Möbius strip spacetime. Locally, it does look like Möbius strip spacetime. It is simply connected and temporally orientable. But just as obviously it is not temporally oriented, if temporal orientation is a matter of how asymmetric processes (broadly construed) are distributed in spacetime (let $\uparrow$ = a tree going from seed $\rightarrow$ tree, T = a sugar cube in coffee $\rightarrow$ dissolved in coffee). In any case, if we only require local orientability as in point one, this will not be a problem.

Third, note that non-temporal orientability and the existence of closed timelike curves are independent properties. Cylinder spacetime (with a spacelike axis) and Gödel spacetime are temporally orientable and contain closed timelike curves; elliptic de Sitter—identify antipodal points in de Sitter spacetime represented as a hyperboloid of revolution in flat spacetime—is non-temporally orientable but does not contain closed timelike curves.

Point one suggests that lack of temporal orientability is not that serious. Point three, then, suggests that we can still have topology change without the price of closed timelike curves by admitting non-temporal orientability. And indeed, Sorkin (1986b) shows that giving up temporal orientability eliminates the need for the Reinhardt–Sorkin selection rule in odd dimensions. Unfortunately, this move to non-orientable spacetimes does not buy us much freedom. As Borde shows, there are still severe restrictions on topology change: apart from cases in which space is appearing or disappearing into nothing, topology change *always* involves the existence of closed timelike curves (so long as the spacetime is compact). This route to topology change, therefore, is not as interesting as we might have hoped.

However, that is not to say there is nothing of interest here (Friedman, 1991). It is well known that the Hartle–Hawking 1983 model of the universe is one wherein the universe has no past boundary. One can devise simple examples of this model that have topology change without closed timelike curves (Sorkin, 1986a, b). One example is (RP^4 − ball), since $\chi(RP^4) = 1$, $\chi(RP^4 - \text{ball}) = 0$. This is the four-dimensional analogue of the Möbius strip, as the reader will recall from Figs. 9 and 10. The Möbius strip has a single boundary that we may take as the final hypersurface of the spacetime. By imposing a timelike direction field normal to this boundary we describe a Lorentzian model of the Hawking–Hartle cosmology with topology change but no timelike curves.

How bad is the other option, closed timelike curves? The physics community, even the quantum gravity community who are so keen to allow topology change, often view spacetimes with closed timelike curves as physically unreasonable. Thus Hawking and others have hoped to prove Chronology Protection Theorems that would rule out their existence. Closed timelike curves are (mistakingly) viewed as opening the door to logical paradoxes, such as the infamous 'Grandfather paradox', whereby a time traveller following a closed timelike curve goes back in time and kills his grandfather before the grandfather impregnates the traveller's grandmother. Of course, as those who have read Lewis (1976) or Earman (1995) know, there is no real danger here. Logical contradictions cannot happen. If you go back in time to kill your grandfather, you already tried it. Your actions in the past may have contributed to the present being the way it is, but they cannot undermine it. Since you are around, evidently you did not kill your grandfather. And your failure to be able to bring about the impossible does not imply consistent time travel scenarios cannot occur.

A less philosophical source of concern about closed timelike curves is the idea that it will be impossible to define a quantum field on a spacetime with such curves. Anderson and DeWitt (1986), for instance, show that two-dimensional 'trousers' bifurcating spacetime ($S^2 - 3$ discs) is unable to support a consistent field theory. However, as the reader will recall from our discussion of two-dimensional topology change, topology change on such a spacetime is impossible. The authors recognise this, of course, and define a Lorentz metric on the spacetime that is singular at a point (more on this in a moment). The resulting spacetime experiences an infinite burst of energy at the 'crotch' singularity when one imposes a scalar quantum field on it. However, as Friedman (1991) points out, the trouble with defining a field on this indecent spacetime arises from the singularity, not the closed timelike curves! So this example cannot be used as part of an argument against topology change.

Without delving into the issue of the compatibility of closed timelike curves and energy conditions, it seems the most we can say against them is that we do not see them. But this only militates against common 'medium-sized' closed time like curves. Large-scale or microscopic-sized closed timelike curves may still exist without our knowing it.

Not sharing our open-minded attitude toward closed timelike curves, the physics community keen to accommodate topology change is left in a bit of a pickle. The now almost standard response to Geroch's theorem is to evade it entirely with a drastic move shown to be available by Horowitz (1991) and others. As Horowitz shows, if one lets the metric go to zero at isolated points, that is, if one considers degenerate metrics, topology change can occur without closed timelike curves (even on temporally orientable manifolds). Because this move essentially rejects the 'equivalence principle' at select points, it takes us outside the realm of classical general relativity and thus our present discussion. Yet we can mention that this approach to topology change faces many challenges. Central among them is defining the causal structure of spacetime ordinarily associated with a non-vanishing Lorentz metric. Work on this and on related issues concerning the characterisation of these singular points using Morse theory are being actively pursued (see e.g. Dowker and Surya, 1998).

Another means of escaping Geroch's theorem is by moving to topology change occuring within non-compact regions of spacetime. (Geroch's theorem holds for compact spacetimes, but it can easily be extended to arbitrary spacetimes, so long as the topology changing region occurs within a compact set.) These cases are not as interesting as the compact case to those researching topological geons or to some other programmes, where one thinks of the topology change as occuring 'within a box'. However, they may be of interest to cosmological questions, such as the one with which we began—namely, whether the universe could divide into two new ones. Along these lines Karsnikov (1995) has shown that there exist topology changing non-compact spacetimes that do not violate the so-called weak energy condition or the stable causality condition.

Topology change, therefore, is not without a price. In the non-compact case, many interesting topological transitions are lost. And in the compact case, for topological transitions not originating or terminating in the null set, we require either a degenerate metric or closed timelike curves.

Finally, let us make a few brief remarks about spacetime dimensions greater than four. These are of interest because extra spatial dimensions are often involved in theories that attempt, in some sense, to unify the different fundamental interactions. We mention two such theories. First, Kaluza–Klein theory attempts to give a geometrical explanation of gauge symmetries in terms of geometrical symmetries of compact extra dimensions. In the original theory, five-dimensional spacetime is rolled up along the extra spacelike dimension, giving it the topology $R^4 \times S^1$. Electromagnetic gauge transformations then correspond to a geometrical transformation along S^1. If we want to include the weak interaction (the $\mathrm{U}(1) \times \mathrm{SU}(2)$ electroweak theory) we need at least three more compact dimensions. If we want the strong interaction included too, a total of at least seven extra compact dimensions is needed, giving spacetime at least eleven dimensions. Assuming compact physical space and orientability, the selection rule $\chi(\Sigma_1) = \chi(\Sigma_2)$ applies to topology change of the ten-dimensional compact spacelike dimensions. Second, in string theory dynamical considerations (quantising, interactions) single out twenty-six spacetime dimensions for the bosonic string and ten for the fermionic string. Since spacetime here is of even dimension, the selection rule does not apply.

There is much more that can be said about topology change. In particular, we have not considered what restrictions the field equations of general relativity impose on topology change. In general, the geometry of spacetime alone does not significantly restrict the available topologies. However, when one considers realistic matter-energy sources, spinor structure, etc., there are definite results, including Tipler's famous theorem. But these issues are quite controversial and anyway go beyond pure topology. For now we are content to have displayed some of the richness and life a question dating back to Aristotle still enjoys in contemporary physics.[5]

References

Anderson, A. and DeWitt, B. (1986) 'Does the Topology of Space Fluctuate?', *Foundations of Physics* **16,** 91–105.

Baez, J. (2000) 'Higher-Dimensional Algebra and Planck-Scale Physics', in C. Callender and N. Huggett (eds), *Physics Meets Philosophy at the Planck Scale* (Cambridge: Cambridge University Press).

Borde, A. (1994) 'Topology Change in Classical General Relativity', *Los Alamos E-print Archives* gr-qc/9406053.

Dowker, F. and Surya, S. (1998) 'Topology Change and Causal Continuity', *Physical Review* **D 58,** 124019.

Earman, J. (1995) *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes* (Oxford: Oxford University Press).

Friedman, J.L. (1991) 'Spacetime Topology and Quantum Gravity', in A. Ashtekar and J. Stachel (eds), *Conceptual Problems of Quantum Gravity* (Boston: Birkhauser Press).

Geroch, R. (1967) 'Topology in General Relativity', *Journal of Mathematical Physics* **8,** 782–786.

Gowdy, R. (1977) 'Instantaneous Cauchy Surfaces, Exploding Black Holes, and Topology Change', *Journal of Mathematical Physics* **18,** 1798–1801.

Hartle, J. B. and Hawking, S. W. (1983) 'Wavefunction of the Universe', *Physical Review* **D 28,** 2960–2975.

Hawking, S. and Ellis, G. (1971) *The Large-Scale Structure of Spacetime* (Cambridge: Cambridge University Press).

Horowitz, G. (1991) 'Topology Change in Classical and Quantum Gravity', *Classical and Quantum Gravity* **8,** 587–601.

Krasnikov, S. V. (1995) 'Topology Change Without Any Pathology', *General Relativity and Gravitation* **27,** 529–536.

Lewis, D. (1976) 'The Paradoxes of Time Travel', *American Philosophical Quarterly* **13,** 145–152.

Milnor, J. and Stasheff, J. (1974) *Characteristic Classes* (Princeton: Princeton University Press).

Nakahara, M. (1990) *Geometry, Topology and Physics* (Bristol: Adam Hilger Publishing).

[5] We would like to thank Nick Huggett for useful comments on a draft of this paper.

Quinton, A. (1962) 'Spaces and Times', *Philosophy* **27,** 130–147.
Reinhardt, B. L. (1963) 'Cobordism and the Euler Number', *Topology* **2,** 173–177.
Stong, R. (1968) *Notes on Cobordism Theory* (Princeton: Princeton University Press).
Sorkin, R. (1986a) 'Topology Change and Monopole Creation', *Physical Review* **D 33,** 978–982.
Sorkin, R. (1986b) 'Non-Time Orientable Lorentzian Cobordism Allows for Pair Creation', *International Journal of Theoretical Physics* **25,** 877–881.
Yodzis, P. (1972) 'Lorentz Cobordism', *Communications in Mathematical Physics* **26,** 39–52.
Weingard, R. (1976) 'On the Unity of Space', *Philosophical Studies* **29,** 215–220.