Causal continuity in degenerate spacetimes*

A.Borde[a,1], H.F.Dowker[b,2], R.S.Garcia[c,2], R.D.Sorkin[d,3], S.Surya[e,4]

[1] Institute of Cosmology, Department of Physics and Astronomy
Tufts University, Medford, MA 02155, USA

[2] Blackett Laboratory, Imperial College of Science Technology and Medicine,
London SW7 2BZ, United Kingdom

[3] Physics Department, Syracuse University, Syracuse, N.Y., 13244-1130, USA

[4] TIFR, Homi Bhabha Rd, Mumbai 400 005, India.

November 30, 2005

Abstract

A change of spatial topology in a causal, compact spacetime cannot occur when the metric is globally Lorentzian. On any cobordism manifold, however, one can construct from a Morse function f and an auxiliary Riemannian metric $h_{\mu\nu}$ a causal metric $g_{\mu\nu}$ which is Lorentzian almost everywhere except that it degenerates to zero at each critical point of f. We investigate causal structure in the neighbourhood of such a degeneracy, when the auxiliary Riemannian metric is taken to be Cartesian flat in appropriate coordinates. For these geometries, we verify the conjecture that causal discontinuity occurs if and only if the Morse index is 1 or $n-1$.

[a] Email: borde@cosmos2.phy.tufts.edu. Permanent address: Theoretical and Computational Studies Group, Natural Sciences Division, Southampton College of Long Island University, Southampton, NY 11968, USA

*Published in *Classical & Quantum Gravity*, **16:** 3457-3481 (1999). Archived: gr-qc/9901063.

[b] dowker@ic.ac.uk
[c] garciars@ic.ac.uk, Beit Scientific Research Fellow
[d] sorkin@physics.syr.edu
[e] ssurya@tifr.res.in

1 Introduction

Spatial topology change is incompatible with having both a non-degenerate Lorentzian metric and a causal partial order on all spacetime points [1]. One or other of these conditions must be given up if topology change is to occur. However, even if the causal order is abandoned and closed timelike curves (CTCs) allowed, there are certain topology changes, including physically interesting ones such as the pair production of Kaluza-Klein monopoles, that still cannot occur via a globally Lorentzian spacetime [2, 3]. On the other hand if CTCs are excluded, all possible topology changes are permitted at a kinematical level as long as the metric is allowed to degenerate to zero at finitely many isolated "Morse points" [4, 5].

In the Sum-Over-Histories (SOH) framework for quantum gravity, the transition amplitude between two non-diffeomorphic spacelike hypersurfaces V_0 and V_1 is given as a sum over all interpolating geometries (see for example [2, 6, 7, 8, 9, 10, 11]), possibly with a constraint on the total spacetime volume [12]. These geometries are usually taken to be given by everywhere invertible metrics. As discussed in [13, 14], this prescription can be generalised to include the so-called "Morse geometries" on the interpolating cobordism M, where $\partial M = V_0 \amalg V_1$ ($\amalg$ denoting disjoint union). A Morse metric, g, is defined to be

$$g_{\mu\nu} \equiv (h^{\lambda\sigma}\partial_\lambda f \partial_\sigma f)\, h_{\mu\nu} - \zeta \partial_\mu f \partial_\nu f \tag{1}$$

where ζ is a real number greater than one, $h_{\mu\nu}$ is a Riemannian metric on M with inverse $h^{\mu\nu}$, and f is a Morse function $f : M \to [0,1]$ such that $f^{-1}(0) = V_0$ and $f^{-1}(1) = V_1$. A Morse function $f \in C^\infty(M)$ has critical points (points at which $\partial_\mu f = 0$) which are isolated and nondegenerate, i.e. the Hessian of f is non-singular. This metric is Lorentzian throughout the regular set of the Morse function and vanishes at its critical points $\{p_k\}$. The *index* λ_k of p_k is the number of negative eigenvalues of the Hessian of f at p_k. We define a *Morse geometry*[1] as the pair (M, g), where M is a compact cobordism and g a Morse metric on it. We suggest that these Morse geometries should in fact be included in the SOH for quantum gravity.

Motivated by surgery theory, Borde and Sorkin conjectured that only those spacetimes which contain critical points of index 1 or $n-1$ have causal discontinuities [16, 13]. This conjecture, while on one hand of mathematical interest, in

[1] In this paper, the word "geometry" does not imply quotienting by the action of the diffeomorphism group. Indeed, it is unclear whether the appropriate notion of gauge invariance can be expressed by the action of any group. [15]

fact has potential importance in quantum gravity. It was shown by the authors of [17, 18] that scalar quantum field propagation on the $1 + 1$ trousers is singular and hence it was suggested that such a topology will be suppressed in the SOH. Subsequently, the authors of [15] found that, at the level of the gravitational action-integral, causally discontinuous topology changing processes in $1 + 1$ dimensions are indeed suppressed, while causally continuous ones are enhanced. Furthermore, Sorkin conjectured that singular propagation of quantum fields on such backgrounds is related to the causal discontinuity of the spacetime.

It is the purpose of this paper to investigate the first of the above two conjectures. Unfortunately, the existing tools of "causal analysis" [19, 20, 21] posit the existence of a globally Lorentzian metric, and thus exclude Morse geometries with their isolated degeneracies. However, we may investigate the geometries induced in the regular set of the Morse function. Thus we define a *Morse spacetime* as the globally Lorentzian spacetime that results from excising the critical points from a Morse geometry. We note that a Morse spacetime is partially ordered by the causality relation, since the Morse function is a global time function and precludes CTCs. In the discussion section we suggest a way of extending the causal structure to the full Morse geometry with the degenerate points left in.

A general theorem would run along the following lines: a Morse spacetime (M, g) is causally continuous *iff* none of the excised Morse points $\{p_k\}$ have index 1 or $n - 1$. We do not prove the full theorem in this paper, but examine a specific class of spacetimes defined on a neighbourhood of a single critical point and are able to verify the conjecture in these cases.

In section 2 of this paper, we remind the reader of some standard definitions and properties of the causal structure of Lorentzian spacetimes, in particular the definition of causal continuity.

Section 3 contains preliminary general results on the causal continuity of two important classes of spacetimes. First we show that a Morse geometry with no degeneracies, which necessarily has topology $\Sigma \times [0, 1]$ (where Σ is a closed $n - 1$ manifold), is causally continuous. This follows as a corollary to a general result on the equivalence of various causality conditions in the case of a compact cobordism. Another consequence of this general result is that imposing strong causality on the histories of the SOH in the case of a compact product cobordism is equivalent to restricting the sum to non-degenerate Morse metrics. This gives us additional confidence that the proposal to sum over Morse metrics is a good one. Second, the general sphere creation/destruction elementary cobordism, which we refer to as the "yarmulke" spacetime, is shown to be causally continuous.

In section 4 we briefly describe the Morse neighbourhood spacetimes (N_ϵ, g). These are simple ball-neighbourhoods of the critical points of a Morse function f, with the critical point excised, on which the Morse metric is constructed from f and a Riemannian metric which is flat and Euclidean in the coordinates in which f takes its canonical form.

Section 5 contains a detailed analysis, in two spacetime dimensions, of the causal structure of these neighbourhood spacetimes. Up to time reversal, there are in two dimensions only two types of neighbourhood spacetime (and thus, locally, only two types of causal topology change), the trousers ($\lambda = 1$) and the yarmulke ($\lambda = 0, 2$). The general proof in section 3 shows that the two dimensional yarmulke is causally continuous. We verify in section 5 that the trousers neighbourhood spacetime is causally discontinuous.

Section 6 contains our main results on the relation between causal continuity and Morse index. Using appropriate null geodesics in Morse neighbourhood spacetimes (N_ϵ, g), we define the associated spacetimes (Q_δ, g), where $Q_\delta \subset N_\epsilon$ is still a neighbourhood of the Morse point and has null side boundary. These are the spacetimes where we study causal continuity. We show that the neighbourhood spacetime (Q_δ, g) is causally continuous *iff* its Morse point does not have index 1 or $n-1$. The proof makes extensive use of the causal structure of the two dimensional yarmulke and trousers from the previous section.

We summarise our results in section 7 and comment on further aspects of this work that are currently under investigation.

In this paper, the word "spacetime" will be reserved for a C^∞ time-oriented Lorentzian manifold, possibly with spacelike boundary, and our signature convention for the metric will be $(- + + + \cdots +)$.

2 Causal continuity

Causal continuity of a spacetime means, roughly, that the volume of the causal past and future of any point in the spacetime increases or decreases continuously as the point moves continuously around the spacetime. Hawking and Sachs [21] give six concrete characterisations of causal continuity, three of which are equivalent in any globally Lorentzian, time-orientable spacetime, while the equivalence to the remaining three further requires that the spacetime be distinguishing. A spacetime is called distinguishing if any two distinct points have different chronological pasts and different chronological futures.

A time orientation is defined in a spacetime (M,g) by the choice, if possible, of a nowhere vanishing timelike vector field u. We take a timelike or null vector v to be future pointing if $g(v,u) < 0$ and past pointing if $g(v,u) > 0$. We define a *future directed timelike curve* in M to be a smooth function $\gamma : [0,1] \to M$ whose tangent vector is future pointing timelike at $\gamma(t)$ for each $t \in [0,1]$. We also use the phrase future-directed timelike curve and the symbol γ to denote the image, $\{x \in M : x = \gamma(t), t \in [0,1]\}$, of such a function. Strictly we should call the function a "path", say, and reserve "curve" for the image set, but we will ignore this distinction since no ambiguity arises in what follows. Future directed causal curves are defined similarly, but the future directed tangent vector can be null as well as timelike and the curves are allowed to degenerate to a single point. Past directed curves are similarly defined using past pointing tangent vectors.

We write $x << y$ whenever there is a future directed timelike curve γ with $\gamma(0) = x$ and $\gamma(1) = y$ and $x < y$ whenever there is a future directed causal curve γ with $\gamma(0) = x$ and $\gamma(1) = y$. The chronos relation $I \subset M \times M$ is defined by $I \equiv \{(x,y) : x << y\}$ while the causal relation $J \subset M \times M$ is defined by $J \equiv \{(x,y) : x < y\}$. The chronological future and past of a particular point $x \in M$ are, $I^+(x) \equiv \{y : (x,y) \in I\}$ and $I^-(x) \equiv \{y : (y,x) \in I\}$, respectively. The causal future $J^+(x)$ and the causal past $J^-(x)$ of a point are defined similarly.

It can be shown, using local properties of the lightcone, that the chronological relation is transitive (i.e., $x << y, y << z \Rightarrow x << z$) and that it is open as a subset of $M \times M$. The relation J is transitive and reflexive (i.e., $x < x$) [22]. In simple spacetimes such as Minkowski, $J^{\pm}(x)$ is also closed as a set in M, but in general it is not so. Given a subset U of the spacetime, $I^+(x,U)$ denotes the set of points in U that can be reached from x along future directed timelike curves totally contained in U. Note that $I^+(x,U) \subset I^+(x) \cap U$, but the converse is not true in general. $I^-(x,U)$ is defined similarly. Henceforth the dual or time-reversed definitions and statements are understood unless stated otherwise.

We note that Morse spacetimes are time-oriented and distinguishing. The time orientation is given by the timelike vector field normal to the level surfaces of the Morse function f, $u = -\partial f$. Distinguishability is guaranteed by the fact that f is a global time function: if y had the same future set as x, then y would be in $\overline{I^+(x)}$ and it would have to lie in the same level surface $f^{-1}(a)$; but looking at a convex normal neighbourhood [19] of x, we see that x is the only point in $\overline{I^+(x)} \cap f^{-1}(a)$.

Thus, all six characterisations of causal continuity given in [21] are equivalent for Morse spacetimes. Before we list four of these characterisations, we first give a few more definitions.

The *common past*, $\downarrow U$ (*common future*, $\uparrow U$) of an open set U is the interior of the set of all points connected to each point in U along a past (future) directed timelike curve, i.e.,

$$\begin{aligned} \downarrow U &\equiv Int\left(\{x:\ x << u \quad \forall u \in U\}\right) \\ \uparrow U &\equiv Int\left(\{x:\ x >> u \quad \forall u \in U\}\right). \end{aligned} \tag{2}$$

We state two properties of past and future sets that are used later. If x is a point in the time-oriented spacetime M, then: (i) $J^{\pm}(x) \subset \overline{I^{\pm}(x)}$ (Proposition 3.9[19]) (ii) $I^+(x) \subset \uparrow I^-(x)$ and $I^-(x) \subset \downarrow I^+(x)$ (Proposition 1.1[21]).

Let F be a function which assigns to each event x in M an open set $F(x) \subset M$. Then F is said to be *outer continuous* if for any x and any compact set $K \subset M - \overline{F(x)}$, there exists a neighbourhood U of x with $K \subset M - \overline{F(y)}\ \forall y \in U$.

A time-orientable distinguishing spacetime, (M, g) is said to be *causally continuous* if its interior satisfies any of the equivalent properties:

1. for all events x, we have $I^+(x) = \uparrow I^-(x)$ and $I^-(x) = \downarrow I^+(x)$.
2. the "reflecting" property holds, i.e., for all events x and y, $I^-(x) \subset I^-(y)$ *iff* $I^+(y) \subset I^+(x)$;
3. for all events x and y, $x \in \overline{J^-(y)}$ *iff* $y \in \overline{J^+(x)}$;
4. at all events x, $I^+(x)$ and $I^-(x)$ are outer continuous.

Although the last characterisation might seem to capture better our intuitive understanding of causal continuity, it is the first one that we use widely in this paper.[2]

[2]The reason why condition 4 refers only to "outer continuity" is that "inner continuity" of $I^{\pm}(x)$ is automatic [21]. In [21], a spacetime is implicitly assumed to be a manifold without boundary. To transfer these conditions as simply as possible to our case, we have merely imposed them on the interior of M, although condition 3, for example could immediately be extended to ∂M.

For this reason, we introduce the following point-by-point criterion. A spacetime (M,g) is *causally continuous at point* x if $I^+(x) = \uparrow I^-(x)$ and $I^-(x) = \downarrow I^+(x)$. Thus (M,g) is causally continuous *iff* it is causally continuous at every interior point of M.

We also require the definitions of causality, strong causality, stable causality, causal simplicity and global hyperbolicity. *Causality* holds in a subset S of M if there are no causal loops based at points in S. *Strong causality* holds in a subset S of M if for every point $s \in S$ any neighbourhood U of s contains another neighbourhood V of s that no causal curve intersects more than once. A spacetime (M,g) is said to be *stably causal* if there is a metric g', whose lightcones are strictly broader than those of g and for which the spacetime (M,g') is causal. A spacetime is stably causal if and only if it admits a global time function, i.e., a function whose gradient is everywhere timelike. A spacetime is said to be *causally simple* if $J^+(q)$ and $J^-(q)$ are closed for every point q. This is equivalent to the conditions $J^+(q) = \overline{I^+(q)}$ and $J^-(q) = \overline{I^-(q)}$. A spacetime is said to be *globally hyperbolic* if it contains a spacelike hypersurface which every inextendible causal curve in the spacetime intersects exactly once. There is a standard sequence of implications amongst these causality conditions [21]: global hyperbolicity of a spacetime $\Rightarrow$ causal simplicity $\Rightarrow$ causal continuity $\Rightarrow$ stable causality $\Rightarrow$ strong causality $\Rightarrow$ causality.

Finally, we recall a result which will be useful in our proofs: the compactness of the space of causal curves [19]. Let K be the set of all points in M where strong causality holds and $C \subset K$ be a compact set, then for any closed subsets A, B of C the space $\mathcal{C}_C(A,B)$ of causal curves in C from A to B is compact (cf. [23], proof of Theorem 23). Strictly speaking, this holds only if one uses a weaker definition of causal curve than the one we are using since the limit curve γ to which a sequence of smooth causal curves converges is not in general smooth. The existence of the limit curve nevertheless ensures the existence of some smooth causal curve between the endpoints of the limit curve, and this subtlety is ignorable for our purposes.

3 Causal continuity of non-degenerate and yarmulke spacetimes

By "non-degenerate Morse spacetime" we mean a compact Morse spacetime (M,g), *i.e.* a Morse geometry with no degeneracies. A yarmulke spacetime is one arising from a Morse geometry in an n-dimensional elementary cobordism of index $\lambda =$

0 (n), whose initial (final) boundary is empty and final (initial) boundary is S^{n-1}.

In this section all spacetimes are assumed to be time-orientable and distinguishing, with boundary $\partial M = V_0 \coprod V_1$, where V_0 and V_1 are closed $n-1$ manifolds (possibly empty) and are respectively the initial and final spacelike boundaries of M.

Proposition 1 *For a compact spacetime (M, g), the following properties are equivalent:*

1. *It is globally hyperbolic.*
2. *It is causally simple.*
3. *It is causally continuous.*
4. *It is stably causal.*
5. *It is strongly causal.*

Proof: Since each item implies the next, we only need to prove that the last item implies the first.

So suppose that strong causality holds on (M, g). For any spacetime, $J^+(q) \subset \overline{I^+(q)}$. Let $p \in \overline{I^+(q)}$, and consider a sequence of points $p_k \in I^+(q)$ which converges to p. There must be a sequence of future-directed timelike curves, γ_k from q to p_k. Since (M, g) is strongly causal, the compactness of the space of causal curves, with $C = M$, $A = \{q\}$, $B = \{p_k : k = 1, 2, \ldots\} \cup \{p\}$, implies that there is a causal limit curve γ from q to p. Thus, $p \in J^+(q)$, or $\overline{I^+(q)} \subset J^+(q)$ which implies that $J^+(q) = \overline{I^+(q)}$. $J^-(q) = \overline{I^-(q)}$ is proved similarly. So (M, g) is causally simple.

(M, g) is therefore stably causal and thus admits a global time function, f. Let S be any $f =$ constant surface. Let p lie on some later time surface, and suppose that a past inextendible causal curve from p, say λ, does not intersect S. Then λ is trapped in the region of M between the two constant-time surfaces mentioned above. Since this region is compact [24] λ must accumulate at some point q and therefore it intersects the constant-time surface through q more than once. This is not possible. $\square$

Proposition 1 as stated does not ask that V_0 and V_1 be non-empty or have the same topology but the known causality violations in a spacetime where these properties fail makes the result relevant only to spacetimes which satisfy them. This proposition provides us with a proof of the causal continuity of non-degenerate Morse spacetimes (M, g), since they are compact and possess a global time function:

Corollary 1 *A non-degenerate Morse spacetime (M, g) is causally continuous.*

Proposition 2 *Let (M, g) be a compact spacetime. Then M is a non-degenerate Morse spacetime if and only if it is strongly causal.*

Proof: The Morse function of a non-degenerate Morse spacetime is a global time function and so it is stably causal and hence strongly causal.

Suppose (M, g) is strongly causal. Then it is stably causal as above and has a global time function f. Let $b^2 = -g^{\mu\nu}\partial_\mu f \partial_\nu f$. Define the positive definite metric h by

$$h_{\mu\nu} \equiv g_{\mu\nu} + (1 + 1/b^2)\partial_\mu f \partial_\nu f. \tag{3}$$

That h is positive definite may be checked by using the basis $\{T^\mu, S_i^\mu\}$, where $T^\mu = g^{\mu\nu}\partial_\nu f$ is timelike with respect to g and the S_i^μ are chosen to be spacelike and orthogonal to each other and to T^μ (with respect to g). Then $\{T^\mu, S_i^\mu\}$ forms an orthogonal basis with respect to h, and all the vectors have positive norm. Since $h^{\mu\nu}\partial_\mu f \partial_\nu f = 1$ (where $h^{\mu\nu} := g^{\mu\nu} + (b^{-2} + b^{-4})g^{\mu\mu}g^{\nu\nu}\partial_\mu f \partial_\nu f$ is the inverse of $h_{\mu\nu}$), we see that we may invert the above expression and express g as a Morse metric of the form (1) with $\zeta = 1 + 1/b^2$.

It is possible to choose f such that $f^{-1}(0) = V_0$ and $f^{-1}(1) = V_1$. The function $\hat{f}(p) \equiv Volume(\overline{I^-(p)})$ is a time function on (M, g). Let $T^\mu \equiv g^{\mu\nu}\partial_\nu \hat{f}$. For any point $p \in M$, let $q_p \in V_1$ be the future endpoint of the integral curve of T^μ through p. Define $v(p) \equiv Volume(\overline{I^-(q_p)})$. Then $f(p) \equiv (1/v(p))\hat{f}(p)$ is a time function on M with $f^{-1}(0) = V_0$ and $f^{-1}(1) = V_1$. $\Box$

Proposition 2 shows that the restriction to Morse metrics in a Lorentzian compact spacetime with initial and final spacelike boundaries is a reasonable one, since that restriction is equivalent to strong causality. Observe that, without any compactness assumption, the proof also shows that a non-degenerate metric is Morse iff it possesses a time function, hence iff it is stably causal. It would be useful to have a similar characterisation of Morse metrics in the degenerate case as well.

We now prove causal continuity for yarmulke spacetimes in n dimensions. The Morse function has a single critical point p. Suppose that $f : M \to [0,1]$ is the Morse function and that the boundary of M is a final boundary. Then we must have $f(p) = 0$. Excising p from our manifold, the associated Morse spacetime has topology $S^{n-1} \times (0,1]$.

Lemma 1 *The yarmulke spacetime (M, g), with topology $S^{n-1} \times (0,1]$, is causally continuous.*

Proof: Using arguments similar to those in the proof of proposition 2, we see that since the spacetime is stably causal, and therefore strongly causal, the surfaces of constant f are Cauchy surfaces. (The full spacetime is not compact here, but the region between any two level surfaces of f is.) Thus the spacetime is globally hyperbolic and hence causally continuous. $\square$

4 The neighbourhood of a Morse point

In this section we introduce the Morse metrics which we study in the remainder of the paper. They are defined in a small neighbourhood of a Morse point, where the Morse function takes a simple form. Consider an open ball D_ϵ of radius ϵ in local coordinates centred on the point p. Let $\{x^i, y^j : i = 1, \ldots \lambda, j = 1, \ldots n - \lambda\}$ be these local coordinates with $x^i(p) = y^j(p) = 0$, $\sum_i (x^i)^2 + \sum_j (y^j)^2 < \epsilon^2$ and in which the Morse function f takes the following canonical form (Morse lemma [24]):

$$f = f(p) - \sum_i (x^i)^2 + \sum_j (y^j)^2 \ . \tag{4}$$

The spacetime manifold we are concerned with is $N_\epsilon \equiv D_\epsilon - \{p\}$. Henceforth we consider all set closures, *etc.*, to be taken in the manifold N_ϵ except when explicitly stated otherwise. We frequently refer to p as though it were in the spacetime: for example we refer to sets of curves that are "bounded away from p", the meaning of which should be clear.

We define the polar coordinates, (ρ, Θ, r, Φ), where $\rho^2 = \sum_1^\lambda x_i^2$, $r^2 = \sum_1^{n-\lambda} y_j^2$, and the collective coordinates Θ and Φ stand for the angles θ_i, $i = 1 \cdots \lambda - 1$ and ϕ_j,

$j = 1 \cdots n-\lambda-1$ that parameterise the $(\lambda-1)$-sphere and the $(n-\lambda-1)$-sphere, respectively. When $\lambda = 1$ the sphere S^0 is disconnected and our convention for such cases is to adopt a single discrete "coordinate" θ_0 with two possible values: 0 and π. Then the transformation between x^1 and (ρ, θ_0) is $x^1 = \rho \cos\theta_0$. Similarly, when $\lambda = n-1$, the disconnected sphere will be parameterised by the single discrete angle ϕ_0 and $y^1 = r\cos\phi_0$. We will often use the notation whereby the coordinates of a point q are referred to as (x^i_q, y^j_q) or $(\rho_q, \Theta_q, r_q, \Phi_q)$ except that the discrete angles are written $\theta_0(q)$ or $\phi_0(q)$ to avoid double subscripts.

The particular Morse metrics that we will be studying are those for which the auxiliary Riemannian metric h is the flat Cartesian metric in the coordinates $\{x^i, y^j\}$ introduced above. In that case we find, transforming to polar coordinates, that the metric (1) becomes

$$\begin{aligned} ds^2 \;=\; & 4\{(\rho^2+r^2)(\rho^2 d\Theta^2_{\lambda-1} + r^2 d\Phi^2_{n-\lambda-1}) + (r^2-(\zeta-1)\rho^2)d\rho^2 \\ & +(\rho^2-(\zeta-1)r^2)dr^2 + 2\zeta\rho r d\rho dr\}. \end{aligned} \tag{5}$$

In these coordinates ds^2 is seen to be a warped product metric[3] since it decomposes into a radial and angular part, i.e., $ds^2 = ds^2_A + ds^2_R$, with the radial coordinates (r, ρ) warping the angular part. An important property of such a warped product form, is that the geodesics of the metric ds^2_R are also geodesics of the full metric. We note also, that ds^2_A is never negative, so that for any timelike (causal) curve $\gamma(t) = (\rho(t), \Theta(t), r(t), \Phi(t))$, the related curve $\gamma'(t) = (\rho(t), \Theta_a, r(t), \Phi_b)$ at any fixed angles Θ_a and Φ_b, is also timelike (causal). The Morse function $f(r,\rho) = f(p) - \rho^2 + r^2$ must increase along future directed timelike curves so that the future time direction is, roughly speaking, decreasing ρ and increasing r.

We call (N_ϵ, g), where g is given by (5), a *Morse neighbourhood spacetime* of type $(\lambda, n-\lambda)$.

5 Causal structure for $n = 2$

We study in detail the two elementary neighbourhood spacetimes in $n = 2$, namely the trousers ($\lambda = 1$) and the yarmulke ($\lambda = 0$ or $\lambda = 2$), see figure 1. (Of course $\lambda = 2$ is just the time-reverse of $\lambda = 0$.) Both these spacetimes happen to be

[3] A metric $g(x^a, y^A)$ is a warped product metric if its line-element takes the form $ds^2 = g_{ab}(\vec{x},\vec{y})dx^a dx^b + g_{AB}(\vec{y})dy^A dy^B$.

flat, but we do not use this feature explicitly in our analysis (it does not extend to higher dimensions). The causal structure of these two cases turns out to be crucial in studying the general n-dimensional neighbourhood spacetimes. A study of the causal structure of the $1+1$ trousers for a particular choice of flat metric was first carried out by [25].

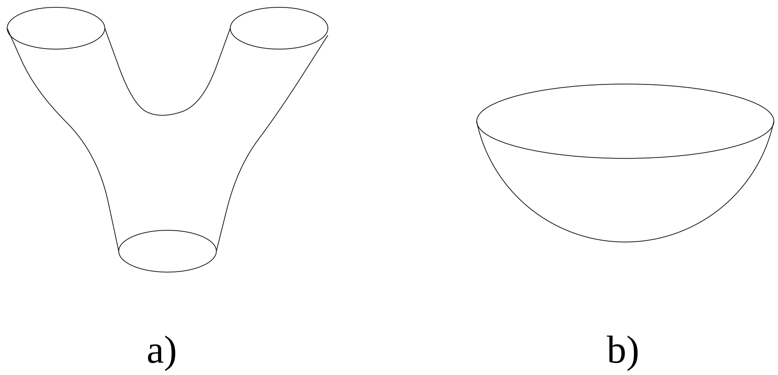

Figure 1: Global topology of the trousers and yarmulke cobordisms. In the trousers N_ϵ is to be regarded as a little region around the saddle point, while in the yarmulke N_ϵ has the same topology as the whole cobordism, which (with p restored) is just a disc.

5.1 Causal structure in the trousers

The neighbourhood Morse metric for the trousers is

$$ds^2 = 4\{(y^2 - (\zeta - 1)x^2)dx^2 + (x^2 - (\zeta - 1)y^2)dy^2 + 2\zeta xydxdy\}. \tag{6}$$

Any radial line with endpoint at the origin is a geodesic so the norm of its tangent vector has the same sign all along. Thus the disc can be partitioned into sectors that are loosely speaking, either future time-like, past time-like or space-like related to the origin (see figure 2). Indeed, by substituting $y = mx$ one obtains:

$$ds^2 = -4((\zeta-1)m^4 - 2(\zeta+1)m^2 + (\zeta-1))x^2dx^2 \quad \begin{cases} > 0 & \text{if} \quad (m_1)^2 < m^2 < (m_2)^2 \\ \\ \leq 0 & \text{otherwise} \end{cases} \tag{7}$$

where the gradients $m_1 = \frac{\sqrt{\zeta-1}}{\sqrt{\zeta-1}} < 1$ and $m_2 = \frac{\sqrt{\zeta+1}}{\sqrt{\zeta-1}} = m_1^{-1} > 1$ mark the transition between the spacelike and timelike character of the radii, with the lines $y = \pm m_1 x$ and $y = \pm m_2 x$ being null and separating the sectors.

We define sets $\mathcal{P}_1, \mathcal{P}_2, \mathcal{F}_1$ and $\mathcal{F}_2$, via

$$\begin{aligned}
\mathcal{P}_1 &\equiv \{(x,y) \in N_\epsilon : |y| < m_1|x| \text{ and } x > 0\} \\
\mathcal{P}_2 &\equiv \{(x,y) \in N_\epsilon : |y| < m_1|x| \text{ and } x < 0\} \\
\mathcal{F}_1 &\equiv \{(x,y) \in N_\epsilon : |y| > m_2|x| \text{ and } y > 0\} \\
\mathcal{F}_2 &\equiv \{(x,y) \in N_\epsilon : |y| > m_2|x| \text{ and } y < 0\} \qquad (8) \\
& \qquad (9)
\end{aligned}$$

Also $\mathcal{P} \equiv \mathcal{P}_1 \cup \mathcal{P}_2$, and $\mathcal{F} \equiv \mathcal{F}_1 \cup \mathcal{F}_2$. We define $\partial\mathcal{P}$ and $\partial\mathcal{F}$:

$$\begin{aligned}
\partial\mathcal{P} &\equiv \{(x,y) \in N_\epsilon : y = m_1 x \text{ or } y = -m_1 x\} \\
\partial\mathcal{F} &\equiv \{(x,y) \in N_\epsilon : y = m_2 x \text{ or } y = -m_2 x\} \qquad (10)
\end{aligned}$$

and $\mathcal{S} \equiv N_\epsilon - (\mathcal{P} \cup \mathcal{F} \cup \partial\mathcal{P} \cup \partial\mathcal{F})$. We see that $\mathcal{F}$ is what we'd expect for the chronological future of the Morse point, p, $\partial\mathcal{F}$ is what we might want to call the future lightcone of p and similarly for the past; $\mathcal{S}$ is the "elsewhere" of p. The status of these sets is discussed further in a later subsection.

To summarise, through all the points in the shaded regions of figure 2 there passes a radial timelike geodesic; $\mathcal{S}$ denotes the points outside the shaded regions through which the radial geodesics are spacelike, while the boundary radial geodesics, with gradient $\pm m_1$ and $\pm m_2$ are null.

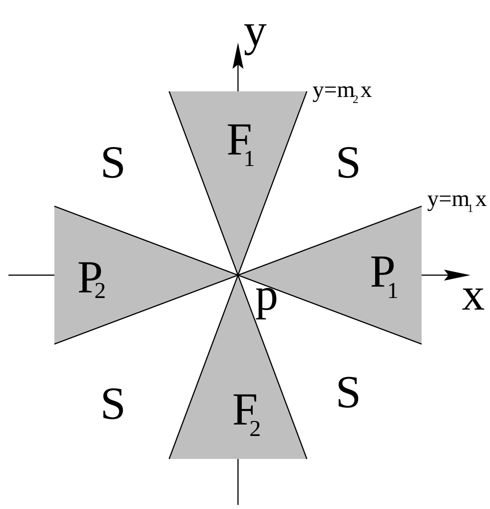

Figure 2: The trousers. Partition of the disc by radial geodesics. The timelike geodesics are in the shaded regions, future outward about the y-axis and future inward about the x-axis. The radial geodesics outside the shaded regions are spacelike.

We return to the general null geodesics which give us the lightcones for an arbitrary point. This is a generalisation to arbitrary ζ of the analysis of [13]. From equation (6) one obtains implicit equations for the null curves:

$$\sqrt{\zeta - 1}(x^2 - y^2) = 2xy + c_+ \tag{11}$$

$$\text{or} \quad \sqrt{\zeta - 1}(x^2 - y^2) = -2xy + c_- \, , \tag{12}$$

where $c_\pm$ are constants. To get a clearer idea of what these curves are, consider a rotated coordinate system (x', y') with the x' axis being the line $y = m_1 x$ at an angle of $\alpha = \tan^{-1} m_1$ with the x axis:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{13}$$

Then (11) takes the form

$$x'y' = \frac{-c_+}{2\sqrt{\zeta}} \tag{14}$$

which shows that they are hyperbolae with $y = m_1 x$ and $y = -m_2 x$ as asymptotes. Rotating to coordinates (x'', y'') by $-\alpha$ instead, (12) becomes

$$x''y'' = \frac{c_-}{2\sqrt{\zeta}} \tag{15}$$

so these are hyperbolae with $y = -m_1 x$ and $y = m_2 x$ as asymptotes.

Through every point q in N_ϵ there passes a curve which satisfies eq. (14) with a particular c_+, we call it v_q^+ and another curve which satisfies eq. (15), we call this v_q^-. As we see shortly, these null geodesics through q suffice to bound its past and future provided q lies in $\mathcal{S}$, but not otherwise. To determine systematically the chronological pasts and futures of all points in the disc, we start by noting that in the rotated coordinate systems,

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{16}$$

the hyperbolae given by $\tilde{x}\tilde{y}$ =constant are timelike if and only if $-\alpha < \psi < \alpha$ or $\frac{\pi}{2} - \alpha < \psi < \frac{\pi}{2} + \alpha$, that is when the $\tilde{x}$ and $\tilde{y}$ axis fall in the interior of the shaded regions in figure 2. We will use segments of such timelike hyperbolae to determine the chronological relation. By symmetry and time-reversal invariance, we need consider only three representative points in the upper right quadrant: (i) $q \in \mathcal{S}$, (ii) $q \in \mathcal{F}_1$ and (iii) $q \in \partial\mathcal{F}_1$.

(i) For $q \in \mathcal{S}$, we claim that $I^+(q)$ is the interior of the horizontally shaded region in figure 3, bounded by the two null hyperbolae through q, υ_q^+ and υ_q^- and that $I^-(q)$ is the interior of the vertically shaded region between the same hyperbolae. Those regions are contained in $I^+(q)$ and $I^-(q)$, because they are swept out by the timelike hyperbolae through q between υ_q^+ and υ_q^-. To see that such regions exhaust all of $I^+(q)$ and $I^-(q)$ is also straightforward. Any future directed timelike curve from q must begin by heading into the horizontally shaded region. If there was one such curve ending at a point s outside the region, it would have to intersect one of the bounding hyperbolae at some point q', but the tangent vector there could not point out of the region and be both timelike and future directed according to the local lightcone at q'. Similarly for $I^-(q)$.

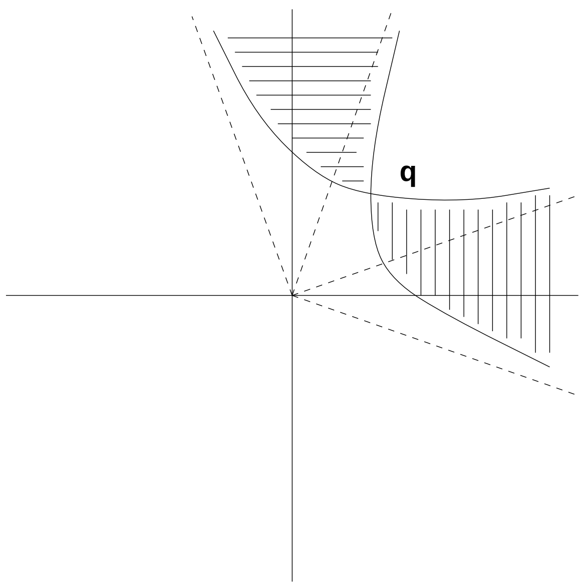

Figure 3: The trousers. $q \in \mathcal{S}$ and $I^+(q)$ ($I^-(q)$) is the horizontally (vertically) shaded region.

(ii) For $q \in \mathcal{F}_1$, $I^+(q)$ is the interior of the horizontally shaded region shown in figure 4 bounded by the two null hyperbolae through q, by the same arguments as in case (i), and we claim that $I^-(q)$ is the interior of the vertically shaded region bounded by the hyperbolae and by lines $y = m_1 x$ and $y = -m_1 x$ with $y < 0$.

To show that this region is indeed contained in $I^-(q)$ we find sequences of timelike hyperbola segments from all points in the region to q. Let us first divide the region in two zones. If $y_q = m_q x_q$, let zone 1 consist of points with $y > -\frac{1}{m_q}x$ and zone 2 of points with $y \leq -\frac{1}{m_q}x$ (see figure 5).

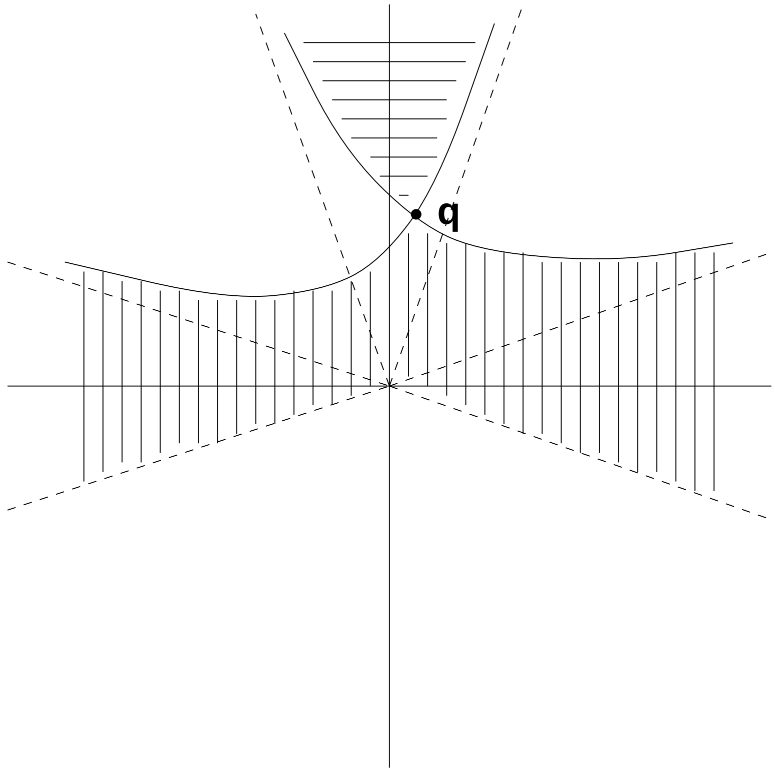

Figure 4: The trousers. $q \in \mathcal{F}$ and $I^+(q)$ ($I^-(q)$) is the horizontally (vertically) shaded region.

A point s in zone 1 can be joined to q by a single arc of timelike hyperbola whose asymptotes are the $\tilde{x}$ and $\tilde{y}$ axes defined by (16) with $-\frac{1}{m_q} < \tan\psi < m_1$, except when $y_s = m_q x_s$ in which case the timelike curve to q is $y = m_q x$.

For a point t in zone 2 two hyperbolic arcs are needed: the first one is part of the hyperbola with asymptotes given by (16), with $-m_1 < \tan\psi < m_t$ where $m_t = y_t/x_t$, and connects t to a point s' in zone 1. Then s' can be connected to q as before.

The argument that these regions comprise all of $I^{\pm}(q)$ is as in (i).

(iii) For $q \in \partial\mathcal{F}_1$, similar arguments show that $I^+(q)$ is the interior of the horizontally shaded region bounded by the null hyperbola $x'y' = x'_q y'_q$ and the null line $y = m_2 x$. $I^-(q)$ is the interior of the vertically shaded region bounded by the hyperbola, $y = m_2 x$ and $y = -m_1 x$ for $x > 0$ (see figure 6). Note that $I^-(q)$ does not contain any point with $x < 0$.

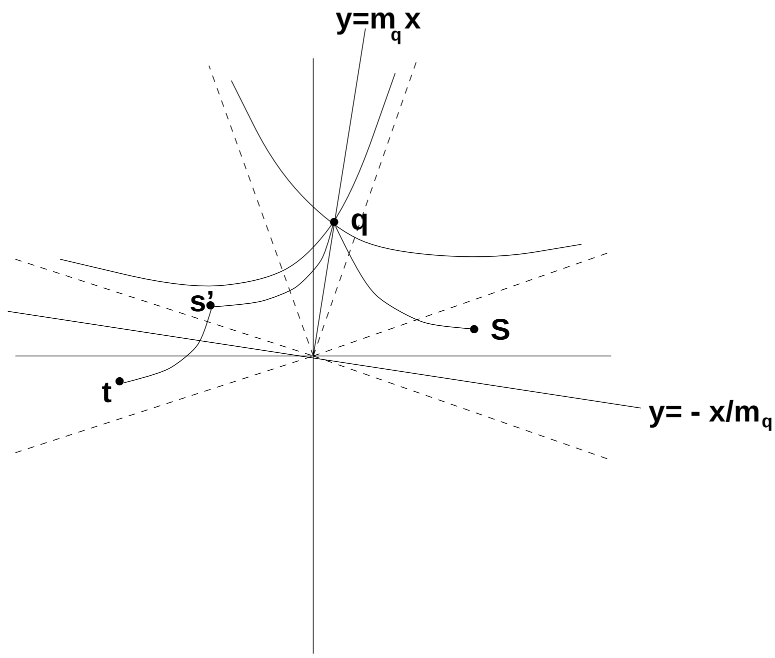

Figure 5: The line $y = -x/m_q$ separates the past of q into zone 1, where $y > -x/m_q$, and zone 2, where $y \leq -x/m_q$. A point $s \in I^-(q)$ in zone 1 can be joined to q with a single arc of timelike hyperbola. For a point $t \in I^-(q)$ in zone 2 we need two such arcs to reach q.

Using $J^\pm(q) \subset \overline{I^\pm(q)}$ the causal pasts and futures of these representative points are easy to find. Since all the points in a null geodesic through q are in its causal past or future, we just need to decide whether the additional straight lines bounding the $I^\pm$ of points type (ii) and (iii) are in $J^\pm$. For definiteness, consider $q \in \mathcal{F}_1$ and a point s on $y = -m_1 x$, $y < 0$. Now q is not in $\overline{I^+(s)}$, so it doesn't belong to $J^+(s)$ either: there is no causal curve from s to q. It follows that the lines that bound $I^-(q)$ in the lower hemiplane are not in $J^-(q)$. Neither is the line $y = -m_1 x$ contained in $J^-(q)$ when $q \in \partial\mathcal{F}_1$. Summarising, for points $q \in S$ we have $J^\pm(q) = \overline{I^\pm(q)}$, otherwise the causal sets $J^\pm(q)$ are not closed.

This completes our analysis of the causal structure around the crotch singularity in the $1+1$ trousers, which we use extensively later. For the moment it allows us to establish the following.

Lemma 2 *The neighbourhood spacetime* (N_ϵ, g) *of type* $(1,1)$ *is causally discontinuous.*

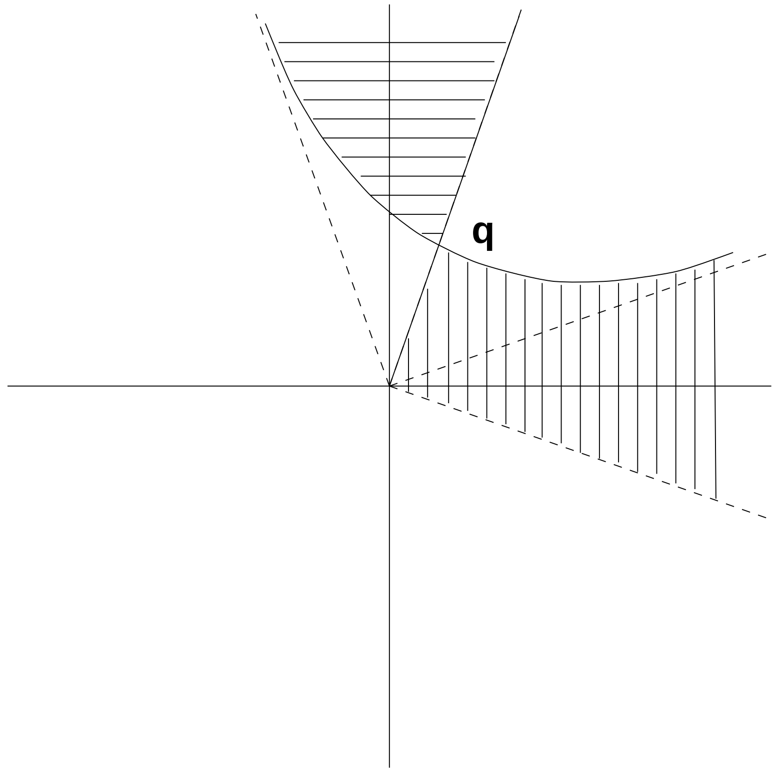

Figure 6: The trousers. $q \in \partial\mathcal{F}$ and $I^+(q)$ $(I^-(q))$ is the horizontally (vertically) shaded region.

Proof: Let $q \in \partial\mathcal{F}_1$ with $x_q > 0$, then $\downarrow I^+(q) \neq I^-(q)$ since any s on the negative x-axis satisfies $s \in \downarrow I^+(q)$ but $s \notin I^-(q)$.$\Box$

In figure 7 we illustrate the failure of causal continuity in each of the remaining three equivalent definitions given in section 2, in the hope to give the reader an intuition for their meaning.

5.2 Causal structure in the yarmulke

We now undertake a similar analysis for the neighbourhood spacetime $(0, 2)$. The Morse metric on the punctured disc is

$$ds^2 = 4(-(\zeta - 1)r^2 dr^2 + r^4 d\phi^2) \ . \tag{17}$$

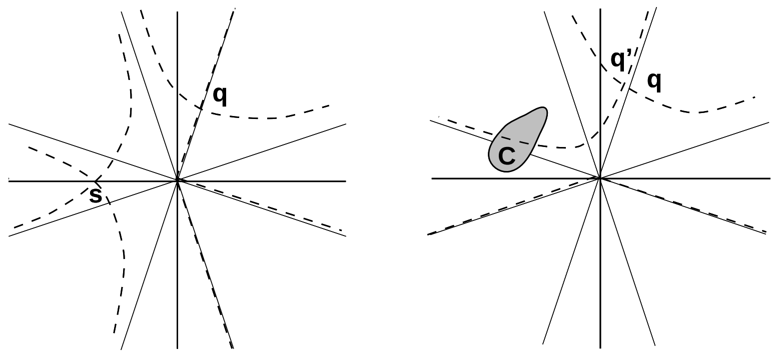

Figure 7: On the left, the futures and pasts of a pair of points, (s, q) are shown. This pair violates the reflecting property, since $I^+(q) \subset I^+(s)$ but $I^-(s) \not\subset I^-(q)$. Also $q \in \overline{J^+(s)}$ but $s \notin \overline{J^-(q)}$. On the right the compact set $C \subset M - \overline{I^-(q)}$ intersects $\overline{I^-(q')}$ for a point $q' \in \mathcal{F}_1$ which can be taken arbitrarily close to q.

Because of the $U(1)$ symmetry of this metric, there is only one class of points. Since $f(r) = r^2$ is the time function, timelike and null tangent vectors are past pointing if their radial component is inwards and future pointing if it is outwards. As before, in two dimensions the geodesic equation is not really needed to find the null geodesics through a point; solving for a tangent vector with vanishing norm suffices. The solutions are the null spiraling curves $\sigma_0^\pm$ given by,

$$r(\phi) = r_0 e^{\frac{\pm\phi}{\sqrt{\zeta - 1}}} \tag{18}$$

Again these spirals can be shown to be consistent with the geodesic equations. For spacetimes with $\zeta \to 1$, the lightcones are squeezed onto the radial direction, while for $\zeta \to \infty$, the lightcones widen to become circles.

The future direction along the spirals σ_q^+ and σ_q^- through a point q corresponds to an increase in the radial coordinate, which, along σ_q^+ is achieved by increasing ϕ and along σ_q^- by decreasing ϕ. The radial straight lines $\phi =$ constant also satisfy the geodesic equations. These geodesics are timelike, since $d\phi = 0$ along them.

Forgetting for the moment that we are confined to a disc of finite radius, the two spirals through a point q converge again at both an earlier q' and at a later q'', beyond which these null geodesics no longer bound the past or future of q (see figure 8). Let L_q be the interior of the little heart-shaped region bounded by the two past directed null spirals from q to q' and B_q the big heart-shaped region bounded by the future directed null spirals from q to q''. Then we claim (i) $I^-(q) = L_q$ and (ii) $I^+(q) = N_\epsilon - B_q$.

By the symmetry of this metric, all points are equivalent so that our task reduces to considering the representative point $q = (r_q, 0)$. We only give the argument for (i) since (ii) is similar[4]. Let $s = (r_s, \phi_s) \in L_q$. By the symmetry of L_q, we can assume without loss of generality that $0 \leq \phi_s \leq \pi$. Thus, $r_s < r_q e^{\frac{-\phi_s}{\mathrm{Z}}}$, where $\mathrm{Z} = \sqrt{\zeta - 1}$. In particular for points s such that $\phi_s = 0$, the radius $\phi = 0$ is a timelike curve from s to q and therefore they belong to $I^-(q)$. If $\phi_s \neq 0$, consider the curve γ from s to q given by

$$r(\phi) = u(\phi) e^{\frac{-\phi}{\mathrm{Z}}} \tag{19}$$

where the function $u(\phi)$,

$$u(\phi) = r_q \frac{\phi_s - \phi}{\phi_s} + r_s e^{\frac{\phi_s}{\mathrm{Z}}} \frac{\phi}{\phi_s} \tag{20}$$

decreases smoothly from r_q to $r_s e^{\frac{\phi_s}{\mathrm{Z}}}$. Then

$$\begin{aligned} \frac{dr}{d\phi} &= -\left(r_q - \frac{\phi}{\phi_s}(r_q - r_s e^{\frac{\phi_s}{\mathrm{Z}}}) \right) \frac{e^{\frac{-\phi}{\mathrm{Z}}}}{\mathrm{Z}} - \frac{1}{\phi_s}(r_q - r_s e^{\frac{\phi_s}{\mathrm{Z}}}) e^{\frac{-\phi}{\mathrm{Z}}} \\ &= -\frac{r}{\mathrm{Z}} - \frac{r_q - r_s e^{\frac{\phi_s}{\mathrm{Z}}}}{\phi_s} e^{\frac{-\phi}{\mathrm{Z}}} \end{aligned} \tag{21}$$

so that, along γ,

$$g(\dot\gamma, \dot\gamma) = 4r^2(-\mathrm{Z}^2 \dot{\mathrm{r}}^2 + \mathrm{r}^2 \dot\phi^2) = 4\mathrm{r}^2 \dot\phi^2 \mathrm{Z}^2 \left(-(\frac{\mathrm{dr}}{\mathrm{d}\phi})^2 + \frac{\mathrm{r}^2}{\mathrm{Z}^2} \right) \tag{22}$$

Since $r_q - r_s e^{\frac{\phi_s}{\mathrm{Z}}} > 0$ we have $\left|\frac{dr}{d\phi}\right| > \frac{|r|}{Z}$, so that the tangent to γ is everywhere timelike. We can choose the direction of parameterisation so that γ is future directed. Thus all points in L_q belong to $I^-(q)$. Moreover, no point outside L_q belongs to $I^-(q)$, since the curve would have to cross the small heart boundary at some point s; but according to the local lightcone at s any vector in its tangent space T_s pointing into L_q is either spacelike or past directed.

The causal past and future of q in this case are simply the closure of their chronological analogues since the bounding curves (the lightcones) are always geodesics through q.

That the above spacetime is causally continuous follows from the general result (lemma 1) about causal continuity of yarmulke spacetimes. But the reader can verify

[4]When the future directed null-spirals through a point q don't meet again within N_ϵ, the large heart B_q is not defined in N_ϵ. However, it is defined in some larger neighbourhood $(N_{\epsilon'}, g)$ with the same metric. Clearly, what our arguments would show in this case is that $I^+(q) = N_\epsilon - B_q \cap N_\epsilon$.

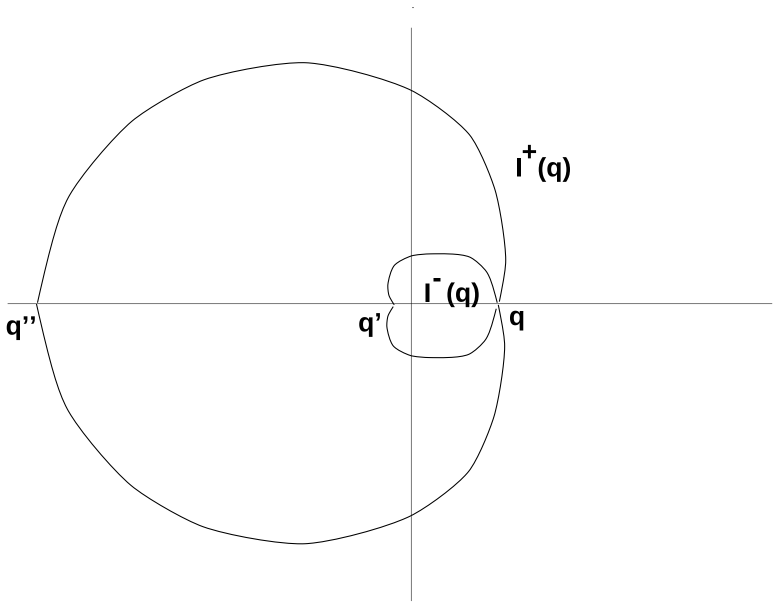

Figure 8: The yarmulke. Segments of the null geodesics through $q = (a, 0)$ that bound q's past and future. Their first intersection in the past occurs at the point $q' = (ae^{-\pi}, \pi)$ and their first intersection in the future occurs at $q'' = (ae^{+\pi}, \pi)$.

it graphically by finding the intersections of the past of all points to the future of q and vice-versa.

5.3 Pasts and futures for the Morse point

Violation of causal continuity in the trousers occurs along the null geodesics that have the origin as a limit point. Looking back at figure 2, one is inclined to regard these lines as the null cone of the critical point p, with $\mathcal{F}$ constituting its chronological future and $\mathcal{P}$ its chronological past. To make such a statement meaningful, we would have to restore p to our spacetime and extend the causal relation to the full Morse geometry. Although there is indeed a natural scheme for doing this [26], we don't need it for present purposes, since here we study the spacetimes without the critical points. If we nevertheless give a definite meaning to the past and future of p in this subsection, it is only for the convenience of being able to refer to $\mathcal{F}$, $\mathcal{P}$ and their higher dimensional generalisations in a simple way.

So let us temporarily consider the Morse points to be present and extend our definition of causal curve to embrace any piecewise smooth curve whose tangent vector v^a obeys $g_{ab}v^a v^b \leq 0$. With this definition, a curve emanating from, or terminating at a Morse point p can be causal (since $g_{ab}v^a v^b \leq 0$ automatically vanishes there), even though there is no way to tell from $g_{ab}(p)$ alone whether v^a is timelike or spacelike there. For the Morse point p we can then conveniently define $I^+(p)$ to be the interior of $J^+(p)$.

Let us apply these definitions to the $(1+1)$ dimensional examples from the previous sections. In the trousers it is easy to see that $I^+(p) = \mathcal{F}$ and $I^-(p) = \mathcal{P}$. For example, let us find $I^+(p)$. The straight lines joining p to points in $\mathcal{F} \cup \partial\mathcal{F}$ are future-directed and causal in the sense just described, so that $\mathcal{F} \subset I^+(p)$. To show the converse, we must check that $J^+(p) \cap (\mathcal{P} \cup \mathcal{S}) = \emptyset$. Now a future-directed causal curve in D_ϵ from p to a point $q \in N_\epsilon$ is equivalent, as far as the spacetime (N_ϵ, g) is concerned, to a past-directed causal curve starting at q which runs off to p, reaching points arbitrarily close to it. It follows that no causal curve starting at p can end at a point $q \in \mathcal{P} \cup \mathcal{S}$, since the causal past of any such q is bounded away from p. In the 1+1 yarmulke, on the other hand (with $\lambda = 0$), every radial geodesic starting at p is causal, whence $I^+(p)$ is the entire punctured disk, while $I^-(p)$ is empty.

In the causally discontinuous $(1+1)$ trousers, both $\mathcal{P}$ and $\mathcal{F}$ consist of two disconnected components. Moreover $\mathcal{F}$ "separates" $\mathcal{P}$, in the sense that any curve joining the two disconnected components of $\mathcal{P}$ necessarily traverses $\mathcal{F}$. Similarly $\mathcal{P}$ separates $\mathcal{F}$. The discontinuity of $I^-(\cdot)$ is manifest in that the past of any point on $\partial\mathcal{F}$ intersects only one of the two components of $\mathcal{P}$, while any point in $\mathcal{F}$ contains the whole of $\mathcal{P}$ in its past. Dual statements can be made regarding the discontinuity of $I^+(\cdot)$. In the causally continuous $(1+1)$ yarmulke and its time reverse, both $\mathcal{F}$ and $\mathcal{P}$ are trivially connected. In neither case is the topology of the "light cone" that seen in the neighbourhood of a non-degenerate point of spacetime.

All this seems to indicate that the source of causal discontinuity for the neighbourhood geometries resides in the topology of the pasts and the futures of the Morse point p, while the discontinuity itself, when present, occurs on their boundaries, $\partial J^{\pm}(p)$ — the "light cones of p". In the next section, we base our investigation of the higher dimensional cases on these intuitions.

Finally we remark that the chronology of the critical point can also be described in the language of TIPs and TIFs, as formulated in [27, 20]. A TIP or "terminal indecomposable past set" is defined as the past $I^-(\gamma)$ of a future-inextendible timelike curve γ (that is a curve γ which has no future limit point in the spacetime). The TIP $I^-(\gamma)$ can be represented by an imaginary point which serves as future endpoint of γ, and of any other inextendible timelike curve with the same past as γ. A TIF is the dual concept. A procedure can then be introduced for identifying those imaginary future and past endpoints –TIPs and TIFs– whose associated curves approach, in an appropriate sense, the same ideal point.

In the punctured trousers it is straightforward to show that there are two distinct TIPs associated with p, $\mathcal{P}_1$ and $\mathcal{P}_2$, corresponding to each of the future-inextendible timelike curves $\omega_1(t) = (x(t), y(t)) = (\frac{\epsilon}{2}e^{-t}, 0)$ and $\omega_2(t) = (-\frac{\epsilon}{2}e^{-t}, 0)$, $t \in [0, \infty)$. Similarly, there are two distinct TIFs, $\mathcal{F}_1$ and $\mathcal{F}_2$, generated by the past inextendible timelike curves $\gamma_1(t) = (0, \frac{\epsilon}{2}e^{t})$ and $\gamma_2(t) = (0, -\frac{\epsilon}{2}e^{t})$, $t \in (-\infty, 0]$. Thus each connected component of $I^+(p)$ is a single TIP and dually for $I^-(p)$. In the 1+1 yarmulke, on the other hand, no TIP belongs to the Morse point, while the entire punctured disk is its single TIF (generated, for example, by any radial geodesic). So in our neighbourhood spacetimes $I^-(p)$ can also be defined as the union of TIPs associated with p (namely the set of points $q \in N_\epsilon$ through which there passes a future-inextendible timelike curve running off to p). And dually for $I^+(p)$.

6 Causal continuity in higher dimensional neighbourhood spacetimes

We now examine the causal continuity of neighbourhood spacetimes with arbitrary λ for $n \geq 3$. For the yarmulke neighbourhoods lemma 1 shows that they are causally continuous without further work. Thus in this section we assume that $\lambda \neq 0, n$. The causal structure in these cases can be analysed by identifying higher dimensional analogues of the sets $\mathcal{P}$ and $\mathcal{F}$ discussed in section 5 and appropriately utilising the causal structures of the 2-dimensional spacetimes studied there.

6.1 Pasts and futures of the Morse point

Guided by the form of $\mathcal{F}$ and $\mathcal{P}$ in the trousers, it is easy to guess what are their analogues in the higher dimensional neighbourhood geometries. In fact, each radial line is either spacelike, timelike, or null, and we may define $\mathcal{F}$ as the union of the future-directed timelike radial lines and $\mathcal{P}$ as the union of the past-directed ones. That $\mathcal{F}$ is then truly $I^+(p)$, p being the Morse point, should be clear from the following discussion. We can characterise $\mathcal{P}$ and $\mathcal{F}$ in terms of the spherical coordinates used in equation (5). The angular coordinates (including the discrete angle if $\lambda = 1$ or $n-1$) are understood to vary over all their possible values.

$$
\begin{aligned}
\mathcal{P} &\equiv \{q \in N_\epsilon : r_q < m_1 \rho_q\} \\
\mathcal{F} &\equiv \{q \in N_\epsilon : r_q > m_2 \rho_q\}
\end{aligned} \tag{23}
$$

where ρ_q is the ρ coordinate of q, *etc.*, and m_1 and m_2 are as in (7). We refer to $R^2 = \rho^2 + r^2$ as the squared distance from the origin.

If λ is neither 1 nor $n-1$, $\mathcal{P}$ and $\mathcal{F}$ are connected sets. However when $\lambda = 1$ $\mathcal{P}$ comprises two disjoint components: $\mathcal{P}_1$, which has $\theta_0 = 0$ and $\mathcal{P}_2$, with $\theta_0 = \pi$. Moreover $\mathcal{F}$, which is connected ($n \geq 3$), then separates $\mathcal{P}$ in the sense that any curve between $\mathcal{P}_1$ and $\mathcal{P}_2$ necessarily intersects the $r = 0$ hyperplane, which is contained in $\mathcal{F}$. When $\lambda = n-1$, in the time-reversed geometry, $\mathcal{P}$ separates $\mathcal{F} = \mathcal{F}_1 \coprod \mathcal{F}_2$ similarly.

Let us define the quadrant χ_q to be the set of points with the same angular coordinates as q. (χ_q is undefined when either ρ_q or r_q vanishes.) The warped product form of (5) with respect to the pair (ρ, r), ensures that the geodesics in this quadrant are geodesics of the full metric.

We introduce an operator P_q which projects the points of our manifold into the quadrant χ_q, such that for $x \in N_\epsilon$ with $x = (\rho_x, \Theta_x, r_x, \Phi_x)$, $P_q x = (\rho_x, \Theta_q, r_x, \Phi_q)$. Similarly, if $\gamma(t)$ is a curve with coordinates $\gamma(t) = (\rho(t), \Theta(t), r(t), \Phi(t))$, then $P_q\gamma(t) = (\rho(t), \Theta_q, r(t), \Phi_q)$. A timelike curve γ projects to a timelike curve $P_q\gamma$ in the quadrant and a causal curve projects to a causal curve. It follows that $I^\pm(q, \chi_q) = I^\pm(q) \cap \chi_q$ and moreover that for any point y in $I^\pm(q)$ its projection $P_q y$ to χ_q must lie in $I^\pm(q, \chi_q)$. Hence, the causal structure in χ_q is that of the upper right quadrant of the trousers and this is illustrated in figures 9–11 in which we give the chronological pasts and futures of three representative points.

We again define the sets

$$
\begin{aligned}
\partial\mathcal{P} &\equiv \{x \in N_\epsilon : r_x = m_1 \rho_x\} \\
\partial\mathcal{F} &\equiv \{x \in N_\epsilon : r_x = m_2 \rho_x\} \\
\mathcal{S} &\equiv N_\epsilon - (\mathcal{P} \cup \mathcal{F} \cup \partial\mathcal{P} \cup \partial\mathcal{F})
\end{aligned} \tag{24}
$$

and let $\mathcal{P}_q$ be the intersection $\mathcal{P} \cap \chi_q$. Similarly, define $\partial\mathcal{P}_q \equiv \partial\mathcal{P} \cap \chi_q$, *etc.*

It is now clear why $\mathcal{P}$ and $\mathcal{F}$ are respectively $I^-(p)$ and $I^+(p)$ as defined in the previous section. Take for example $\mathcal{F}$. From the timelike character of the null radii

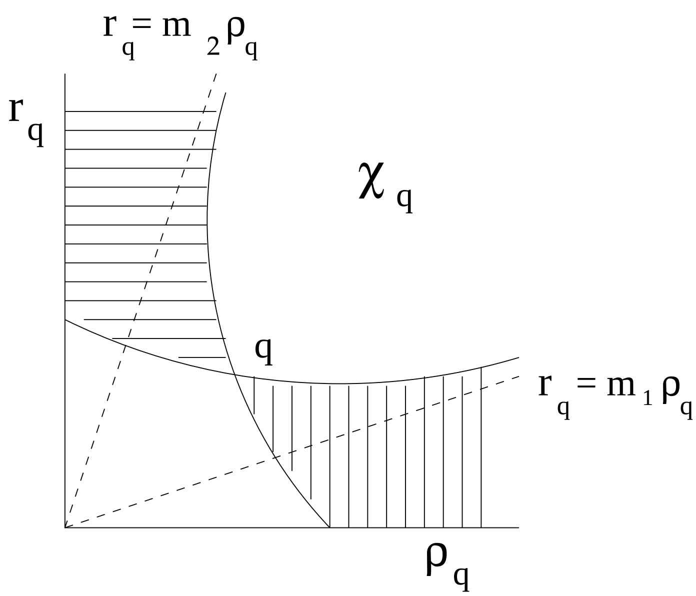

Figure 9: χ_q. $q \in \mathcal{S}$ and $I^+(q, \chi_q)$ ($I^-(q, \chi_q)$) is the horizontally (vertically) shaded region.

through points in $\mathcal{F}$, we immediately conclude $\mathcal{F} \subset I^+(p)$. To show the converse, namely that $J^+(q) \cap (\mathcal{P} \cup \mathcal{S}) = \emptyset$, it suffices to note that for any point $q \in \mathcal{P} \cup \mathcal{S}$, the set $J^-(q)$ is bounded away from the origin. This is because any causal curve through q projects to a causal curve in the quadrant χ_q, where the radial distance R along γ is seen to be bounded below.

Claim 1 *(i) If $q \in \mathcal{F}$, then $\mathcal{P} \subset I^-(q)$. (ii) If $q \in \partial\mathcal{F}$, then $I^+(q) \subset \mathcal{F}$. (iii) If $q \in \partial\mathcal{F}$, then (a) $\lambda \neq 1$ implies $\mathcal{P} \subset I^-(q)$, while (b) $\lambda = 1$ implies either $I^-(q) \cap \mathcal{P}_2 = \emptyset$ or $I^-(q) \cap \mathcal{P}_1 = \emptyset$.*

Proof: (i) We are to show that $z \ll q$ for every $z \in \mathcal{P}$ and $q \in \mathcal{F}$. Using the notation introduced in section 4, we can write $q = (x^i_q, y^j_q) = (\rho_q, \Theta_q, r_q, \Phi_q)$ and similarly $z = (x^i_z, y^j_z)$. Then for sufficiently small ϵ we have

$$z = (x^i_z, y^j_z) \ll (\epsilon x^i_z, 0) \ll (0, \epsilon y^j_q) \ll (x^i_q, y^j_q) = q$$

from which $z \ll q$ results by transitivity. Here, the first $\ll$ follows from the fact that both z and $(\epsilon x^i_z, 0)$ lie in the 2-d quadrant χ_z, which has the causal structure of the $1 + 1$ trousers, as represented by figures 9-11 (unless $y^i_z = 0$, in which case $\ll$ is trivial); and the last $\ll$ follows analogously. Finally, $(\epsilon x^i_z, 0) \ll (0, \epsilon y^j_q)$ follows

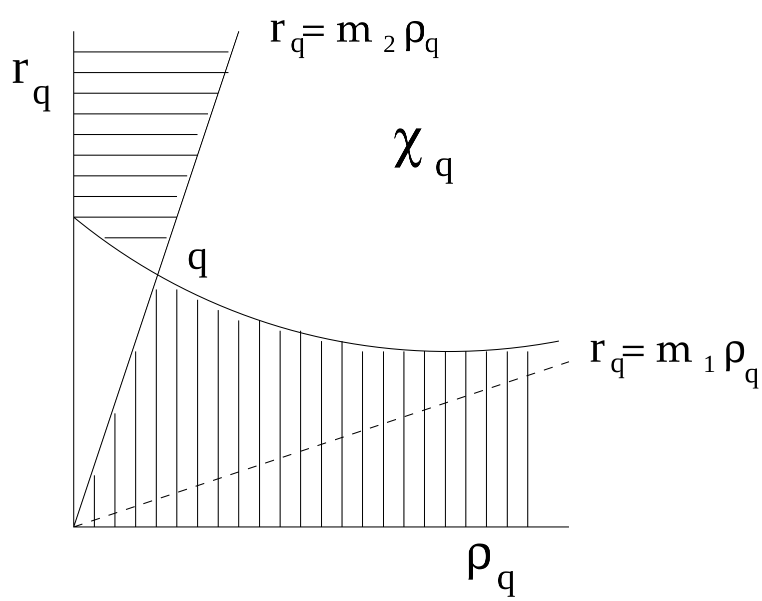

Figure 10: χ_q. $q \in \partial\mathcal{F}$. $I^+(q, \chi_q)$ $(I^-(q, \chi_q))$ is the horizontally (vertically) shaded region.

from the causal structure of the 2-d quadrant χ spanned by these two vectors (i.e. the quadrant with $\Theta = \Theta_z$ and $\Phi = \Phi_q$), see figures 4 and 11.

(ii) Let $q \in \partial\mathcal{F}$ and $z \in I^+(q)$. Consider the projection $P_q z$ into χ_q which must lie in $I^+(q, \chi_q)$. This gives us $r_z > m_2 \rho_z$ and so $z \in \mathcal{F}$. (We didn't actually need to prove this point explicitly, since $\mathcal{F}$ is easily seen to be a future set, namely $\mathcal{F} = I^+(\mathcal{F})$, and for any future set A we have $I^+(\partial A) \subset A$ [19].)

(iii) Let $q \in \partial\mathcal{F}$.

(a) $\lambda \neq 1$. Let $z \in \mathcal{P}$. Then for sufficiently small ϵ and δ, we have

$$z = (x_z^i, y_z^j) \ll (\epsilon x_z^i, 0) \ll (\epsilon\delta\, x_q^i, 0) \ll (x_q^i, y_q^j) = q$$

where this time, the final $\ll$ follows from the causal structure of χ_q (see figures 6 and 10), and the middle $\ll$ follows from that of the plane χ spanned by $(x_z^i, 0)$ and $(x_q^i, 0)$, which has the yarmulke causal structure (see figure 8). In this last step, we have assumed that x_z^i and x_q^i are linearly independent. If this is not the case (the only real danger being that they are *anti*-parallel), then we can take χ to be spanned by $(x_q^i, 0)$ and $(x_w^i, 0)$ for any x_w^i independent of x_q^i. (Such an x_q^i exists because $\lambda \geq 2$.)

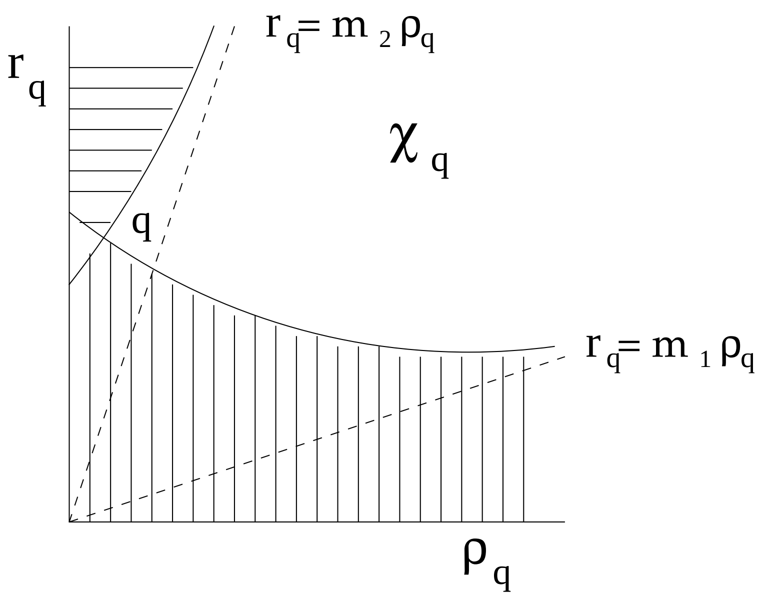

Figure 11: χ_q. $q \in \mathcal{F}$ and $I^+(q, \chi_q)$ $(I^-(q, \chi_q))$ is the horizontally (vertically) shaded region.

(b) $\lambda = 1$. Suppose, without loss of generality, that $\theta_0(q) = 0$ and let $z \in I^-(q)$. If $\theta_0(z) = \pi$ then there exists a future directed timelike curve from z to q which passes through $\rho = 0$. Considering the projection of this curve into χ_q shows this to be a contradiction. So $\theta_0(z) = 0$ and $I^-(q) \cap \mathcal{P}_2 = \emptyset$. $\square$

Finally, we remark that the relation of $\mathcal{P}$ and $\mathcal{F}$ to TIPs and TIFs is topologically just as in the 2-dimensional case: each connected component of $\mathcal{P}$ [resp. $\mathcal{F}$] is a single TIP [resp. TIF] associated with the absent point p in the neighbourhood spacetime N_ϵ.

6.2 Causal continuity in neighbourhood spacetimes of general index and dimension

Here we finally establish the conjectured relation between Morse index and causal continuity [16] for neighbourhood spacetimes around Morse points of arbitrary index. Instead of working directly in the Morse neighbourhood N_ϵ, we define a subset $Q_\delta \subset N_\epsilon$ which is still homeomorphic to a punctured ball around p, but which has null side boundaries. Remember that ultimately we would like to establish that a

Morse geometry (an entire compact cobordism) is causally continuous if and only if it contains no index 1 or $n-1$ points. Here we solve the analogue problem in an isolated punctured neighbourhood of a single Morse point. That causal continuity fails in (N_ϵ, g) when λ is 1 or $n-1$ is already apparent from claim 1. In fact this is true for the induced spacetime in any neighbourhood $U \subset N_\epsilon$ surrounding p, irrespective of the shape of U.

To show causal continuity for the remaining values of λ we have to be more careful. Indeed, by treating the neighbourhood as a separate spacetime we are introducing an artificial boundary, and with it a new potential source of causal discontinuity, absent in the Morse spacetime where the neighbourhood is actually embedded. It is to avoid this kind of artifact that we consider Q_δ instead of N_ϵ. The boundary ∂Q_δ is manifestly harmless to causal continuity, as will be clear in the proof of claim 3 below[5].

Our construction of Q_δ is inspired by the notion of "projection to the quadrant" introduced in the last section. Let N_ϵ be a round punctured neighbourhood of a critical point p of index λ, $c = f(p)$ be the critical value and $V_c = \{q \in N_\epsilon : f(q) = c\} \equiv \{q \in N_\epsilon : r(q) = \rho(q)\}$ be the critical surface in N_ϵ. Consider the sphere $S_\delta^{n-1} = \{q \in N_\epsilon : R(q) = \delta\}$ for some $\delta < \epsilon$. From equation (4) we see that the intersection $V_c \cap S_\delta^{n-1}$ is a product of spheres $S^{\lambda-1} \times S^{n-\lambda-1}$. Pick a point q in this set, with coordinates $(\delta/\sqrt{2}, \Theta_q, \delta/\sqrt{2}, \Phi_q)$, and draw the null hyperbolae through q in χ_q, as in figure 12. For any value of ζ we can choose δ small enough so that the segment of υ_q^- from q to the line $r = m_1\rho$ is contained in N_ϵ and so is the segment of υ_q^+ from q to the line $r = m_2\rho$. Call their respective endpoints q_1 and q_2. We take these null arcs to be the intersection of the side boundary of Q_δ with χ_q. Note that by construction $\partial Q_\delta \cap \chi_q = P_q \partial Q_\delta$.

The side boundary of Q_δ is the congruence of hyperbolic segments obtained by repeating the same procedure for all points in $V_c \cap S_\delta^{n-1}$. It has topology $S^{\lambda-1} \times S_\delta^{n-\lambda-1} \times D^1$. To define the lower and upper boundary of Q_δ we simply cap off the ends of this cylinder using the level surfaces V_a and V_b, where $a = f(q_1) = c - \alpha^2$, $b = f(q_2) = c + \alpha^2$. It is easy to show, using equation (14), that $\alpha^2 = \frac{1-m_1^2}{4m_1}\delta^2$. The whole boundary ∂Q_δ is then an $n-1$ sphere with corners and we define Q_δ to be the open region inside it[6]. In figure 13 we illustrate the construction of ∂Q_δ in a

[5] From the shape of the radial null geodesics, it seems reasonable to believe that the sphere bounding N_ϵ would not disrupt causal continuity either. This could be shown by studying the behaviour of R along general null geodesics at different values of r/ρ.

[6] Topologically Q_δ is just an n-ball. The side boundary resembles the standard surgery cylinder defined through integral curves of the gradient-like vector field in a neighbourhood of the critical

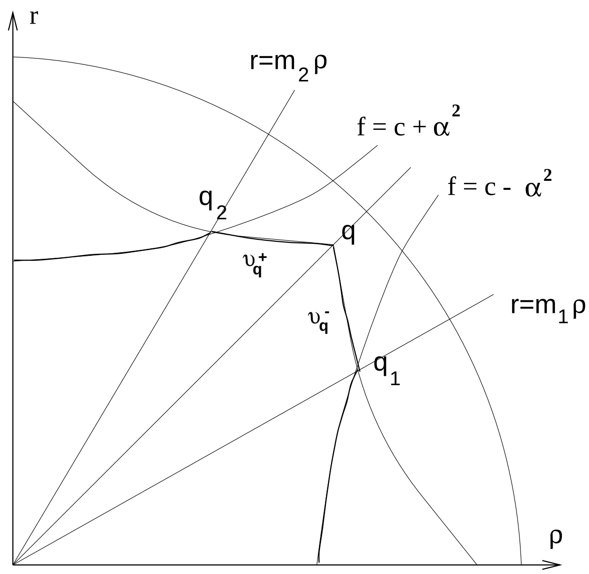

Figure 12: For given $q \in V_c \cap S_\delta^{n-1}$, the intersection $\partial Q_\delta \cap \chi_q$ is given by the null hyperbolae through q and by hyperbolae of constant f in χ_q.

neighbourhood of type $(1,2)$. The thick lines in figure 12 represent the intersection of ∂Q_δ with χ_q for general dimension n and $\lambda \neq 0, n$.

Like for N_ϵ, we say that the neighbourhood spacetime (Q_δ, g), obtained by restricting the metric (5) to Q_δ, is of type $(\lambda, n-\lambda)$. Let us examine causal continuity in these spacetimes for the different values of $\lambda \neq 0, n$.

Consider the spacetime (Q_δ, g) inside a Morse neighbourhood N_ϵ. Let V_c be the critical surface, δ be the value of R for points in $\partial Q_\delta \cap V_c$, as above, and a, b be the values of f at respectively the initial and final components of the spacelike part of ∂Q_δ. The whole set ∂Q_δ is a closed subset of N_ϵ. We have:

Claim 2 *Given any point $q \in Q_\delta$, the sets $I^+(q) \cap f^{-1}((a,c])$ and $I^-(q) \cap f^{-1}([c,b))$ are bounded away from ∂Q_δ.*

Proof: We prove that $I^+(q) \cap f^{-1}((a,c])$ is bounded away from ∂Q_δ. This is trivial when $f(q) \geq c$. So suppose there were a q with $f(q) < c$ for which

point [28]. Here the pair of hyperbolic segments through a point q play the role of the integral curve through q.

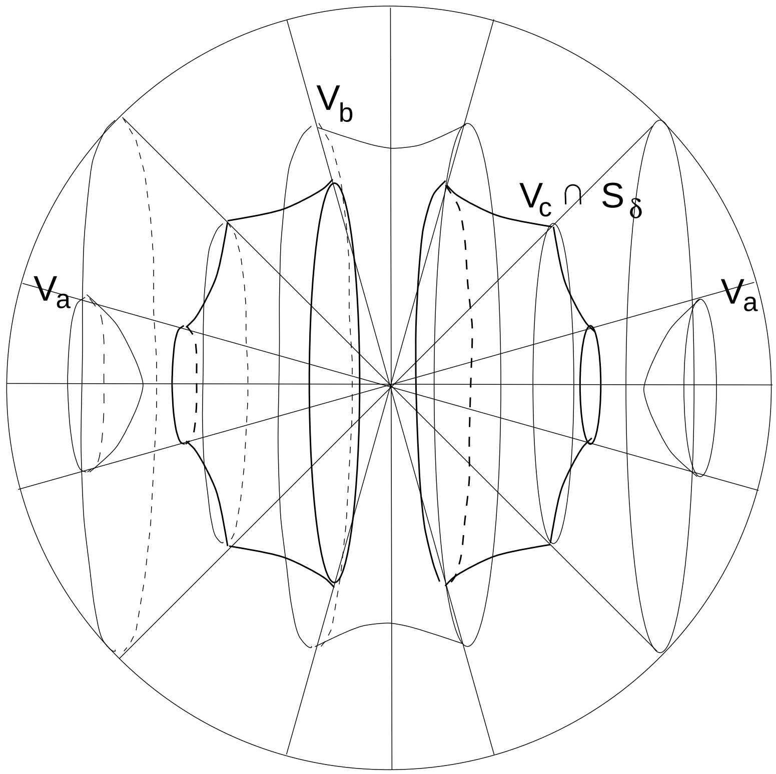

Figure 13: In a neighbourhood type $(1, 2)$ the side boundary is a cylinder $S^0 \times S^1 \times D^1$, drawn in thick lines. We close this cylinder to obtain ∂Q by gluing a subset $S^0 \times D^2$ of V_a at the "bottom" and a subset $S^1 \times D^1$ of V_b at the "top". We have drawn the bottom and top boundaries in thinner lines, to emphasize the distinction between null and spacelike parts of ∂Q_δ.

$I^+(q) \cap f^{-1}((a, c])$ was not bounded away from ∂Q_δ. Then there would be a sequence of timelike curves γ_k starting at q and reaching points y_k arbitrarily near ∂Q_δ before crossing V_c. Consider their projections $\gamma'_k = P_q \gamma_k$ to the quadrant χ_q, which we know are also timelike. At least one of the γ'_k must traverse the null hyperbola υ_q^-, since below V_c this remains at finite distance from ∂Q_δ. But this is not possible because the tangent vector to γ'_k at the intersection point would not be both future-directed and timelike. $\square$

Claim 3 *Consider a point q in a neighbourhood spacetime Q_δ. If there is a point $q' \in I^+(q)$ such that $I^-(q')$ is bounded away from the critical point p, then $\downarrow I^+(q) = I^-(q)$. Dually, if there is a point $q' \in I^-(q)$ such that $I^+(q')$ is bounded away from p, then $\uparrow I^-(q) = I^+(q)$.*

Proof: Suppose that the past of a point $q' \in I^+(q)$ is bounded away from p and consider a point $y \in \downarrow I^+(q)$. It has a neighbourhood U_y also contained in $\downarrow I^+(q)$. We want to show that $y \in I^-(q)$. Let the $q_k \to q$ be a sequence of points in $I^+(q)$ with $q_1 = q'$ and $q_{k+1} << q_k$. Pick a point $y' \in U_y \cap I^+(y)$ and a sequence of timelike curves γ_k from y' to q_k.

We plan to use the compactness of the space of causal curves in some compact subset of Q_δ to construct a timelike curve from y to q. To do this we must identify a compact region where all the γ_k's are contained. We already know that the γ_k's are bounded away from p. Indeed, there must be an open ball $D_{\epsilon'}$ around p for which $I^-(q') \cap D_{\epsilon'} = \emptyset$. Then $\gamma_k \subset Q_\delta - D_{\epsilon'}\ \forall\, k$, since $I^-(q_k) \subset I^-(q')$.

We now show that the γ_k's are moreover bounded away from ∂Q_δ. Since $\gamma_k \subset I^+(y') \cap I^-(q')\ \forall\, k$, it suffices to show that this set is bounded away from ∂Q_δ. Now $I^+(y') \cap I^-(q')$ is contained in the union of $I^+(y') \cap f^{-1}((a,c])$ and $I^-(q') \cap f^{-1}([c,b))$, which by claim 2 are both bounded away from ∂Q_δ. Therefore there must be a $\delta' < \delta$ such that $\gamma_k \subset Q_{\delta'}\ \forall\, k$.

We conclude that all of the γ_k are contained in the compact set $\overline{Q_{\delta'}} - D_{\epsilon'} \subset Q_\delta$. Since the spacetime (Q_δ, g) is strongly causal, the space of causal curves in $\overline{Q_{\delta'}} - D_{\epsilon'}$ is compact. This implies that the γ_k have a limit causal curve γ between y' and q, contained in $\overline{Q_{\delta'}} - D_{\epsilon'}$. Therefore $y << y' < q$ in (Q_δ, g), so that $y \in I^-(q)$ (proposition 2.18 [19]), as desired. The proof of the dual statement is similar. $\Box$

The sets $\mathcal{P}$, $\mathcal{F}$ and $\mathcal{S}$ intersect Q_δ in the obvious regions, which we also call $\mathcal{P}$, $\mathcal{F}$ and $\mathcal{S}$. There is no ambiguity, since from now on we will only be talking about the spacetime (Q_δ, g). It is straightforward to verify that claim 1 also applies for the $\mathcal{P}$ and $\mathcal{F}$ in the spacetime (Q_δ, g).

It should also be clear from the construction of Q_δ that for any $x \in Q_\delta$ and any number $0 < a \leq 1$ the point ax is also in Q_δ. We use the following notation to denote scale transformations by a real number a: given a point $q = (\rho_q, \Theta_q, r_q, \Phi_q)$ and a curve $\gamma(t) = (\rho(t), \Theta(t), r(t), \Phi(t))$, we write aq for the point $aq = (a\rho_q, \Theta_q, ar_q, \Phi_q)$ and $a\gamma$ for the curve $a\gamma(t) = (a\rho(t), \Theta(t), ar(t), \Phi(t))$. Notice that the timelike character of a curve is preserved under this scaling.

We are now ready to prove that in a neighbourhood spacetime (Q_δ, g) of type $(\lambda, n-\lambda)$ with $\lambda \neq 0, n$ causal continuity always holds at a point $q \in \mathcal{P} \cup \mathcal{F} \cup \mathcal{S}$. It fails at points $q \in \partial\mathcal{F}$ only when $\lambda = 1$, for then $I^-(\cdot)$ changes abruptly at q, and at points $s \in \partial\mathcal{P}$ only when $\lambda = n-1$, for then $I^+(\cdot)$ changes abruptly at s.

Remember that all our statements here refer to the spacetime (Q_δ, g) and subsets therein.

Lemma 3 *For any value of the Morse index λ, causal continuity holds at all points $q \in \mathcal{S}$.*

Proof: We show that $I^-(q) = \downarrow I^+(q)$. By claim 3 we just need to find a point $q' \in I^+(q)$ whose past is bounded away from p. So let q' be any point in $I^+(q) \cap \mathcal{S}$. Clearly such a point exists and its past bounded away from p, as can be seen by considering $I^-(q') \cap \chi_{q'} = I^-(q', \chi_q)$.

$I^+(q) = \uparrow I^-(q)$ is proved similarly. □

In our final two lemmas we use the intuitively clear result that if for a point x in a spacetime M the difference $\downarrow I^+(x) - I^-(x)$ is not empty, neither is its interior. More explicitly:

Claim 4 *Given a spacetime (M, g), suppose that for a point $x \in M$ there exists a $y \in \downarrow I^+(x)$ such that $y \notin I^-(x)$, then there also exists a $z \in \downarrow I^+(x)$ such that $z \notin \overline{I^-(x)}$. Dually, if $\uparrow I^-(x) - I^+(x) \neq \emptyset$, then there exists a point z in $\uparrow I^-(x)$ such that $z \notin \overline{I^+(x)}$.*

Proof: We prove the first part of the claim. Suppose $\downarrow I^+(x) - I^-(x)$ is not empty and all of its points lie in $\partial I^-(x)$. Pick one such y. By the obvious generalisation of Proposition 6.3.1[20] to a Lorentzian metric in general dimension, the boundary $\partial I^-(x)$ is an $n-1$ dimensional submanifold, so that any ball centered on the boundary intersects both $I^-(x)$ and the complement of $\overline{I^-(x)}$. Thus every neighbourhood of y intersects the complement of $\overline{I^-(x)}$; hence no neighbourhood of y is contained in $\downarrow I^+(x)$, contradicting the fact that $\downarrow I^+(x)$ is open. The proof of the second part is similar. □

Lemma 4 *For any index λ, causal continuity holds at all points $q \in \mathcal{P} \cup \mathcal{F}$.*

Proof: Let $q \in \mathcal{F}$. First, we prove that $\downarrow I^+(q) = I^-(q)$. Suppose not. Then there is a point y and a neighbourhood U_y of y such that $U_y \subset \downarrow I^+(q)$ and $U_y \cap I^-(q) = \emptyset$.

Define the sequence of points $q_k = a_k q$ where $a_k = (1 + \frac{\delta}{k})$, $k = 1, 2, \ldots$ and $\delta > 0$ is small enough that $q_1 \in N_\epsilon$. The q_k tend to q and lie along the radial timelike line from the origin through q. So $q_k \in I^+(q)$, $\forall k$. Thus there exists a future directed timelike curve γ_k from y to q_k. Let $\gamma'_k = a_k^{-1}\gamma_k$, again a future directed timelike curve. The final point of each of these scaled curves is q and the initial point of γ'_k is $y_k = a_k^{-1}y$. Choose k large enough so that $y_k \in U_y$. This is a contradiction.

To prove that $\uparrow I^-(q) = I^+(q)$ we just need to find a point $q' \in I^-(q)$ whose future is bounded away from p, since then we can use claim 3. Now for any point $q' \in I^-(q) \cap \mathcal{F}$, the set $I^+(q')$ is bounded away from p. This is because f increases along every future-directed timelike curve and there are small neighbourhoods of p, $D_{\epsilon'}$, such that every $x \in D_{\epsilon'}$ has $f(x) < f(q)$.

Causal continuity at points in $\mathcal{P}$ is proved similarly. □

The only potential obstructions to global causal continuity therefore lie in the remaining region $\partial\mathcal{F} \cup \partial\mathcal{P}$. Combining the method of the previous lemma and the results gathered in the last section we can prove.

Lemma 5 *(a) When $\lambda \neq 1$ causal continuity holds at all points $q \in \partial\mathcal{F}$. (b) When $\lambda = 1$, if $q \in \partial\mathcal{F}$ then $I^-(q) \neq \downarrow I^+(q)$, while $I^+(q) = \uparrow I^-(q)$. (c) When $\lambda \neq n-1$ causal continuity holds at all points $q \in \partial\mathcal{P}$. (d) When $\lambda = n-1$, if $q \in \partial\mathcal{P}$ then $I^+(q) \neq \uparrow I^-(q)$, while $I^-(q) = \downarrow I^+(q)$.*

Proof: Consider $q \in \partial\mathcal{F}$.

(a) $\lambda \neq 1$. We first show that $\downarrow I^+(q) = I^-(q)$. Suppose not. Then as usual, there is a point y with a neighbourhood, U_y, such that $U_y \subset \downarrow I^+(q)$ and $U_y \cap I^-(q) = \emptyset$. First $y \notin \mathcal{P}$ by claim 1 (iiia); also $y \notin \partial\mathcal{P}$ since otherwise some other point in U_y would be in $\mathcal{P}$. Secondly $y \notin \mathcal{F}$, since otherwise considering a sequence of points $q_k \to q$, with $q_k \in I^+(q, \chi_q)$ and the fact that $P_q y \in I^-(q_k, \chi_q)$ $\forall k$ would lead to a contradiction with the known causal structure of χ_q; again $y \notin \partial\mathcal{F}$ since otherwise $U_y \cap \mathcal{F} \neq \emptyset$ and the same contradiction would arise. Finally if y lies in $\mathcal{S}$ then using arguments similar to those of lemma 3 we would obtain once more a contradiction. That $\uparrow I^-(q) = I^+(q)$ is proven as for points in $\mathcal{F}$ (lemma 4), since $f(q) > c$.

(b) $\lambda = 1$. Without loss of generality, assume $\theta_0(q) = 0$. Let $y \in \mathcal{P}_2$. Then $x \in I^+(q) \Rightarrow x \in \mathcal{F}$ by claim 1(ii) $\Rightarrow y \in I^-(x)$ by claim 1(i) and similarly for any point y' in a neighbourhood, U_y, of y such that $U_y \subset \mathcal{P}_2$. Thus $y \in \downarrow I^+(q)$. But

$y \notin I^-(q)$ by the proof of claim 1(iiib).

Parts (c) and (d) are proved similarly. $\Box$

Putting together the partial results in lemmas 1, 2 4, 3 and 5, we have a proof of the following proposition.

Proposition 3 *Let* (Q_δ, g) *be a neighbourhood spacetime of type* $(\lambda, n-\lambda)$*. Then*

(i) If $\lambda \neq 1, n-1$*, then* (Q_δ, g) *is causally continuous.*

(ii) If $\lambda = 1$ *or* $\lambda = n-1$*, then* (Q_δ, g) *is not causally continuous.*

7 Discussion

We have made progress towards proving the conjecture [16] that a Morse geometry presents causal discontinuities when, and only when, it contains critical points of index $\lambda = 1, n-1$. In particular we have proved this for a class of spacetimes defined in neighbourhoods of single Morse points. We have also proved some more general results. The causal continuity of the yarmulke spacetimes has been demonstrated for *any* Morse metric on these cobordisms. Moreover we have seen that any strongly causal compact product spacetime must be a Morse metric. This indicates that the class of Morse spacetimes is in fact quite general and hence supports the proposal to sum over Morse metrics in the SOH.

In order to obtain a full proof, we first need to extend our results in the neighbourhood to more general metrics than those built from the Cartesian flat Riemannian auxiliary metric. The next step would be to understand how the causal properties of the individual neighbourhoods affect the causal properties of the entire Morse spacetime within which they are embedded. We address these issues in a forthcoming paper [29].

We may take note here of the intuitive meaning of Lemma 5. At a Morse point of index $\lambda = n-1$ (i.e. $+ - - \ldots -$), the universe intuitively is splitting into two disconnected parts. In this case, Lemma 5 tells us that past(x) varies continuously with x, but future(x) jumps when x crosses $\partial\mathcal{P}$. When the index is 1, the situation is time reversed; two components of the universe come together and it is the pasts that vary discontinuously. In all other types of topology change, both pasts and futures vary continuously throughout.

We mentioned in the introduction our intention, eventually, to consider the Morse points as part of the spacetime and not to excise them. A causal order that includes the critical points is desirable, especially since it does not seem plausible that isolated points should be relevant in the quantum context. Most simply we can add in the degenerate point, p, and extend the causal relation by hand as done in section 5.3, where J was extended simply by allowing causal curves to originate or terminate at the Morse point. For our neighbourhood spacetimes, the result was that $J^-(p) = \overline{\mathcal{P}}$ and $J^+(p) = \overline{\mathcal{F}}$, where the closure is taken in the unpunctured disc.

Though simple, this prescription may appear a little ad hoc. We believe, however, that the causal order obtained in this way coincides with a robust generalisation of the usual causal relation proposed in [23]. This new relation, called K, is defined in terms of the chronology relation I, which can be extended to the points of degeneracy essentially at will, for example by using the strict definition for chronology, i.e., by putting $I^\pm(p) = \phi$. One can immediately verify that the causal relation proposed above for the neighbourhood geometries coincides with K. Our simple extension of the chronology to the critical points thus seems justified by the robustness of K, which is indifferent to the presence or absence of isolated points in a spacetime. A remarkable feature of this new setting is that the property of causal continuity appears as the condition that the pair (I, K) constitute a causal structure in the axiomatic sense of [22]. Details of this analysis will appear elsewhere [26].

Even if further work is required before we can understand the physical relevance of our results, in particular the effect of causal (dis)continuity on the propagation of quantum fields, we can already make an interesting observation: if the universe did begin in a big bang that could be described in its earliest moments by a Morse metric with an index 0 point, then the causal structure of the yarmulke is such that there are no particle horizons: all points on a given level surface of the Morse function have past points in common. That would mean that there would no longer be a cosmological "horizon problem".

8 Acknowledgments

We would like to thank Andrew Chamblin and Bernard Kay for interesting discussions. AB thanks the Inter-University Center for Astronomy and Astrophysics, Pune, India, and the Institute of Cosmology at Tufts University for their hospitality, and the Faculty Research Awards Committee at Southampton College for partial financial support. SS was supported by a postdoctoral fellowship from T.I.F.R., Mumbai. HFD was supported by an EPSRC Advanced Fellowship. RSG is supported by a Beit Fellowship. RDS was partly supported by NSF grant PHY-9600620 and by a grant from the Office of Research and Computing of Syracuse University.

References

[1] R.P.Geroch. *J. Math. Phys*, 8:782, 1967.

[2] R.D. Sorkin. *Phys. Rev.*, D33:978, 1986.

[3] R.D.Sorkin. *Int.J.Theor.Phys.*, 25:877, 1986.

[4] P. Yodzis. *Gen.Rel. and Grav.*, 4:299, 1973.

[5] R.D.Sorkin. In A.Coley, F.Cooperstock, and B.Tupper, editors, *Proceedings of the Third Canadian Conference of General Relativity and Relativistic Astrophysics (Victoria, Canada, May 1989)*, Singapore, 1990. World Scientific.

[6] S.W. Hawking. *Phys.Rev.D*, 18:1747–1753, 1978.

[7] S.W. Hawking. *Nucl.Phys.B*, 144:349–362, 1978.

[8] Gary T. Horowitz. *Class.Quant.Grav.*, 8:587–602, 1991.

[9] G.W. Gibbons and S.W. Hawking. *Commun.Math.Phys.*, 148:345–352, 1992.

[10] Alexander Vilenkin. *Phys.Rev. D*, 50:2581–2594, 1994. gr-qc/9403010.

[11] S. Carlip. *Class.Quant.Grav*, 15:2629–2638, 1998. gr-qc/9710114.

[12] Jorma Louko Alan Daughton and Rafael D. Sorkin. "Instantons and unitarity in quantum cosmology with fixed four-volume" *Phys. Rev. D*, 58:084008, 1998. gr-qc/9805101.

[13] H.F. Dowker and S.Surya. *Phys.Rev. D*, 58:124019, 1998. gr-qc/9711070.

[14] H.F. Dowker and R.S. Garcia. *Class.Quant.Grav.*, 15:1859–1879, 1998. gr-qc/9711042.

[15] Jorma Louko and Rafael D. Sorkin. "Complex actions in two-dimensional topology change" *Class.Quant.Grav.*, 14:179–204, 1997. gr-qc/9511023.

[16] A. Borde and R.D. Sorkin. Unpublished.

[17] A. Anderson and B. DeWitt. *Found. Phys.*, 16:91–105, 1986.

[18] C.A. Manogue, E. Copeland, and T. Dray. *Pramana*, 30:279, 1988.

[19] R.Penrose. *Techniques of differential topology in relativity.* Society for Industrial and Applied Mathematics, Pennsylvania, 1973.

[20] S.W. Hawking and G.F.R. Ellis. *The large scale structure of space-time.* Cambridge University Press, Cambridge, 1973.

[21] S.W. Hawking and R.K. Sachs. *Comm. Math. Phys*, 35:287–296, 1974.

[22] E.H.Kronheimer and R.Penrose. *Proc.Cam.Phil.Soc*, 63:481, 1967.

[23] R.D.Sorkin and E.Woolgar. *Class.Quant.Grav.*, 13:1971–1994, 1996. gr-qc/9508018.

[24] John Willard Milnor. *Morse theory.* Princeton University Press, 1963.

[25] S.G. Harris and T. Dray. *Class. Quantum Grav.*, 7:149, 1990.

[26] H.F. Dowker, R.S. Garcia, and S. Surya. In preparation.

[27] R.P.Geroch, E.H.Kronheimer, and R.Penrose. *Proc.Roy.Soc.Lond.A*, 327:545, 1972.

[28] John Willard Milnor. *"Lectures on the h-cobordism theorem".* Princeton University Press, 1965.

[29] H.F. Dowker, R.S. Garcia, and S. Surya. In preparation.