CONSEQUENCES OF SPACETIME TOPOLOGY *

Rafael D. Sorkin
Department of Physics
Syracuse University
Syracuse, NY 13244-1130

In the last decade or so, topological questions have increasingly attracted the interest of physicists. Indeed the amount of work devoted to topological issues has grown so explosively that even a review talk restricting itself to the domain of General Relativity can't hope to give an adequate account of recent work. The review I am about to give will therefore be very superficial, although I *will* try to dwell a bit longer on one or two areas that I am most familiar with. As indicated by my title, I will also try, insofar as I can, to emphasize the physical *consequences* of spacetime topology−the possible effects that have excited much of the recent intense interest of which I spoke. Finally let me apologize in advance for the incompleteness of the references. Rather than attempt a comprehensive bibliography for this lecture, I am just going to give a relatively small number of unsystematically chosen citations, hoping in particular to mention a few papers that you can follow up for more information[1)].

Topology enters General Relativity through the fundamental assumption that spacetime exists and is organized as a manifold. This means in the first place that spacetime has a well-defined dimension, but it also carries with it the inherent possibility of modified "patterns of global connectivity", such as distinguish a sphere from a plane, or a torus from a surface of higher genus. Such modifications can be present in the spatial topology without affecting the time direction, but they can also have a genuinely spacetime character in which case the spatial topology *changes with time*, if indeed it is well-defined at all.

* Published in A. Coley, F. Cooperstock, B. Tupper (eds.), Proceedings of the Third Canadian Conference on General Relativity and Relativistic Astrophysics held May, 1989, Victoria, Canada, 137-163 (World Scientific, 1990).

Strictly speaking the spacetime of General Relativity is a *differentiable* manifold rather than just a topological one, the extra structure being required in order that the manifold can play host to a Riemannian (or rather Lorentzian) metric. The contemplation of a modified topology thus leads naturally to the idea of a modified differentiable structure, and indeed the latter has recently shown signs of mathematical life, especially in four dimensions. However, nobody seems to understand yet the physical consequences of replacing, for instance, the standard $\mathbf{R}^4$ by one of its more exotically differentiable variants. In any case there will be no time to say more about such possibilities in this talk.[2)]

Another set of related possibilities which I will have to neglect involves the topology, not solely of spacetime itself, but of the non-gravitational fields which live in it. Topological effects involving such fields have proved important in many settings, as have effects involving the topology of quantum configuration spaces and of non-trivial bundles over these configuration spaces. In themselves such possibilities do not involve spacetime topology, but when the latter is non-trivial, it will in general "interact" with them to produce new possibilities not coming from either type of topology alone. Unfortunately I won't have time to say any more about this either.[3)]

Finally I will all but ignore various speculations that spacetime is actually some kind of topological space *more general* than a manifold. This means neglecting the possibility of boundaries or small holes in spacetime (which, strictly speaking, do not yet transcend the manifold concept), as well as the more radical type of possibility that spacetime is, for example, a set of fractal dimension, or a locally finite topological space.[4)5)]

Large-Scale Topology

Within the limitations I have just set out, the consequences of spacetime topology can be categorized broadly under three headings, depending on the length-scale at which they occur. At the biggest scales—say galactic scales and larger—any deviations from Minkowskian topology are likely to be purely spatial in the sense that spacetime as a whole is the Cartesian product of a time-direction with some unchanging three-manifold. The reason for this limitation is that quantum effects are presumably insignificant at large scales, while classically the Einstein equations forbid topology-change if spacetime is to remain non-singular (and if the energy density of matter is positive, which it is for astrophysical matter).[6] In fact, even without assuming any field equations, one knows that a changing topology entails the existence of closed time-like paths ("time machines"), which are almost certainly of very small size if they exist at all.[7] (Of course there might well have been dynamical topology on cosmologically relevant scales in the very early universe, when what is now big was very small; but that is another matter.)

In speaking about large-scale topology, then, I will forget about time and worry only about the topology of space. It appeared for a while that the field equations (namely the constraints) might rule out all but a few of the conceiveable spatial topologies, but in fact Don Witt has shown there is actually no restriction at all: consistent initial data exists for any closed 3-manifold and therefore presumably for any open one too.[8] (I say presumably only because it is impossible to make a more definite statement without deciding what cosmological boundary conditions one wants to impose in the spatially open situation.)

In view of this infinity of possibilities we might well ask why the actual topology of our universe is so simple. Or is it really as simple as we think? The existing observational evidence is apparently rather weak. We do see a high degree of apparent uniformity, but even if we restrict ourselves to homogeneous and (locally) isotropic metrics, there remains an infinite family of possibilities, each obtained by quotienting one of the three standard spatial slices (spherical, flat, or hyperbolic) by some discrete group of freely acting isometries.

If we lived in such a quotient universe then many of the "distant" galaxies we seem to see might be multiple images of much closer galaxies. If you ask how far out we can look, and be certain that we are not seeing such multiple images, the answer is apparently only about one tenth the radius of the visible universe. It has even been suggested[9)] that the statistical averaging associated with this multiple imaging might account for the high degree of isotropy we observe in the microwave background. If this were correct then we would already possess observational evidence of multiple connectivity without yet recognizing it as such. We could presumably get definitive evidence by identifying unambiguous multiple images, but I don't know how difficult that would be in practice.

Incidentally, this is a good place to correct an impression that is sometimes given that you can tell whether the universe is bounded or not by looking at the sign of the spatial curvature. It is true that in the spherical case the universe must be spatially finite, but in the flat and hyperbolic cases the universe can be either open *or* closed. Indeed there are in some sense "more" closed hyperbolic topologies than there are spherical ones[10)]. (These hyperbolic manifolds also have a curious property I can't resist mentioning: they support spatially closed "cosmologies" that expand from an initial singularity, but are entirely empty and flat, containing neither matter nor gravitational radiation.[11)])

Medium-Scale Topology

On intermediate scales (by which I mean anything from 10^{-16}cm to galactic sizes!) it is again to be expected that topology change is ruled out for the same reasons as before; so again I will worry only about the topology of space. On these scales cosmological boundary conditions become inappropriate, and they must be replaced by ones suited to an isolated system, namely that the metric be asymptotically flat at infinity. Nevertheless, asymptotic flatness *also* permits any spatial topology to occur[8)], so again we can ask why the topology we actually see is so simple.

In fact it probably would not be possible to recognize an isolated piece of spatial topology at all. If we stick to the classical Einstein equations with positive energy density (which seems reasonable at the length scales in question), then one of the singularity theorems (Gannon's theorem) guarantees that any piece of non-trivial topology would form a singularity, and presumably be hidden from view behind an event horizon. In other words it would appear to us as a black hole.

So do we see such black holes? Perhaps the ones thought to inhabit galactic nuclei are hiding non-trivial topology, but we should really be asking why such black holes aren't all around us. In this form the question is recognizeable as an aspect of the old puzzle of why the entropy around us is not much higher than it "might have been". For, as emphasized by R. Penrose, a sprinkling of black holes in our neighborhood would increase the entropy so dramatically that it is hard to understand why such a "thermodynamically likely" state of affairs has not actually been realized[12)]. The relative topological simplicity of our local environment may thus be related to the "relative simplicity" of the initial conditions that gave rise to our region of the universe. (I have used quotation marks because it is not really clear in such arguments what is being compared with what when we speak of relative simplicity. For a local system of bounded energy there probably exists a state of maximum entropy, but when it is a question of cosmological boundary conditions, then any initial state we could think up is sure to be infinitely simpler than an infinite number of possible alternatives.)

Another suggested consequence of intermediate scale topology is much more dramatic than the presence of black holes. We can imagine taking advantage of so-called handles (="wormholes"="bridges") in the spatial topology to travel directly to distant galaxies, or even into the past (which of course tends to follow from the former possibility). For this to really work we would have to create negative energy densities over macroscopically large regions of spacetime, which means that even topology change might then become possible. The proposal of reference 13 uses the

Casimir effect to try to do this, but their arrangement is strongly reminiscent of the "self-accelerating box" of reference 14. As with that paradox, it is not clear whether we are dealing with some contradiction inside quantum field theory, or whether some general principle prevents fabrication of the kind of material needed to construct the walls of the "Casimir chamber", or whether, lastly, time machines and unlimited energy sources are actually physically self-consistent in some subtle way.

Small-Scale Topology

Most of the recent work on spacetime topology relates to scales at, or only somewhat larger than, the so-called Planck length of about 10^{-32}cm. On these truly tiny length-scales it is hard to believe that spacetime could remain topologically trivial, for the following reason. Dimensional analysis tells us that any structure of Planckian size will have a gravitational Action of order h-bar, or in other words of order one in natural units (although structures on much smaller scales might again have much larger Actions). But an Action of order unity leads to very little suppression of the corresponding quantum process, whence Planck-scale modifications of the spacetime topology ought to occur copiously, among them relatively long-lasting ones that could function as particles, and evanescent or "virtual" ones that would pervade the quantum vacuum.

In particular—and in contrast to the situation at "large" and "medium" scales—a dynamical, changing topology is now the rule, since effective negative energies would be easy to create, either by the virtual violations of the field equations just invoked, or due to a locally negative energy density of "on shell" gravitons, gluons, etc. (Although the total energy of a quantum field must be positive, it is common for the expectation value of the energy density to be negative over some region.)

[Of course this reasoning would be much less convincing if we ceased to grant that spacetime is still a manifold at Planck scales. In reality it seems to me at least as plausible that some discrete structure takes over near the Planck length. Even if that is true, however, we are still OK as long as the notion of topology continues to apply down to distances small enough that the above heuristic arguments can come into play. This is not an unreasonable assumption, I think, but on the other hand it could well be false also.]

In the remainder of this lecture, I want to say something about topology as particles, about some possible effects of virtual topology, and finally about the theoretical description of topology-change as it has developed so far. Before taking up these topics, however, I want to list for you some other topics that I am going to have to omit from this discussion of small-scale topology.

First on this list is Kaluza-Klein theory, which not only concerns topology in the sense that it posits a different dimension for spacetime, but also has topological excitations whose existence stems from a novel kind of interaction between the "internal" topology and that of the effective four-dimensional spacetime.[15)]

Next, there is string theory, in which two distinct topologies of "spacetime" come into play: that of the so-called 'target manifold' (which is spacetime in the usual sense) and that of the two-dimensional surface representing the string world-sheet itself. Topological questions have proved crucial in both contexts, but it is perhaps in its role as a theory of "two dimensional gravity" that string theory has fostered the work that has the most to offer to the rest of quantum gravity. One might mention in particular the possibility of generalizing manifolds to orbifolds, the role of unitarity in fixing the relative amplitudes of different topologies in the sum-over-histories, the role of the "large diffeomorphisms" (modular group) as analogs of the large gauge transformations in non-abelian gauge theories, and the study of topology change (the string vertex) in a setting where exact results are much easier to come by than in ordinary General Relativity.[16)]

Going up one dimension, there is (2+1)-dimensional gravity, which recently has been the subject of a fair amount of work, and where the absence of gravitons leads to a much more intimate and perspicuous coupling of the metric to the topology than one finds in four dimensions.[17)]

And finally there are proposed theories that continue to attribute a topological structure to spacetime, but transcend the manifold framework in a more or less essential way. Among such suggestions are the ideas of Finkelstein on "quantum topology"[18)], the suggestion of Hartle to use "unruly" simplicial complexes in Regge calculus[19)], and my own suggestion to use finite topological spaces as models for bounded portions of spacetime[20)] (not to be confused with the suggestion I made this morning that the so-called "causal set" is the true discrete structure.)

Topology as particles

Very early in the history of Relativity it was suggested that some of the then-known elementary particles might represent singularities in the metric, or from a topological point of view "holes in spacetime " resulting from the removal of those points where the fields ceased to be defined. Without going so far as mutilate the manifold in this way, however, it is still possible to introduce topological structures which are "quasi-localized" and therefore capable of behaving like particles. Following a suggestion of Finkelstein and Misner[21)], I will call such particles 'topological geons', or just 'geons' for short.

The relevant topology here is again purely spatial (not spacetime) and is real as opposed to virtual, except insofar as the geons in question are themselves virtual. Also, it seems reasonable to require on physical grounds that a geon be "quasi-localized" in the sense that you can isolate it from the remainder of any space-like hypersurface it inhabits by surrounding it with a single embedded two-sphere. Within this framework one knows that there exist (in 3+1 dimensions) an infinite number of "elementary" geons of which every other possible geon is uniquely a combination. Such elementary manifolds are called 'prime', and it is a theorem that any closed 3-manifold is uniquely (modulo a small ambiguity involving non-orientable handles) a so-called connected sum of primes.

[Notice, by the way, that the condition of quasi-localization eliminates J. Wheeler's "single wormhole mouth" as a geon (but not necessarily Markov's "portal to a hidden universe" [22]) because a sphere surrounding only one of the "mouths" will not isolate the wormhole from the rest of the manifold.]

It is known that geons can be spinorial (carrying half-integer quantum spin), fermionic, and possibly chiral as well[23] [24]. (Moreover, in a Kaluza-Klein realization of gauge symmetry, higher dimensional geons can carry in addition the quantum numbers of any fundamental charged fermion of the theory![25]) Thus, geons possess an unexpected kinematical flexibility that recommends them in certain ways as candidate elementary particles (perhaps so-called preons??[26]). As an illustration of this kind of "flexibility" let me take a minute to explain briefly a recent example in which the two identical geons present are neither bosons nor fermions, but exhibit rather a certain novel combination of those two statistical behaviors[27].

The *exchange* of two isomorphic topological geons can be represented mathematically by a suitable diffeomorphism of the 3-manifold, ${}^3M$ in which they both reside. In addition to diffeomorphisms of this sort there are also ones resulting from a process in which one geon "passes through" another geon by moving along a loop that links the topology of the latter in a non-trivial way. Diffeomorphisms of this second sort are called 'slides', and, together with another class of diffeomorphisms representing "internal symmetries" of the geons, they and the exchanges generate the entire diffeomorphism group of M.

Now each irreducible unitary representation of the diffeomorphism group of M, or more precisely of the group $G = \pi_0(\mathrm{Diff}^\infty(M))$ of (isotopy) equivalence classes of diffeomorphisms of M that become trivial at infinity, induces a possible superselection sector of quantum gravity on the spacetime manifold $M \times \mathbf{R}$. In each such quantum sector there is a close (but not entirely straightforward!) relation between the eigenvalues of the operator representing an exchange diffeomorphism and the statistics of the corresponding geons. But since slides and exchanges do not commute, a slide will mix the exchange-eigenvalues ± 1 unless it is trivially represented itself.

As I mentioned before, this effect has been confirmed explicitly in the case where $M = RP^3 \# RP^3 \# R^3$, or in other words when precisely two identical geons are present, each carrying the topology of projective 3-space, RP^3 (which itself is just the 3-sphere modulo antipodal identification). For this manifold the group G is generated by the exchange E together with a slide S of one of the RP^3's through the other one. The only independent relations are $E^2 = 1$ and $S^2 = 1$, so G has the presentation in terms of generators and relations, $G = < E, S : EE = SS = 1 >$. It turns out that the irreducible representations of G are all either of dimension one or two. In the latter case E is not diagonal, but has both of the eigenvalues ± 1. Hence fermi and bose statistics occur together in the same representation, and they will be mixed dynamically by the action of the quantum Hamiltonian. (Notice that this mixing has nothing to do with so-called parastatistics, which can only come into play when three or more particles are present.)

At first hearing this might sound exotic, but certainly it is not in itself unusual for a pair of identical bosons to turn into a pair of identical fermions via some interaction (say via mutual annihilation followed by creation of a pair of self-conjugate fermions). What does seem peculiar here is the role of the slide S in mediating the conversion of bosons to fermions and vice-versa. My guess is that the mathematics is telling us that a change of statistics can occur when the geons collide with sufficient force that there is an appreciable tunnelling amplitude for one of them to "pass through the other", thereby physically implementing the slide-diffeomorphism S.

Be that as it may, there is another aspect of the statistics-mixing in this particluar case that is not only novel but disconcertingly so. If the RP^3 geons are to be fermions at certain times (and in some of the quantum sectors they are fermions *all* the time), then we would expect them to be able to carry half-integer angular momentum as well. However a geon can do that in one of our superselection sectors only if its associated group of "internal symmetry diffeomorphisms" includes a non-trivial element implementing spatial rotation through 2π. But the manifold RP^3 is so simple that it has no internal symmetries at all, and in particular no non-trivial 2π-rotation diffeomorphism. Hence the corresponding geon will be tensorial (with integer spin) no matter what irreducible representation of G we choose. It can thus be a fermion only at the cost of violating the normal connection between spin and statistics.

Personally I would not conclude from this that we should expect a breakdown of the familiar spin-statistics correlation in quantum gravity. Rather I believe that this counter-intuitive result reflects an incompleteness in the physical assumptions underlying the above analysis. By supposing that the spacetime was simply a product ${}^3M \times \mathbf{R}$, we automatically excluded the possibility of topology change, and therefore of the appearance or disappearance of geons. In other words we have not yet "second quantized" the geons, even though the metric itself is a fully fledged quantum field. Since any theory without the possiblility of pair creation and virtual particles is likely to be physically inconsistent, I think we can conclude that topology change is an inescapable part of the framework of quantum gravity.

Possible consequences of virtual topology

Once we admit the possiblity of a dynamical topology, we must ask what physical effects will be associated with small-scale virtual fluctuations in the connectivity of the spacetime manifold. Among the most intriguing consequences that people have suggested are a gradual loss of coherence of the quantum state-vector, and the automatic forcing of the cosmological constant to zero.

Both these effects are related to the so-called " baby universes", which conceivably branch off from (or join onto) our own universe in certain situations, such as the final moments of existence of an evaporating black hole. If the parent black hole (or is it a midwife?) were of astrophysical size, the "baby"could be relatively large; but one can also imagine that microscopic "virtual branchings" participate in the ceaseless activity underlying the vacuum, just as virtual pair creations and annihilations of ordinary particles do. By extension, one can also imagine vacuum processes in which tiny "tubular" connections are continually forming and breaking between our own (component of the) universe and other universes that may be present "alongside us" (or between pairs of far-apart locations within our own universe.)

Now, any information falling into a black hole is at least temporarily lost to the external universe. If some sort of topological branching were to render this loss permanent, then for all practical purposes, those of us left behind would be forced to "trace out" the degrees of freedom of the child universe, and therefore to replace any unitarily evolving quantum state-vector by a non-unitarily evolving statistical state (or "density matrix"). It seems plausible, therefore, that, even in the absence of real black holes, there is an effective continual loss of quantum coherence associated with virtual topological branching processes.[28)]

If true, this would require some change to the rules of effective quantum field theories such as QCD. Whether such a change is technically feasible—or even whether it undermines quantum mechanics entirely—has been a matter of controversy, but recently Coleman has claimed that it is not really necessary after all, in a certain simplified, Euclidean-signature implementation of the baby universe idea[29)]. My personal view is that a change from state-vector evolution to density-matrix evolution would not be a big conceptual jump in any case, although the technical impediments in its way might be trying to tell us something about the need for a more significant reworking of the foundations of quantum theory.

Whatever its ultimate validity may be, the reasoning leading from topological branching to the expectation of quantum incoherence is relatively straightforward, as argumentation within quantum gravity goes. In contrast, the conclusion that virtual topology can give us a small cosmological constant rests on more tentative foundations, including in particular the basic assumption that quantum gravity is correctly described by a Euclidean-signature path integral.

The first step is to argue that the "gas of baby universes surrounding our own universe" will induce effective vertices (in the sense of quantum field theory) whenever its constituents join onto, or split off from, the main universe(s). For example, an up quark might fall into a virtual black hole, which then branched off from us, taking up life as an independent baby universe. This would induce a baryon non-conserving term in the effective lagrangian, or renormalize the coefficient of any such term that might be present already. In a similar way, every other coefficient ("coupling constant") in the Lagrangian will be renormalized by the baby universes, including the spacetime-volume term, whose coefficient is, of course, the cosmological constant Λ.

Once we have made Λ variable in this way, we argue that the path integral will pick out that value (as conditioned by the state of the baby universe gas) that is most dynamically viable. Adopting a "semi-classical approximation", we are instructed to look for the *solution* of the Euclidean Einstein equation that gives the largest combined Action, and therefore the greatest contribution to the path integral. This extremum turns out to be a sphere whose radius varies inversely with the value of Λ, and the larger the sphere, the larger the resulting probability. But to get a big sphere we need a small (and positive) Λ. Hence the dynamics makes it infinitely probable that the baby universes are in a state that induces an infinitely small Λ so that the corresponding extremum of the Action can be infinitely large. Therefore, it is claimed, the cosmological constant is (infinitely close to) zero.[30)]

An even more far reaching extension of this argument has been put forward by Coleman, who argues that not only Λ, but also *every other coefficient in the Lagrangian* is fixed by this same mechanism, at least in principle! In apparent contradiction to this conclusion, Giddings and Strominger argue that baby universes will cause the coupling constants to become operators subject to uncertainty relations, and therefore incapable of having (simultaneously) any precise values at all.[31]

Before leaving the question of how virtual topology might influence the effective quantum field theory of "low energy" physics, I want to mention a rather different mechanism which could dramatically modify certain coupling constants of the "coarse-grained" Lagrangian. In this case the relevant process is one in which an elementary fermion reverses its handedness by tunnelling through a virtual non-orientable geon. Obviously this process resembles the ones just discussed in connection with "baby universes", but here the fermion re-emerges from the virtual geon back into our own universe, rather than disappearing into a disconnected region of space.

Now the effect of such an orientation-reversal would be described in the effective field theory by a chirality-reversing term in the Action, the simplest example of which is the mass term in the Dirac Lagrangian. Thus one would expect the virtual tunnelling to manifest itself as an increase in the effective mass of the fermion.

Originally, Zel'dovich[32] had suggested that this effect would be suppressed by a small tunneling-probability, but in fact there is no evidence of such suppression in a recent model calculation by Friedman et. al.[33]. Instead they find just what one would expect on dimensional grounds, namely an induced mass of order one in natural units (i.e. of order 10^{19}GeV). Since fermions much lighter than this certainly exist, the mechanism must be neutralized for them. (This tendency to produce inordinately large fermion masses is an example of what is called the 'hierarchy problem' by particle physicists.)

One way out of the difficulty is to assume that the light fermions are really not elementary, but are composites (or better collective excitations) of some more fundamental fields. Another possibility is that non-orientable topology is ruled out by some unknown principle of quantum gravity. And a third way out would be to specify that all spin-structures must be summed over with equal weight in forming the fermionic path-integral; for in such a sum the offending amplitudes would cancel in pairs. (A spin-structure is in effect a rule for resolving the sign ambiguity associated with transport around non-trivial loops in spacetime.) If any one of these possibilities is found to be correct, we will have learned something important about the internal structure of quarks and leptons, or about the basic kinematical or dynamical rules for quantum gravity.

Theory of Topology Change

I hope that what I have said so far has been able to bring out, first of all, the potentially great importance that topological issues have for quantum gravity, and secondly the reasons why we should not expect the topology of spacetime to be static, any more than its metric is, or any of the other spacetime fields are. In order to understand a dynamical topology like this, we will have to broaden our concept of Lorentzian manifold so as to allow the connectivity of spacetime to fluctuate and change; and then we will have to discover the dynamical principles that guide such processes. In other words we need a theory of topology-change.

Of course, we are still far from such a theory, but I think the outlines at least of a possible framework have emerged, and a few suggestive results (even quantitative ones) have also been obtained. In this last part of my lecture, I would like to describe one possible approach to topology change, and to say a little bit more about one or two of the results I just alluded to.

A spacetime in which the topology does *not* change is by definition a manifold homeomorphic to $\mathbf{R}\times\Sigma$, with $\mathbf{R}$ being a timelike dimension, and Σ spacelike. Conversely a portion of spacetime (between initial and final hypersurfaces) within which the topology changes, will not be a simple product manifold. It might not be a manifold at all, ultimately; but it is natural to begin by assuming that it is, at least on scales larger than any fundamental graininess that spacetime itself might possess. With this assumption, the kinematic description of topology-change becomes a branch of *cobordism* theory, this word denoting a manifold whose boundary has been conceptually separated into "initial" and "final" pieces (either of which can be empty, by the way).

Of course a cobordism (say a compact one for definiteness) is only a manifold; it is not yet a physical spacetime until it has acquired a Lorentzian metric. In attempting to add a metric, however, we face in effect, a choice between causality and the equivalence principle, a choice forced upon us by Geroch's theorem that any (time-oriented) Lorentzian metric (on a compact cobordism with spacelike initial and final boundaries) must admit closed timelike curves. If we choose to accept such "causality violations" in order to maintain the Lorentzian signature everywhere, then we get what I will call a 'Lorentzian cobordism'. If, on the other hand, we choose to maintain a globally well-defined causal ordering then we get what I will call a 'causal cobordism'.

[As always there are more radical possibilities which I am neglecting here. One of them is to "Euclideanize" the metric, thereby sacrificing *both* the causal order *and* the Lorentzian character of the metric. Another is to admit non-differentiable or even discontinuous metrics, as suggested by the rough behavior of the histories contributing to the "path integral" in quantum mechanics and quantum field theory. This might invalidate the above-quoted theorem, but I'm not sure it would make the opposition between the "Lorentzian" and "causal" alternatives any less sharp.]

In choosing between the above two alternatives, I personally have been guided by my belief that causality, inasmuch as it is related to the partial order defining a causal set, is a more fundamental notion than is the local light-cone structure that we see in Minkowski space. But there is also supporting evidence coming from five-dimensional Kaluza-Klein theory.

In that theory there exist, as I have already mentioned, topological solitons carrying a charge conjugate to that associated with the symmetry group of the fibers ("magnetic monopoles"). Now everything would lead us to expect that pair-production of such monopoles ought to be a possible process in the quantum version of the theory. And indeed, there does exist a cobordism whose initial boundary has the vacuum topology, and whose final boundary has the topology of a monopole-anti-monopole pair. However, no globally Lorentzian metric exists on this cobordism, or on any other cobordism mediating the same topological transition.[34)]

Finally, there is the practical problem with Lorentzian cobordism that, so far, no-one knows how to set up quantum field theories in a spacetime with a background metric containing closed time-like curves. This makes it difficult to assess the dynamical effect of field fluctuations in topology-changing spacetimes, even though there is evidence (which I will discuss in a moment) that the influence of such fluctuations may well be very great.

The alternative to a Lorentzian cobordism (what I called above a 'causal cobordism') would have to be a spacetime which tolerates some amount of singularity in the metric, but not so much that the past-future relationship induced by a regular metric is lost. It is not clear a priori what the right realization of this requirement is, but there is one particularly simple class of metrics which looks promising, especially because of its relationship with the elementary cobordisms out of which all other cobordisms can be built[35)]. Before describing the type of metric I have in mind, let me say a bit more about the elementary cobordisms themselves.

A priori, you might expect that a general cobordism would be something very complicated, and indeed we know that cobordisms (say the four-dimensional compact ones) are not even classifiable, precisely because 4-manifolds themselves are not classifiable. But from another point of view, cobordisms are really very simple, because each of them can be presented as a succession of elementary modifications, of which there are only five distinct types (or n+1 in n dimensions).

The proof of this result begins by introducing a "time function" into the cobordism, in the hope that surfaces of constant t would evolve continuously from the initial boundary to the final one. As one might expect, the evolution's continuity actually gets interrupted at certain "Morse points", and this discrete collection of spacetime points can be regarded as comprising the "events at which the topology changes." As I mentioned before, Morse points come in only five varieties, each corresponding to one of the five possible signatures of the (non-degenerate) quadratic form $\partial_a\partial_b t$ at an extremum of t. It seems remarkable that the inherently global process of topology change, can in this way be presented as a series of separate events, each localized near a single point of the spacetime in question. In this local simplicity there is hope for understanding topology change at the level of generality it requires.

Now, once you have a Morse function (as t is commonly called), you can go on to concoct a metric g_{ab} for the cobordism, M. I won't review how this can be done, but only remark that the particular construction used in reference 35 yields metrics which are smooth everwhere, zero at the Morse points, and Lorentzian everyplace that they don't vanish. Moreover—and it is in this sense that our cobordism becomes causal—every point of M inherits well-defined past and future "light-cones" from g_{ab}, and these pasts and futures in turn define a coherent partial ordering among the elements of M. (At a Morse point the topological character of the local division into future, past, and present differs from what it is normally; and this difference constitutes a "signature" specifying which of the five possible types the Morse point belongs to.)

If we admit cobordisms of this "causal" sort then we immediately resolve the problem with pair creation that I mentioned in connection with Lorentzian cobordisms. In fact, any pair creation in any dimension can be accomplished via some cobordism (a canonical one, no less); and since every cobordism admits a Morse function, there can never be an obstruction to finding a metric of the desired type, once the cobordism is given. (Incidentally there is no obstruction for Lorentzian cobordism either, as long as we stick to four-dimensions, nor can there be an obstruction for Euclideanized gravity in any dimension.)

We still have to worry, however, about the singular curvature we have introduced at the Morse points, which could cause trouble in at least two ways. First it might give an infinite contribution to the classical gravitational Action (the amplitude at "tree level"). Second it might induce an infinity indirectly by exciting divergent graviton fluctuations (" higher-loop terms"), or fluctuations in other matter fields.

Now as you approach a Morse point P, the metric g_{ab} takes on a canonical form depending only on the Morse-signature of t at P. Thus, any trouble stemming from the infinite curvature (or any other effects arising from the non-Lorentzian charcter of g_{ab} at P) ought to show up if we examine conveniently chosen, local metrics of each of the five canonical types. In fact, the only completed work that I know of refers to (1+1)-dimensions, where there are not five but three distinct elementary cobordisms, or really only two if you forget about the direction of time.

Let us take first the classical Action, whose value depends of course on what terms we choose to include in the Lagrangian. If we use simply the scalar-curvature Action, then it is easy to see by power-counting that the Lagrangian-density will have a non-integrable singularity only in dimension $1+1$. In all other dimensions, the singularity is integrable, and there are then no δ-function contributions to worry about either, presumably.

In 1+1 dimensions, the density $R(-g)^{1/2}$ has a singularity of δ-function strength, but it is not really well-defined as a distribution at all. Nonetheless, it is possible to integrate it plausibly using Regge calculus (or alternatively a complex regularization), and the answer turns out to be finite all right, but complex! Interestingly enough, the sign of its imaginary part differs for the two types of Morse point. For the $(+-)$ type, which corresponds to the "crotch point" of a "trousers" cobordism, the sign is such as to induce *damping* in the functional integral (assuming the sign of the "gravitational constant" is positive). For the $(++)$ type, which corresponds to the point of origination of a two-dimensional "big bang" (this can be called a "yarmulke cobordism", to preserve the sartorial imagery), the sign is such as to *enhance* the contribution to the functional integral, other things being equal. In the $(--)$ case we again get enhancement, since it is just the time-reverse of $(++)$.

The question of fluctuations is naturally more complicated than that of the classical Action alone. To address it in full quantum gravity would be very difficult, but perhaps the behavior of a free massless scalar field ϕ in a background metric of one of our canonical types can serve as a good guide to what one can expect in general (in linear order). Even in this setting of "quantum fields in curved spacetime" the quantization rules become ambiguous in the presence of a Morse point, but different approaches seem to lead to equivalent results, at least qualitatively.

The first calculation of this kind was performed by Anderson and DeWitt[36)], using a method that appears to apply only to the case of Morse signature $(+-)$. By evaluating the Bogoliubov transformation between initial and final vacuum states, these authors computed the net particle creation in a trousers spacetime with a metric flat everywhere except at the Morse point. Choosing a particular "shadow" rule to propagate solutions of the wave equation past this exceptional point, they found an infinite "flash of light" emerging from the crotch. The calculation was improved by Dray and Manogue, who criticized the original choice of propagation rule, but the class of rules that they proposed also led to an infinite burst of energy propagating out along the future light cone of the Morse point[37)].

A somewhat different approach to the same problem tries to avoid specifying an explicit propagation rule for classical solutions, by working directly with the quantum "partition function" $Z = Z(g)$, rather than with a quantum operator $\Phi(x)$ expanded in modes of the classical field[38)]. In this approach it is natural to imagine that $Z(g)$ arises by beginning with the full functional integral for quantum gravity, but then doing only the ϕ-integration, reserving the integral over metrics to a later stage of evaluation. And instead of worrying about a divergence in the expectation value of the stress-energy operator T^{ab}, we worry about a divergence in the response of $\log Z$ to a local variation δg_{ab} of the metric. (In ordinary quantum field theory this response would be identified with an "out-in" *matrix element* of T^{ab}, as opposed to an in-vacuum *expectation value*, which is what the first approach computes.)

Now formally, $\langle \mathcal{T}^{ab} \rangle := \delta \log Z / \delta g_{ab}$, can be written as the coincidence limit of the double gradient of an appropriate Green's function G for the wave-operator. Its value is infinite, of course, but one can renormalize it by subtracting off its flat-space value, and then it depends only on the choice of boundary conditions for G. The precise nature of these conditions should not be too important, except at the Morse point, where we specify that G be finite. (I like to interpret this condition as saying that the Morse point is physically present despite the fact that the metric vanishes there: it has not been "removed from spacetime".) Using in addition Feynman-like boundary conditions in the distant future and past (or whichever of these regions exists in the given case), we can obtain values for $\langle \mathcal{T}^{ab} \rangle$ in both the trousers and yarmulke cobordisms.

For the trousers-type spacetime (Morse signature $(+-)$) the simplest metric g_{ab} constructed in the manner described above turns out to be flat everywhere (except, of course, at the Morse-point, where it vanishes). The renormalized $\langle \mathcal{T}^{ab} \rangle$ blows up as one approaches the light cone of the Morse-point, and now the blow-up occurs in both future and past, corresponding to the time-symmetry of Feynman boundary conditions. The precise nature of the singularity on the light-cone itself remains to be determined, but the integrated "energy" diverges in any case.

For yarmulke spacetimes (Morse signature $(++)$ or $(--)$) the metric again turns out to be flat, making the yarmulke look more like a dunce cap: a geometrical cone. This time, however, the renormalized "stress-energy" $\langle \mathcal{T}^{ab} \rangle$ is everywhere *finite*, except at the Morse-point itself, where it diverges like $1/t^2$. The total energy is then also finite, when evaluated on any hypersurface to the future of the initial singularity (respectively, past of the final singularity).

How should one interpret these preliminary results concerning topology-change via causal cobordism? Both the behavior of the classical Action and of the "one-loop matter terms" indicate that an elementary cobordism of type $(+-)$ is likely to be suppressed relative to the trivial cobordism, and also relative to one of type $(++)$ or $(--)$, insofar as any definite meaning can be attached to this second comparison. DeWitt has even suggested that the infinite energy-production in the trousers means that no such topological transition can occur at all; but such a conclusion would be unwarranted in my view.

In the first place there is the technical objection that the partition function of full quantum gravity can still be finite, even if a certain order of doing the functional integral leads to an apparent divergence in the one-loop terms.

However, I think that a more important consideration here is the potentially misleading character of the two-dimensional problem as a model for the higher dimensional cases. Not only is gravity a degenerate theory below dimension four, but so in a sense is the theory of cobordism itself. Indeed there is reason to believe[35)] that the divergent fluctuations found in the trousers are characteristic of Morse signatures with precisely a single plus sign or a single minus sign. Certainly this hypothesis is consistent with the above results for two-dimensions, but its first real test would come in four-dimensions with the elementary cobordisms of type $(++--)$. Fortunately both of the two methods just described ought to be adaptable to that case.

If indeed certain types of Morse-point were to turn out to be disfavored, or even forbidden, that would be an important discovery about the dyanmics of topology-change. But what if *every* type of elementary cobordism turned out to be infinitely damped in the functional integral (with the possible exception, as indicated by our two-dimensional results, of the all-plus or all-minus ones)? How would this square with my earlier argument that quantum gravity without dynamical topology is physically inconsistent? Personally, I would conclude in that event that spacetime must be discrete in order to cutoff the high frequency fluctuations responsible for the divergences. Thus, one of the theoretical consequences of the existence of topology-change would be the breakdown of spacetime continuity on sufficiently small scales.

Actually, a large (but finite) burst of energy associated with a change of topology might even be a good thing. Recently Kandrup and Mazur[39)] studied particle creation during cobordisms, using an instanton approximation, and were led to suggest that the cosmological big bang was the result of quantum tunnelling from a topologically complicated state to a relatively simple one. According to this suggestion, the radiation "emitted" by the Morse-points involved in the tunneling event, would be responsible for the initial heating of the universe, and therefore might still be present today as the 3-degree microwave background. It would be hard to imagine a more dramatic consequence of topology-change than that!

This research was partly supported by NSF grant PHY 8700651.

References

1. Friedman, J.L., "Spacetime Topology and Quantum Gravity", in *Proceedings of the Osgood Hill Conference on Conceptual Problems in Quantum Gravity*, edited by A. Ashtekar and J. Stachel, (Boston: Birkhäuser, to appear).
2. Freedman, M., " There is no room to spare in four dimensional space", *AMS Notices*, **31**, 3 (1984).
3. Srivastava, A.M., *Classical Topology and Quantization*, Ph.D. thesis (Syracuse University, 1989);

Balachandran, A.P., Jo, S. and Srivastava, A.M., "What Becomes of the Vacuum Angle in Generally Covariant Gauge Theories and the Vielbein Approach to Gravity?", *Int. J. Mod. Phys.*, **A4**, 2995 (1989); a related example is:
Sorkin, R.D., "On the Relation Between Charge and Topology", *J. Phys.*, **A10**, 717 (1977) and **A11**, 795 (1978).

4. Sorkin, R.D., "A Finitary Substitute for Continuous Topology?", Institute for Advanced Study preprint IASSNS-HEP-87/39 (1987); also:
"Posets as Lattice Topologies", in *General Relativity and Gravitation* (Part I of the proceedings of the GR10 conference, held Padova) vol. I, 635, edited by B. Bertotti, F. de Felice and A. Pascolini (Consiglio Nazionale delle Ricerche, Roma 1983).
5. Crane, L. and Smolin, L., "Renormalizability of General Relativity on a Background of Space-time Foam", *Nuc. Phys.*, **B267**, 714 (1986).
6. Tipler, F., *Ann. Phys.*, **108**, 1 (1977);
Borde, A., "Topology Change in General Relativity", preprint.
7. Geroch, R., "Topology in General Relativity", *J. Math. Phys.* , **8**, 782 (1967).
8. Witt, D.M.,"Vacuum space-times that admit no maximal slice", *Phys. Rev. Letts.*, **57**(12), 1386 (1986); also:
"Topological obstructions to maximal slices", to appear in *J. Math. Phys.*; and:
"New asymptotically flat spacetimes", preprint.
9. Sato, H., in Proceedings of the Third Marcel Grossmann Meeting on General Relativity, edited by Hu Ning (Science Press, Beijing and North Holland Press, Amsterdam, 1983).
10. Thurston, W.P. and Weeks, J.R., "The Mathematics of Three-dimensional Manifolds", *Scientific American*, **251** no.1, 108 (July, 1984).
11. Crane, L., Yale University preprint.

12. Penrose, R. "Singularities and Time-asymmetry", in *General Relativity, an Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, 1979).

13. Morris, M.S., Thorne, K.S. and Yurtsever, U., "Wormholes, time machines, and the weak energy condition", *Phys.Rev. Lett.*, **61**, 1446 (1988).

14. Unruh, W.G. and Wald, R.M., "Acceleration radiation and the generalized second law of thermodynamics", *Phys. Rev.* **D25**, 942 (1982).

15. Sorkin, R.D., "A Kaluza-Klein Monopole", *Phys. Rev. Lett.*, **51**, 87 (1983); and **54**, 86 (1985);

    Witten, E., *Nuc. Phys.*, **B195**, 481 (1982);

    Friedman, J.L. and Higuchi, A., "State vectors in higher-dimensional gravity with quantum numbers of quarks and leptons", *Nuc. Phys.*, **B** (in press);

    Carlip, S., "Topological solitons in Kaluza-Klein theories", *Phys. Rev.*, **D33**, 1638 (1986).

16. Green, M., Schwarz, J.H. and Witten, E., *Superstring Theory* (2 volumes), (Cambridge University Press, 1987); see also:

    Krichever, M. and Novikov, S.P., "Virasoro-type algebras, Riemann surfaces and strings in Minkowsky space," *Funktsional'nyi Analiz i Ego Prilozheniya*, **21**, 47 (1987). (Translated with the same title in a 1988 issue of the journal, *Functional Analysis and Applications*, p.294.)

17. Witten, E., "(2+1)-Dimensional Gravity As an Exactly Soluble System", *Nuc. Phys.*, **B311**, 46 (1988).

18. Finkelstein, D. and Rodriguez, E., "The Quantum Pentacle", *Int. J. Theor. Phys.*, **23**, 1065 (1984); also:

    Finkelstein, D., " 'Superconducting' Causal Nets", *Int. J. Theor. Phys.*, **27**, 473 (1988).

19. Hartle, J.B., "Unruly Topologies in Two Dimensional Quantum Gravity", *Classical and Quant. Grav.*, **2**,707 (1985).

20. Sorkin, R.D., "Posets as Lattice Topologies", in *General Relativity and Gravitation* (Part I of the proceedings of the GR10 conference, held Padova) vol. I, 635, edited by B. Bertotti, F. de Felice and A. Pascolini (Consiglio Nazionale delle Ricerche, Roma 1983); and:
"A Finitary Substitute for Continuous Topology?", Institute for Advanced Study (Princeton) preprint IASSNS-HEP-87/39; see also:
Isham, C.J., " An introduction to general topology and quantum topology", Imperial College preprint 88-89/30 (to appear in the proceedings of the Advanced Summer Institute on Physics, Geometry and Topology, held in Banff August 1989).
21. Finkelstein, D. and Misner, C., *Ann. Phys.*, **6**, 230 (1959).
22. Frolov, V.P., Markov, M.A. and Mukhanov, V.F., "Through a Black Hole into a New Universe?", *Phys. Lett.*, **B216**, 272 (1989).
23. Sorkin, R.D., "Classical Topology and Quantum Phases: Quantum Geons", in *Geometrical and Algebraic Aspects of Nonlinear Field Theories* (Proceedings of the conference held Amalfi, Italy, May 1988) edited by S. DeFilippo, M. Marinaro, G. Marmo and G. Vilasi (North-Holland, 1989); also:
"Introduction to Topological Geons", in *Topological Properties and Global Structure of Space-Time*, edited by P.G. Bergmann and V. de Sabbata (Plenum, 1986).
24. Aneziris, Ch., Balachandran, A.P., Bourdeau, M. , Jo S., Ramadas, T.R. and Sorkin, R.D., "Aspects of Spin and Statistics in Generally Covariant Theories", *Int. J. Mod. Phys.*, (in press).
25. see the paper by J.L. Friedman in reference 15.
26. Greenberg, O.W., "The Next Layer of the Onion? Composite Models of Quarks and Leptons", *Orbis Scientiae*, 333 (1982).

27. Aneziris, Ch., Balachandran, A.P., Bourdeau, M. , Jo S., Ramadas, T.R. and Sorkin, R.D., “Statistics and General Relativity”, *Mod. Phys. Lett.*, **A4**, 331 (1989); see also reference 24 and the first paper in reference 23.

28. Hawking, S.W., “Coherence Down the Wormhole”, *Phys. Lett.*, **195B**, 337 (1987);
Strominger, A., “Vacuum Topology and Incoherence in Quantum Gravity”, *Phys. Rev. Lett.* **52**, 1733 (1984); see also in connection with black holes:
Sorkin, R.D., “Toward an Explanation of Entropy Increase in the Presence of Quantum Black Holes”, *Phys. Rev. Lett.*, **56**, 1885 (1986).

29. Coleman, S., “Black Holes as Red Herrings: Topological Fluctuations and the Loss of Quantum Coherence”, *Nuc. Phys.*, **B307**, 854 (1988).

30. Hawking, S.W., “The Cosmological Constant is Probably Zero”, *Phys. Lett.*, **134B**, 403 (1984).

31. a good review of these developments is: Strominger, A., “Baby Unverses”, in Proceedings of the Theoretical Advanced Study Institute, Brown University, held June 1988, edited by A. Jevicki (World Scientific, to appear).

32. Zel’dovich, Ya.B., “Cosmology and the Early Universe”, in *General Relativity, an Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, 1979).

33. Friedman, J.L., Papastamatiou, N., Parker, L. and Zhang H., “Non-orientable foam and an effective Planck mass for point-like fermions”, *Nuc. Phys.*, **B309**, 533 (1988).

34. Sorkin, R.D., “On Topology Change and Monopole Creation”, *Phys. Rev.*, **D33**, 978 (1986); and:
“Non-Time-Orientable Lorentzian Cobordism Allows for Pair Creation”, *Int. J. Theor. Phys.*, **25**, 877-881 (1986).

35. Bombelli, L., Borde, A. and Sorkin, R.D., "Causal Cobordisms", to be published; see also the first paper in reference 34 and unpublished lectures of C.J. Isham.

36. Anderson, A. and DeWitt, B., *Found. Phys.*, **16**, 91 (1986).

37. C.A. Manogue, E. Copeland and T. Dray "The trousers problem revisited", *Pramana–J. Phys*, **30**, 279 (1988).

38. Daughton, A., Gu Zh-Ch., Ó Connor, D.J. and Sorkin, R.D., "Topology-change and Field Fluctuations in Two Dimensions", to be published.

39. Kandrup, H.E. and Mazur, P.O., "Particle Creation and Topology Change in Quantum Cosmology", Florida preprint UFTP-88-10.