On the Geometry and Topology of Initial Data Sets in General Relativity

Greg Galloway
University of Miami

GeLoMa 2016 – 8th International Meeting on Lorentzian Geometry
Málaga, Spain, September 2016

References for Causal Theory

[1] G. J. Galloway, *Notes on Lorentzian causality*, ESI-EMS-IAMP Summer School on Mathematical Relativity (available at: `http://www.math.miami.edu/~galloway/`).

[2] B. O'Neill, *Semi-Riemannian geometry*, Pure and Applied Mathematics, vol. 103, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983.

****

[3] J. K. Beem, P. E. Ehrlich, and K. L. Easley, *Global Lorentzian geometry*, second ed., Monographs and Textbooks in Pure and Applied Mathematics, vol. 202, Marcel Dekker Inc., New York, 1996.

[4] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*, Cambridge University Press, London, 1973, Cambridge Monographs on Mathematical Physics, No. 1.

[5] R. Penrose, *Techniques of differential topology in relativity*, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1972, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 7.

[6] Robert M. Wald, *General relativity*, University of Chicago Press, Chicago, IL, 1984.

Lorentzian manifolds

We start with an (n+1)-dimensional Lorentzian manifold (M, g). Thus, at each $p \in M$,

$$g : T_pM \times T_pM \to \mathbb{R}$$

is a scalar product of signature $(-, +, ..., +)$. With respect to an orthonormal basis $\{e_0, e_1, ..., e_n\}$, as a matrix,

$$[g(e_i, e_j)] = \mathrm{diag}(-1, +1, ..., +1)\,.$$

Example: Minkowski space, the spacetime of Special Relativity. Minkowski space is $\mathbb{R}^{n+1}$, equipped with the Minkowski metric η: For vectors $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^i \frac{\partial}{\partial x^i}$ at p, (where x^i are standard Cartesian coordinates on $\mathbb{R}^{n+1}$),

$$\eta(X, Y) = \eta_{ij} X^i X^j = -X^0 Y^0 + \sum_{i=1}^{n} X^i Y^i\,.$$

Thus, each tangent space of a Lorentzian manifold is isometric to Minkowski space. This builds in the locally accuracy of Special Relativity in General Relativity.

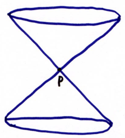
Causal character of vectors.

At each point, vectors fall into three classes, as follows:

$$X \text{ is } \begin{cases} \text{timelike} & \text{if } g(X,X) < 0 \\ \text{null} & \text{if } g(X,X) = 0 \\ \text{spacelike} & \text{if } g(X,X) > 0 \,. \end{cases}$$

A vector X is *causal* if it is either timelike or null.

The set of null vectors $X \in T_pM$ forms a double cone $\mathcal{V}_p$ in the tangent space T_pM:



called the null cone (or light cone) at p.
Timelike vectors point inside the null cone and spacelike vectors point outside.

Time orientability.

At each $p \in M$ we have a double cone; label one cone the *future* cone and the other a *past cone*.

If this assignment of a past and future cone can be made in a continuous manner over all of M then we say that M is *time-orientable*.

Time orientability (cont.).

There are various ways to make the phrase "continuous assignment" precise (see e.g., O'Neill p. 145), but they all result in the following:

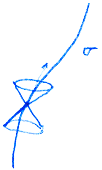
Fact: A Lorentzian manifold M^{n+1} is time-orientable iff it admits a smooth timelike vector field T.

If M is time-orientable, the choice of a smooth time-like vector field T fixes a time orientation on M: A causal vector $X \in T_pM$ is *future pointing* if it points into the same half-cone as T, and *past pointing* otherwise.

(Remark: If M is not time-orientable, it admits a double cover that is.)

By a spacetime we mean a connected time-oriented Lorentzian manifold (M^{n+1}, g).

Lorentzian Causality



Causal character of curves.

Let $\gamma : I \to M$, $t \to \gamma(t)$ be a smooth curve in M.

- γ is said to be timelike provided $\gamma'(t)$ is timelike for all $t \in I$.

  In GR, a timelike curve corresponds to the history (or *worldline*) of an observer.

- Null curves and spacelike curves are defined analogously.

  A causal curve is a curve whose tangent is either timelike or null at each point.

- The length of a causal curve $\gamma : [a, b] \to M$, is defined by

$$L(\gamma) = \text{Length of } \gamma = \int_a^b |\gamma'(t)| dt = \int_a^b \sqrt{-\langle \gamma'(t), \gamma'(t) \rangle}\, dt\,.$$

  Owing to the Lorentz signature, causal geodesics locally maximize length.

  If γ is timelike one can introduce arc length parameter along γ. In general relativity, the arc length parameter along a timelike curve is called proper time, and corresponds to time kept by the observer.

Lorentzian Causality

Futures and Pasts

Let (M, g) be a spacetime. A timelike (resp. causal) curve $\gamma : I \to M$ is said to be *future directed* provided each tangent vector $\gamma'(t)$, $t \in I$, is future pointing. (*Past-directed* timelike and causal curves are defined in a time-dual manner.)

Causal theory is the study of the causal relations $\ll$ and $<$:

Definition 1.1

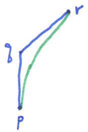
For $p, q \in M$,

1. *$p \ll q$ means there exists a future directed timelike curve in M from p to q (we say that q is in the timelike future of p),*
2. *$p < q$ means there exists a (nontrivial) future directed causal curve in M from p to q (we say that q is in the causal future of p),*

We shall use the notation $p \leq q$ to mean $p = q$ or $p < q$.

The causal relations $\ll$ and $<$ are clearly transitive. Also, from variational considerations, it is heuristically clear that the following holds,

$$\text{if } p \ll q \text{ and } q < r \text{ then } p \ll r\,.$$



Proposition 1.1 (O'Neill, p. 294)

In a spacetime M, if q is in the causal future of p ($p < q$) but is not in the timelike future of p ($p \not\ll q$) then any future directed causal curve γ from p to q must be a null geodesic (when suitably parameterized).

Now introduce standard causal notation:

Definition 1.2

Given any point p in a spacetime M, the timelike future and causal future of p, denoted $I^+(p)$ and $J^+(p)$, respectively, are defined as,

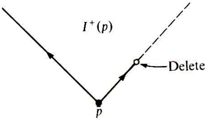
$$I^+(p) = \{q \in M : p \ll q\} \quad \textit{and} \quad J^+(p) = \{q \in M : p \leq q\} .$$

The timelike and causal *pasts* of p, $I^-(p)$ and $J^-(p)$, respectively, are defined in a time-dual manner in terms of past directed timelike and causal curves.

With respect to this notation, the above proposition becomes:

Propostion *If $q \in J^+(p) \setminus I^+(p)$ ($q \neq p$) then there exists a future directed null geodesic from p to q.*

In general, sets of the form $I^+(p)$ are open (see e.g. Gal-ESI). However, sets of the form $J^+(p)$ need not be closed, as can be seen by removing a point from Minkowski space.



For any subset $S \subset M$, we define the timelike and causal future of S, $I^+(S)$ and $J^+(S)$, respectively by

$$I^+(S) = \bigcup_{p \in S} I^+(p) = \{q \in M : p \ll q \text{ for some } p \in S\} \tag{1}$$

$$J^+(S) = \bigcup_{p \in S} J^+(p) = \{q \in M : p \leq q \text{ for some } p \in S\}\,. \tag{2}$$

Note:

- $S \subset J^+(S)$.
- $I^+(S)$ is open (union of open sets).

$I^-(S)$ and $J^-(S)$ are defined in a time-dual manner.

Achronal Boundaries

Achronal sets play an important role in causal theory.

Definition 1.3

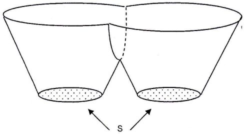
A subset $S \subset M$ is achronal provided no two of its points can be joined by a timelike curve.

Of particular importance are *achronal boundaries*.

Definition 1.4

An achronal boundary is a set of the form $\partial I^+(S)$ (or $\partial I^-(S)$), for some $S \subset M$.

The following figure illustrates some of the important structural properties of achronal boundaries.



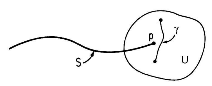
Proposition 1.2

An achronal boundary $\partial I^+(S)$, if nonempty, is a closed achronal C^0 hypersurface in M.

Claim A: An achronal boundary $\partial I^+(S)$ is achronal.

Definition 1.5

Let $S \subset M$ be achronal. Then $p \in \overline{S}$ is an edge point of S provided every neighborhood U of p contains a timelike curve γ from $I^-(p, U)$ to $I^+(p, U)$ that does not meet S.



We denote by $\mathrm{edge}\, S$ the set of edge points of S. If $\mathrm{edge}\, S = \emptyset$ we say that S is *edgeless*.

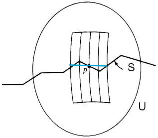
Claim B: An achronal boundary is *edgeless*.

Claims A and B follow easily from the following simple fact; see Gal-ESI for details.

Fact: If $p \in \partial I^+(S)$ then $I^+(p) \subset I^+(S)$, and $I^-(p) \subset M \setminus \overline{I^+(S)}$.

Claim C: An edgeless achronal set S, if nonempty, is a C^0 hypersurface in M.

Proof: See O'Neill, p. 413.



Causality conditions

A number of results in Lorentzian geometry and general relativity require some sort of causality condition.

Chronology condition: A spacetime M satisfies the *chronology condition* provided there are no closed timelike curves in M.

Compact spacetimes have limited interest in general relativity since they all violate the chronology condition.

Proposition 1.3

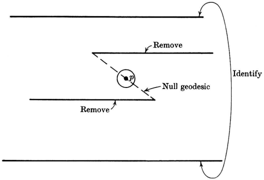
Every compact spacetime contains a closed timelike curve.

Proof: The sets $\{I^+(p); p \in M\}$ form an open cover of M from which we can abstract a finite subcover: $I^+(p_1)$, $I^+(p_2)$, ..., $I^+(p_k)$. We may assume that this is the minimal number of such sets covering M. Since these sets cover M, $p_1 \in I^+(p_i)$ for some i. It follows that $I^+(p_1) \subset I^+(p_i)$. Hence, if $i \neq 1$, we could reduce the number of sets in the cover. Thus, $p_1 \in I^+(p_1)$ which implies that there is a closed timelike curve through p_1. □

A somewhat stronger condition than the chronology condition is the

Causality condition: A spacetime M satisfies the causality condition provided there are no closed (nontrivial) causal curves in M.

A spacetime that satisfies the causality condition can nontheless be on the verge of failing it, in the sense that there exist causal curves that are “almost closed”, as illustrated by the following figure.



Strong causality is a condition that rules out almost closed causal curves.

Definition 1.6

An open set U in spacetime M is said to be causally convex provided every causal curve segment with end points in U lies entirely within U.

Definition 1.7

Strong causality is said to hold at $p \in M$ provided p has arbitrarily small causally convex neighborhoods, i.e., for each neighborhood V of p there exists a causally convex neighborhood U of p such that $U \subset V$.

Note that strong causality fails at the point p in the figure above. It can be shown that the set of points at which strong causality holds is open.

Strong causality condition: A spacetime M is said to be strongly causal if strong causality holds at all of its points.

This is the "standard" causality condition in spacetime geometry, and, although there are even stronger causality conditions, it is sufficient for most applications.

The following lemma is often useful.

Lemma 1.4

Suppose strong causality holds at each point of a compact set K in a spacetime M. If $\gamma : [0, b) \to M$ is a future inextendible causal curve that starts in K then eventually it leaves K and does not return, i.e., there exists $t_0 \in [0, b)$ such that $\gamma(t) \notin K$ for all $t \in [t_0, b)$.

(γ is future inextendible if it cannot be continuously extended, i.e. if $\lim_{t \to b^-} \gamma(t)$ does not exist.)

We say that a future inextendible causal curve cannot be "imprisoned" in a compact set on which strong causality holds.

Global hyperbolicity

We now come to a fundamental condition in spacetime geometry, that of global hyperbolicity.

Mathematically, global hyperbolicity is a basic 'niceness' condition that often plays a role analogous to geodesic completeness in Riemannian geometry. Physically, global hyperbolicity is closely connected to the issue of classical determinism and the strong cosmic censorship conjecture.

Definition 1.8

A spacetime M is said to be globally hyperbolic provided

- *M is strongly causal.*
- *(Internal Compactness) The sets $J^+(p) \cap J^-(q)$ are compact for all $p, q \in M$.*

Remarks:

1. Condition (2) says roughly that M has no holes or gaps.
2. In fact, as shown by Bernal and Sanchez [4], internal compactness + causality imply strong causality.

Proposition 1.5

Let M be a globally hyperbolic spacetime. Then,

1. *The sets $J^{\pm}(A)$ are closed, for all compact $A \subset M$.*
2. *The sets $J^{+}(A) \cap J^{-}(B)$ are compact, for all compact $A, B \subset M$.*

Global hyperbolicity is the standard condition in Lorentzian geometry that ensures the existence of maximal causal geodesic segments.

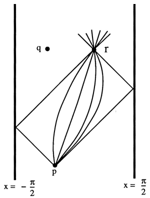
Theorem 1.6

Let M be a globally hyperbolic spacetime. If $q \in J^{+}(p)$ then there is a maximal future directed causal geodesic from p to q (i.e., no causal curve from p to q can have greater length).

See Gal-ESI for discussions of the proofs.

Contrary to the situation in Riemannian geometry, geodesic completeness does not guarantee the existence of maximal segments.

<ins>Ex.</ins> Two-dimensional anti-de Sitter space:
$M = \{(t, x) : -\pi/2 < x < \pi/2\}$, $ds^2 = \sec^2 x(-dt^2 + dx^2)$.
All future directed timelike geodesics emanating from p refocus at r.
There is no timelike geodesic from p to $q = I^{+}(p)$.



Cauchy hypersurfaces

Global hyperbolicity is closely related to the existence of certain 'ideal initial value hypersurfaces', called Cauchy surfaces. There are slight variations in the literature in the definition of a Cauchy surface. Here we adopt the following definition.

Definition 1.9

A Cauchy surface for a spacetime M is an achronal subset S of M which is met by every inextendible causal curve in M.

Observations:

- If S is a Cauchy surface for M then $\partial I^+(S) = S$. (Also $\partial I^-(S) = S$.) It follows from Proposition 1.2 that a Cauchy surface S is a closed achronal C^0 hypersurface in M.
- If S is Cauchy then every inextendible timelike curve meets S exactly once.

Theorem 1.7 (Geroch)

If a spacetime M is globally hyperbolic then it has a Cauchy surface S.

We make some comments about the proof. (As discussed later, the converse also holds.)

- Introduce a measure μ on M such that $\mu(M) = 1$, and consider the function
  $f : M \to \mathbb{R}$ defined by
  $$f(p) = \frac{\mu(J^-(p))}{\mu(J^+(p))}\,.$$
- Internal compactness is used to show that f is continuous, and strong causality is used to show that f is strictly increasing along future directed causal curves.
- Moreover, if $\gamma : (a, b) \to M$ is a future directed inextendible causal curve in M, one shows $f(\gamma(t)) \to 0$ as $t \to a^+$, and $f(\gamma(t)) \to \infty$ as $t \to b^-$.
- It follows that 'slices' of f, e.g., $S = \{p \in M : f(p) = 1\}$, are Cauchy surfaces for M. $\square$

Remark: The function f constructed in the proof is what is referred to as a *time function*, namely, a continuous function that is strictly increasing along future directed causal curves.

Bernal and Sanchez [3] have shown how to construct *smooth time functions*, i.e., smooth functions with (past directed) timelike gradient, which hence are necessarily time functions. (See also Chruściel, Grant and Minguzzi [6] for related developments.)

Proposition 1.8

Let M be globally hyperbolic.

- *If S is a Cauchy surface for M then M is homeomorphic to $\mathbb{R} \times S$.*
- *Any two Cauchy surfaces in M are homeomorphic.*

Proof: To prove the first, one introduces a future directed timelike vector field X on M. Each integral curve of X meets S exactly once. These integral curves, suitably parameterized, provide the desired homeomorphism. A similar technique may be used to prove the second. $\square$

Remark: In view of Proposition 1.8, any nontrivial topology in a globally hyperbolic spacetime must reside in its Cauchy surfaces.

The following fact is often useful.

Proposition 1.9

If S is a compact achronal C^0 hypersurface in a globally hyperbolic spacetime M then S must be a Cauchy surface for M.

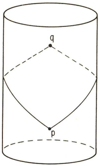
Proposition 1.9

If S is a compact achronal C^0 hypersurface in a globally hyperbolic spacetime M then S must be a Cauchy surface for M.

Comments on the proof:

- We have that $M = J^+(S) \cup J^-(S)$: $J^+(S) \cup J^-(S)$ is closed by Proposition 1.5, and is also easily shown to be open.
- Let γ be an inextendible causal curve. Suppose γ meets $J^+(S)$ at a point p. Then the portion of γ to the past of p must meet S, otherwise it is imprisoned in the compact set $J^+(S) \cap J^-(p)$.

<u>Ex.</u> $\partial I^+(p)$ in the flat spacetime cylinder closed in space.



Domains of Dependence

Definition 1.10

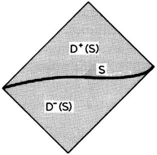
Let S be an achronal set in a spacetime M. The future domain of dependence of $D^+(S)$ of S is defined as follows,

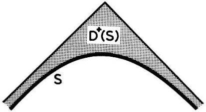
$$D^+(S) = \{p \in M : \text{ every past inextendible causal curve from } p \text{ meets } S\}$$

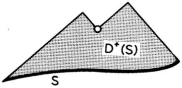
In physical terms, since information travels along causal curves, a point in $D^+(S)$ only receives information from S. Thus if physical laws are suitably causal, initial data on S should determine the physics on $D^+(S)$.

The past domain of dependence of $D^-(S)$ is defined in a time-dual manner. The (total) domain of dependence of S is the union, $D(S) = D^+(S) \cup D^-(S)$.

Below we show a few examples of future and past domains of dependence.







Note: $D^+(S) \supset S$.

The following characterizes Cauchy surfaces in terms of domain of dependence.

Proposition 1.10

Let S be an achronal subset of a spacetime M. Then, S is a Cauchy surface for M if and only if $D(S) = M$.

Proof: Exercise.

The following basic result ties domains of dependence to global hyperbolicity.

Proposition 1.11

Let $S \subset M$ be achronal.

(1) *Strong causality holds on* $\operatorname{int} D(S)$.

(2) *Internal compactness holds on* $\operatorname{int} D(S)$, *i.e., for all* $p, q \in \operatorname{int} D(S)$, $J^+(p) \cap J^-(q)$ *is compact.*

Proof: See Gal-ESI for a discussion of the proof. A few heuristic remarks:

(1): Suppose γ is a closed timelike curve through $p \in \operatorname{int} D^(S)$. By repeating loops we obtain an inextendible timelike which hence must meet S, infinitely often, in fact. But this would violate the achronality of S. More refined arguments show that strong causality holds on $\operatorname{int} D(S)$.

(2): A failure of internal compactness suggests the existence of a "hole" in $\operatorname{int} D(S)$. This, one expects, would lead to the existence of a past inextendible (resp. future inextendible) causal curve starting at $p \in \operatorname{int} D^+(S)$ (resp. $p \in \operatorname{int} D^-(S)$) that does not meet S.

By Proposition 1.11, for $S \subset M$ achronal, $\operatorname{int} D(S)$ is globally hyperbolic. Hence, we can now address the converse of Theorem 1.7.

Corollary 1.12

If S is a Cauchy surface for M then M is globally hyperbolic.

Proof: This follows immediately from Propositions 1.10 and 1.11: S Cauchy $\implies D(S) = M \implies \operatorname{int} D(S) = M \implies M$ is globally hyperbolic. $\square$

Thus we have that: M is globally hyperbolic if and only if M admits a Cauchy surface.

Cauchy horizons

We conclude this section with some comments about Cauchy horizons. If S is achronal, the *future Cauchy horizon* $H^+(S)$ of S is the future boundary of $D^+(S)$.

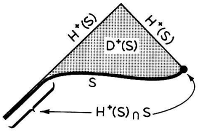
This is made precise in the following definition.

Definition 1.11

Let $S \subset M$ be achronal. The future Cauchy horizon $H^+(S)$ of S is defined as follows

$$\begin{aligned} H^+(S) &= \{p \in \overline{D^+(S)} : I^+(p) \cap D^+(S) = \emptyset\} \\ &= \overline{D^+(S)} \setminus I^-(D^+(S)) \,. \end{aligned}$$

The past Cauchy horizon $H^-(S)$ is defined time-dually. The (total) Cauchy horizon of S is defined as the union, $H(S) = H^+(S) \cup H^-(S)$.



We record some basic facts about domains of dependence and Cauchy horizons.

Proposition 1.13

Let S be an achronal subset of M. Then the following hold.

1. $H^+(S)$ *is achronal.*
2. $\partial D^+(S) = H^+(S) \cup S$.
3. $\partial D(S) = H(S)$.

Point 3 provides a useful mechanism for showing that an achronal set S is Cauchy: S is Cauchy iff $D(S) = M$ iff $\partial D(S) = \emptyset$ iff $H(S) = \emptyset$.

Cauchy horizons have structural properties similar to achronal boundaries, as indicated in the next two results.

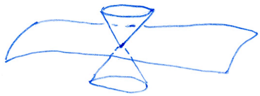
Proposition 1.14

Let $S \subset M$ be achronal. Then $H^+(S) \setminus \mathrm{edge}\, H^+(S)$, if nonempty, is an achronal C^0 hypersurface in M.

Proposition 1.15

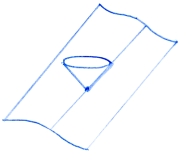
Let S be an achronal subset of M. Then $H^+(S)$ is ruled by null geodesics, i.e., every point of $H^+(S) \setminus \mathrm{edge}\, S$ is the future endpoint of a null geodesic in $H^+(S)$ which is either past inextendible in M or else has a past end point on $\mathrm{edge}\, S$.

In addition to curves, one can discuss the causality of certain higher dimensional submanifolds. For example, a *spacelike* hypersurface is a hypersurface all of whose tangent vectors are spacelike, or, equivalently, whose normal vectors are timelike:



In other words, a hypersurface is spacelike iff the induced metric is positive definite (i.e. Riemannian). In GR, a spacelike hypersurface represents space at a given instant of time.

A null hypersurface is a hypersurface such that the null cone is tangent to it at each of its points:



Null hypersurfaces play an important role in GR as they represent horizons of various sorts. Null hypersurfaces have an interesting geometry which we would like to discuss in this section.

Comments on Curvature and the Einstein Equations

- Let $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, $(X, Y) \to \nabla_X Y$, be the Levi-Civita connection with respect to the Lorentz metric g. ∇ is determined locally by the Christoffel symbols,

$$\nabla_{\partial_i}\partial_j = \sum_k \Gamma_{ij}^k \,\partial_k, \qquad (\partial_i = \frac{\partial}{\partial x^i}, etc.)$$

- *Geodesics* are curves $t \to \sigma(t)$ of zero acceleration,

$$\nabla_{\sigma'(t)}\sigma'(t) = 0\,.$$

  Timelike geodesics correspond to *free falling* observers.
- The Riemann curvature tensor is defined by,

$$R(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$$

  The components $R^\ell{}_{kij}$ are determined by,

$$R(\partial_i, \partial_j)\partial_k = \sum_\ell R^\ell{}_{kij}\partial_\ell$$

- The Ricci tensor Ric and scalar curvature R are obtained by taking traces,

$$R_{ij} = \sum_\ell R^\ell{}_{i\ell j} \quad \text{and} \quad R = \sum_{i,j} g^{ij} R_{ij}$$

The geometry of null hypersurfaces

- The Einstein equations, the field equations of GR, are given by:

$$R_{ij} - \frac{1}{2} R g_{ij} = 8\pi T_{ij} ,$$

  where T_{ij} is the energy-momentum tensor.

- From the PDE point of view, the Einstein equations form a system of second order quasi-linear equations for the g_{ij}'s. This system may be viewed as a (highly complicated!) generalization of Poisson's equation in Newtonian gravity.
- The *vacuum* Einstein equations are obtained by setting $T_{ij} = 0$. It is easily seen that this equivalent to setting $R_{ij} = 0$.

  We will sometimes require that a spacetime satisfying the Einstein equations, obeys an *energy condition*.

- The *null energy condition* (NEC) is the requirement that

$$T(X, X) = \sum_{i,j} T_{ij} X^i X^j \geq 0 \quad \text{for all null vectors } X .$$

- The stronger *dominant energy condtion* (DEC) is the requirement,

$$T(X, Y) = \sum_{i,j} T_{ij} X^i Y^j \geq 0 \quad \text{for all future directed causal vectors } X, Y .$$

Null Hypersurfaces

Definition 2.1

A null hypersurface in a spacetime (M, g) is a smooth co-dimension one submanifold S of M, such that at each $p \in S$, $g : T_pS \times T_pS \to \mathbb{R}$ is degenerate.

This means that there exists a nonzero vector $K_p \in T_pS$ (the direction of degeneracy) such that

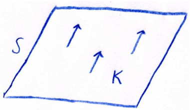
$$\langle K_p, X \rangle = 0 \quad \text{for all} \quad X \in T_pS \qquad\qquad (\langle \cdot, \cdot \rangle = g)$$

In particular,

- K_p is a null vector, $\langle K_p, K_p \rangle = 0$, which we can choose to be future pointing, and
- $[K_p]^\perp = T_pS$.
- Moreover, every (nonzero) vector $X \in T_pS$ that is not a multiple of K_p is spacelike.

Thus, every null hypersurface S gives rise to a smooth future directed null vector field K on S, unique up to a positive pointwise scale factor.

$$p \in S \xrightarrow{K} K_p \in T_pS,$$



Ex. $\mathbb{M}^{n+1}$ = Minkowski space.

- Null hyperplanes in $\mathbb{M}^{n+1}$: Each nonzero null vector $X \in T_p\mathbb{M}^{n+1}$ determines a null hyperplane $\Pi = \{q \in \mathbb{M}^{n+1} : \langle \overline{pq}, X\rangle = 0\}$.
- Null cones in $\mathbb{M}^{n+1}$: The past and future cones, $\partial I^-(p)$ and $\partial I^+(p)$, respectively, are smooth null hypersurfaces away from the vertex p.

The following fact is fundamental.

Proposition 2.1

Let S be a smooth null hypersurface and let K be a smooth future directed null vector field on S. Then the integral curves of K are null geodesics (when suitably parameterized),

Remark: The integral curves of K are called the *null generators* of S.

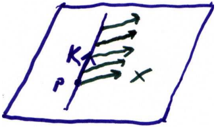
Proof: Suffices to show:

$$\nabla_K K = \lambda K$$

This follows by showing at each $p \in S$,

$$\nabla_K K \perp T_pS\,, \quad \text{i.e.,} \quad \langle \nabla_K K, X\rangle = 0 \quad \forall X \in T_pS$$

Extend $X \in T_pS$ by making it invariant under the flow generated by K,



$$[K, X] = \nabla_K X - \nabla_X K = 0$$

X remains tangent to S, so along the flow line through p,

$$\langle K, X \rangle = 0$$

Differentiating,

$$K\langle K, X \rangle = \langle \nabla_K K, X \rangle + \langle K, \nabla_K X \rangle = 0$$
$$\langle \nabla_K K, X \rangle = -\langle K, \nabla_X K \rangle = -\frac{1}{2} X\langle K, K \rangle = 0. \qquad \square$$

Remark: To study the 'shape' of the null hypersurface S we study how the null vector field K varies along S. Since K is actually orthogonal to S, this is somewhat analogous to how we study the shape of a hypersurface in a Riemannian manifold, or spacelike hypersurface in a Lorentzian manifold, by introducing the shape operator (or Weingarten map) and associated second fundamental form.

Null Weingarten Map/Null 2nd Fundamental Form

- We introduce the following equivalence relation on tangent vectors: For $X, Y \in T_pS$,

$$X = Y \bmod K \iff X - Y = \lambda K$$

Let $\overline{X}$ denote the equivalence class of $X \in T_pS$ and let,

$$T_pS/K = \{\overline{X} : X \in T_pS\}$$

Then,

$$TS/K = \cup_{p\in S} T_pS/K$$

is a rank $n-1$ vector bundle over S ($n = \dim S$). This vector bundle does not depend on the particular choice of null vector field K.

- There is a natural positive definite metric h on TS/K induced from $\langle\,,\rangle$: For each $p \in S$, define $h : T_pS/K \times T_pS/K \to \mathbb{R}$ by

$$h(\overline{X}, \overline{Y}) = \langle X, Y\rangle.$$

Well-defined: If $X' = X \bmod K$, $Y' = Y \bmod K$ then

$$\begin{aligned} \langle X', Y'\rangle &= \langle X + \alpha K, Y + \beta K\rangle \\ &= \langle X, Y\rangle + \beta\langle X, K\rangle + \alpha\langle K, Y\rangle + \alpha\beta\langle K, K\rangle \\ &= \langle X, Y\rangle\,. \end{aligned}$$

- The *null Weingarten map* $b = b_K$ of S with respect to K is, for each point $p \in S$, a linear map $b : T_pS/K \to T_pS/K$ defined by

$$b(\overline{X}) = \overline{\nabla_X K}\,.$$

  b is well-defined: $X' = X \mod K \Rightarrow$

$$\begin{aligned}\nabla_{X'} K &= \nabla_{X+\alpha K} K \\ &= \nabla_X K + \alpha \nabla_K K = \nabla_X K + \alpha\lambda K \\ &= \nabla_X K \quad \mod K\end{aligned}$$

- b is self adjoint with respect to h, i.e., $h(b(\overline{X}), \overline{Y}) = h(\overline{X}, b(\overline{Y}))$, for all $\overline{X}, \overline{Y} \in T_pS/K$.

  Proof: Extend $X, Y \in T_pS$ to vector fields tangent to S near p. Using $X\langle K, Y\rangle = 0$ and $Y\langle K, X\rangle = 0$, we obtain,

$$\begin{aligned}h(b(\overline{X}), \overline{Y}) &= h(\overline{\nabla_X K}, \overline{Y}) = \langle \nabla_X K, Y\rangle \\ &= -\langle K, \nabla_X Y\rangle = -\langle K, \nabla_Y X\rangle + \langle K, [X, Y]\rangle \\ &= \langle \nabla_Y K, X\rangle = h(\overline{X}, b(\overline{Y}))\,.\end{aligned}$$

- The *null second fundamental form* $B = B_K$ of S with respect to K is the bilinear form associated to b via h:

  For each $p \in S$, $B : T_pS/K \times T_pS/K \to \mathbb{R}$ is defined by,

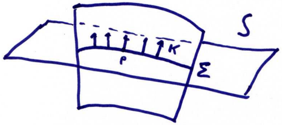
  $$B(\overline{X}, \overline{Y}) := h(b(\overline{X}), \overline{Y}) = h(\overline{\nabla_X K}, \overline{Y}) = \langle \nabla_X K, Y \rangle \,.$$

  Since b is self-adjoint, B is symmetric.

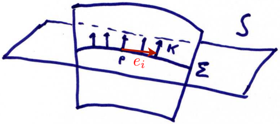
- The *null mean curvature* (or *null expansion scalar*) of S with respect to K is the smooth scalar field θ on S defined by,

  $$\theta = \mathrm{tr}\, b$$

  θ has a natural geometric interpretation. Let Σ be the intersection of S with a hypersurface in M which is transverse to K near $p \in S$; Σ will be a co-dimension two spacelike submanifold of M, along which K is orthogonal.



The geometry of null hypersurfaces

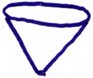


- Let $\{e_1, e_2, \cdots, e_{n-1}\}$ be an orthonormal basis for $T_p\Sigma$ in the induced metric. Then $\{\overline{e}_1, \overline{e}_2, \cdots, \overline{e}_{n-1}\}$ is an orthonormal basis for T_pS/K. Hence at p,

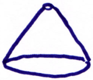
$$\begin{aligned}\theta = \mathrm{tr}\, b = \sum_{i=1}^{n-1} h(b(\overline{e}_i), \overline{e}_i) = \sum_{i=1}^{n-1} \langle \nabla_{e_i} K, e_i \rangle. \\ = \mathrm{div}_\Sigma K \,. \qquad (3)\end{aligned}$$

  where $\mathrm{div}_\Sigma K$ is the divergence of K along Σ.

- Thus, θ measures the overall expansion of the null generators of S towards the future.



$\theta > 0$



$\theta < 0$

- *Effect of scaling:* If $\widetilde{K} = fK$, $f \in C^\infty(S)$, is any other future directed null vector field on S, then $b_{\tilde{K}} = fb_K$, and hence, $\widetilde{\theta} = f\theta$.

  Hence, $b = b_K$ at a point p is uniquely determined by the value of K at p.

The geometry of null hypersurfaces

Comparison Theory

We now study how the null Weingarten map propagates along the null geodesic generators of S.

Let $\eta : I \to M$, $s \to \eta(s)$, be a future directed affinely parameterized null geodesic generator of S. For each $s \in I$, consider the Weingarten map $b = b(s)$ based at $\eta(s)$ with respect to the null vector field K which equals $\eta'(s)$ at $\eta(s)$,

$$b(s) = b_{\eta'(s)} : T_{\eta(s)}S/\eta'(s) \to T_{\eta(s)}S/\eta'(s)$$

Proposition 2.2

The one parameter family of Weingarten maps $s \to b(s)$, obeys the following Riccati equation,

$$b' + b^2 + R = 0\,, \qquad {}' = \nabla_{\eta'} \tag{4}$$

where $R : T_{\eta(s)}S/\eta'(s) \to T_{\eta(s)}S/\eta'(s)$ is given by $R(\overline{X}) = \overline{R(X, \eta'(s))\eta'(s)}$.

Remark on notation: In general, if $Y = Y(s)$ is a vector field along η tangent to S, we define, $(\overline{Y})' = \overline{Y'}$. Then, if $X = X(s)$ is a vector field along η tangent to S, b' is defined by,

$$b'(\overline{X}) := b(\overline{X})' - b(\overline{X'})\,. \tag{5}$$

Proof: Fix a point $p = \eta(s_0)$, $s_0 \in (a, b)$, on η. On a neighborhood U of p in S we can scale the null vector field K so that K is a geodesic vector field, $\nabla_K K = 0$, and so that K, restricted to η, is the velocity vector field to η, i.e., for each s near s_0, $K_{\eta(s)} = \eta'(s)$. Let $X \in T_pM$. Shrinking U if necessary, we can extend X to a smooth vector field on U so that $[X, K] = \nabla_X K - \nabla_K X = 0$. Then,

$$R(X, K)K = \nabla_X \nabla_K K - \nabla_K \nabla_X K - \nabla_{[X,K]} K = -\nabla_K \nabla_K X$$

Hence along η we have,

$$X'' = -R(X, \eta')\eta'$$

(which implies that X, restricted to η, is a Jacobi field along η).

Thus, from Equation (5), at the point p we have,

$$\begin{aligned} b'(\overline{X}) &= \overline{\nabla_X K}\,' - b(\overline{\nabla_K X}) = \overline{\nabla_K X}\,' - b(\overline{\nabla_X K}) \\ &= \overline{X''} - b(b(\overline{X})) = -\overline{R(X, \eta')\eta'} - b^2(\overline{X}) \\ &= -R(\overline{X}) - b^2(\overline{X}), \end{aligned}$$

which establishes Equation (4).

By taking the trace of (4) we obtain the following formula for the derivative of the null mean curvature $\theta = \theta(s)$ along η,

$$\theta' = -\mathrm{Ric}(\eta', \eta') - \sigma^2 - \frac{1}{n-1}\theta^2, \tag{6}$$

where $\sigma := (\mathrm{tr}\, \hat{b}^2)^{1/2}$ is the *shear scalar*, $\hat{b} := b - \frac{1}{n-1}\theta \cdot \mathrm{id}$ is the trace free part of the Weingarten map, and $\mathrm{Ric}(\eta', \eta')$ is the spacetime Ricci tensor evaluated on the tangent vector η'.

Equation 6 is known in relativity as the Raychaudhuri equation (for an irrotational null geodesic congruence). This equation shows how the Ricci curvature of spacetime influences the null mean curvature of a null hypersurface.

We consider a basic application of the Raychaudhuri equation.

Proposition 2.3

Let M be a spacetime which obeys the null enery condition (NEC), $\mathrm{Ric}\,(X, X) \geq 0$ for all null vectors X, and let S be a smooth null hypersurface in M. If the null generators of S are future geodesically complete then S has nonnegative null expansion, $\theta \geq 0$.

Proof: Suppose $\theta < 0$ at $p \in S$. Let $s \to \eta(s)$ be the null generator of S passing through $p = \eta(0)$, affinely parametrized. Let $b(s) = b_{\eta'(s)}$, and take $\theta = \text{tr } b$. By the invariance of sign under scaling, one has $\theta(0) < 0$.

Raychaudhuri's equation and the NEC imply that $\theta = \theta(s)$ obeys the inequality,

$$\frac{d\theta}{ds} \leq -\frac{1}{n-1}\theta^2 \,,$$

and hence $\theta < 0$ for all $s > 0$. Dividing through by θ^2 then gives,

$$\frac{d}{ds}\left(\frac{1}{\theta}\right) \geq \frac{1}{n-1} \,,$$

which implies $1/\theta \to 0$, i.e., $\theta \to -\infty$ in finite affine parameter time, contradicting the smoothness of θ. $\square$

Remark. Let Σ be a local cross section of the null hypersurface S (see earlier figure) with volume form ω. If Σ is moved under flow generated by K then $L_K\omega = \theta\,\omega$, where $L =$ Lie derivative.

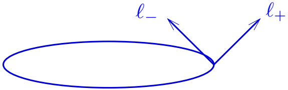
Thus, Proposition 2.3 implies, under the given assumptions, that cross sections of S are nondecreasing in area as one moves towards the future. Proposition 2.3 is the simplest form of Hawking's black hole area theorem [16]. For a study of the area theorem, with a focus on issues of regularity, see [6].

In this section we introduce the important notion of a trapped surface and present the classical Penrose singularity theorem.

- Let (M^{n+1}, g) be an $(n+1)$-dimensional spacetime, $n \geq 3$.
  Let Σ^{n-1} be a closed (i.e., compact without boundary) co-dimension two spacelike submanifold of M.

- Each normal space of Σ, $[T_p\Sigma]^\perp$, $p \in \Sigma$, is timelike and 2-dimensional, and hence admits two future directed null directions orthogonal to Σ.

  Thus, under suitable orientation assumptions, Σ admits two smooth nonvanishing future directed null normal vector fields ℓ_+ and ℓ_- (unique up to positive rescaling).

  By convention, we refer to ℓ_+ as outward pointing and ℓ_- as inward pointing.

- Associated to ℓ_+ and ℓ_-, are the two *null second fundamental forms*, χ_+ and χ_-, respectively, defined as

$$\chi_\pm : T_p\Sigma \times T_p\Sigma \to \mathbb{R}, \qquad \chi_\pm(X, Y) = g(\nabla_X \ell_\pm, Y)\,.$$

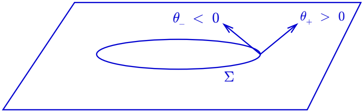
- *The null expansion scalars (or null mean curvatures)* $\theta_\pm$ of Σ are obtained by tracing $\chi_\pm$ with respect to the induced metric γ on Σ,

$$\theta_\pm = \mathrm{tr}_\gamma \chi_\pm = \gamma^{AB} \chi_{\pm AB} = \mathrm{div}\,_\Sigma \ell_\pm\,.$$

  The sign of $\theta_\pm$ does not depend on the scaling of $\ell_\pm$. Physically, θ_+ (resp., θ_-) measures the divergence of the outgoing (resp., ingoing) light rays emanating orthogonally from Σ.

  Remark: There is a natural connection between these null expansion scalars $\theta_\pm$ and the null expansion of null hypersurfaces: ℓ_+ locally generates a smooth null hypersurface S_+. Then θ_+ is the null expansion of S_+ restricted to Σ; θ_- may be described similarly.

- For round spheres in Euclidean slices in Minkowski space (and, more generally, large "radial" spheres in AF spacelike hypersurfaces),



- However, in regions of spacetime where the gravitational field is strong, one can have both

$$\theta_- < 0 \quad \text{and} \quad \theta_+ < 0\,,$$

in which case Σ is called a **trapped surface**. (See e.g. [5, 19] for results concerning the dynamical formation of trapped surfaces.)

- As we now discuss, assuming appropriate energy and causality conditions, if a trapped surface forms, then the development of singularities is inevitable.

Theorem 3.1 (Penrose singularity theorem)

Let M be a globally hyperbolic spacetime which satisfies the NEC, $\mathrm{Ric}(X,X) \geq 0$ for all null vectors X, and which has a noncompact Cauchy surface S. If M contains a trapped surface Σ then M is future null geodesically incomplete.

Proof: We first observe the following.

Claim: $\partial I^+(\Sigma)$ is noncompact.

Proof of Claim: $\partial I^+(\Sigma)$ is an achronal boundary, and hence, by Proposition 1.2, is an achronal C^0 hypersurface. If $\partial I^+(\Sigma)$ were compact then, by Proposition 1.9, $\partial I^+(\Sigma)$ would be a compact Cauchy surface. But this would contradict the assumption that S is noncompact (all Cauchy surfaces are homeomorphic). □

We now construct a future inextendible null geodesic in $\partial I^+(\Sigma)$, which we show must be future incomplete.

- We have that

$$\partial I^+(S) = \overline{I^+(S)} \setminus \operatorname{int} I^+(S) = J^+(S) \setminus I^+(S)\,.$$

  It then follows from Proposition 1.1 that each $q \in \partial I^+(\Sigma)$ lies on a null geodesic in $\partial I^+(\Sigma)$ with past end point on Σ. Moreover this null geodesic meets Σ orthogonally (due to achronality, cf. O'Neill [22, Lemma 50, p. 298]).

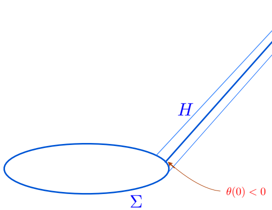
- Since $\partial I^+(\Sigma)$ is closed and noncompact, there exists a sequence of points $\{q_k\} \subset \partial I^+(\Sigma)$ that diverges to infinity. For each k, there is a null geodesic η_k from Σ to q_k, which is contained in $\partial I^+(\Sigma)$ and meets Σ orthogonally.

- By compactness of Σ, some subsequence η_{k_j} converges to a future inextendible null geodesic η contained in $\partial I^+(\Sigma)$, and meeting Σ orthogonally (at p, say).

  η must be future incomplete. Suppose not.

- By achronality of $\partial I^+(\Sigma)$,
  - No other future directed null normal geodesic starting on Σ can meet η.
  - There can be no *null focal point* to Σ along η (cf. O'Neill, Prop. 48, p. 296).

- It follows that η is contained in a smooth (perhaps very thin) null hypersurface $H \subset \partial I^+(\Sigma)$.



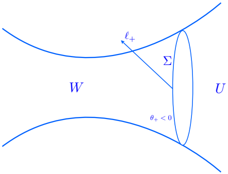
- Let θ be the null expansion of H along η. Since Σ is a trapped surface $\theta(p) < 0$. Arguing just as in the "area theorem" (Proposition 2.3), using Raychaudhuri + NEC, θ must go to $-\infty$ in finite affine parameter time $\rightarrow\leftarrow$. Hence η must be future incomplete. $\square$

Topics

For certain applications, the following variant of the Penrose singularity theorem is useful.

Theorem 3.2

Let M be a globally hyperbolic spacetime satisfying the null energy condition, with smooth spacelike Cauchy surface V. Let Σ be a smooth closed (compact without boundary) hypersurface in V which separates V into an "inside" U and an "outside" W, i.e., $V \setminus \Sigma = U \cup W$ where $U, W \subset V$ are connected disjoint sets. Suppose, further, that $\overline{W}$ is non-compact. If Σ is outer-trapped ($\theta_+ < 0$) then M is future null geodesically incomplete.



Proof: Consider the achronal boundary $\partial I^+(\overline{U})$, and argue similarly to the proof of the Penrose singularity theorem that if M is future null geodesically complete then $\partial I^+(\overline{U})$ is compact. But, by considering the integral curves of a timelike vector field X on M, one can establish a homeomorphism between $\partial I^+(\overline{U})$ and $\overline{W}$ $\rightarrow\leftarrow$.

This version of the Penrose singularity theorem may be used to prove the following beautiful result of Gannon [15] and Lee [20].

Theorem 3.3

Let M be a globally hyperbolic spacetime which satisfies the null energy condition and which contains a smooth asymptotically flat spacelike Cauchy surface V. If V is not simply connected ($\pi_1(V) \neq 0$) then M is future null geodesically incomplete.

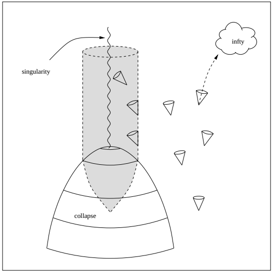
Comment on the proof. Pass to the universal covering spacetime $\overline{M}$. It will contain a Cauchy surface $\overline{V}$, which is the universal cover of V. If $\pi_1(V) \neq 0$ then $\overline{V}$ will have more than one AF end. Now apply Theorem 3.2 to $\overline{M}$ with Cauchy surface $\overline{V}$.

As suggested by Theorem 3.3, nontrivial topology tends to induce gravitational collapse. This in turn leads to the notion of topological censorship. We return to these issues in Section 5.

Introduction

Black holes are certainly one of the most remarkable predictions of General Relativity.

The following depicts the process of gravitational collapse and formation of a black hole.



A stellar object, after its fuel is spent, begins to collapse under its own weight. As the gravitational field intensifies the light cones bend "inward" (so to speak).

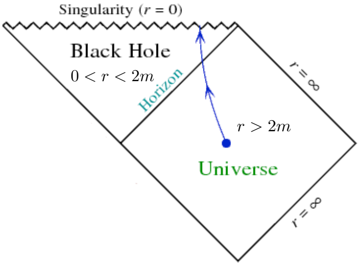
The shaded region is the black hole region. The boundary of this region is the black hole event horizon. It is the boundary between points that can send signals to infinity and points that can't.

_Ex._ *The Schwarzshild solution (1916).* Static (time-independent, nonrotating) spherically symmetric, vacuum solution to the Einstein equations.

$$g = -\left(1 - \frac{2m}{r}\right) dt^2 + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2 d\Omega^2$$

This metric represents the region outside a (collapsing) spherically symmetric star.

The region $0 < r < 2m$ is the black hole region; $r = 2m$ corresponds to the event horizon.



Ex. The Kerr solution (1963). Stationary (time-independent, rotating), axisymmetric, vacuum solution.

$$\begin{aligned} ds^2 &= -\left[1 - \frac{2mr}{r^2 + a^2\cos^2\theta}\right] dt^2 - \frac{4mra\sin^2\theta}{r^2 + a^2\cos^2\theta}\, dt\, d\phi \\ &\quad + \left[\frac{r^2 + a^2\cos^2\theta}{r^2 - 2mr + a^2}\right] dr^2 + (r^2 + a^2\cos^2\theta)\, d\theta^2 \\ &\quad + \left[r^2 + a^2 + \frac{2mra^2\sin^2\theta}{r^2 + a^2\cos^2\theta}\right] \sin^2\theta\, d\phi^2. \end{aligned}$$

The Kerr solution is determined by two parameters: mass parameter m and angular momentum parameter a. When $a = 0$, the Kerr solution reduces to the Schwarzschild solution. The Kerr solution contains an event horizon (provided $a < m$), and hence represents a steady state rotating black hole.

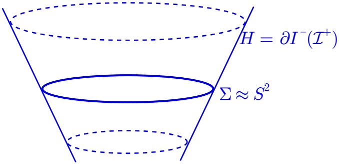
It is a widely held belief that "true" astrophysical black holes "settle down" to a Kerr solution. This belief is based largely on results ("no hair theorems") that establish the uniqueness of Kerr among all asymptotically flat stationary, solutions to the vacuum Einstein equations. (The proof assumes analyticity, but there has been recent progress in removing this assumption; see the excellent recent review article of Ionescu and Klainerman [18].)

Marginally outer trapped surfaces and the topology of black holes

A basic step in the proof of the uniqueness of the Kerr solution is Hawking's theorem on the topology of black holes in $3+1$ dimensions.

Theorem 4.1 (Hawking's black hole topology theorem)

Suppose (M, g) is a $(3+1)$-dimensional asymptotically flat stationary black hole spacetime obeying the dominant energy condition. Then cross sections Σ of the event horizon are topologically 2-spheres.



Comment on the proof: Hawking's proof is variational in nature. Using the dominant energy condition and the Gauss-Bonnet theorem, he shows that if Σ has genus ≥ 1 then Σ can be deformed outward to an *outer trapped surface*. However, there can be no outer trapped surface outside the event horizon. Such a surface would be visible from 'null infinity', but there are arguments precluding that possibility [23, 7].

Higher Dimensional Black Holes

- String theory, and various related developments (e.g., the AdS/CFT correspondence, braneworld scenarios, entropy calculations) have generated a great deal of interest in gravity in higher dimensions, and in particular, in higher dimensional black holes.
- One of the first questions to arise was:

  Does black hole uniqueness hold in higher dimensions?

- With impetus coming from the development of string theory, in 1986, Myers and Perry [21] constructed natural higher dimensional generalizations of the Kerr solution. These models painted a picture consistent with the situation in $3 + 1$ dimensions. In particular, they have spherical horizon topology.

- But in 2002, Emparan and Reall [9] discovered a remarkable example of a $4+1$ dimensional AF stationary vacuum black hole spacetime with horizon topology $S^2 \times S^1$ (the *black ring*).

VOLUME 88, NUMBER 10 PHYSICAL REVIEW LETTERS 11 MARCH 2002

A Rotating Black Ring Solution in Five Dimensions

Roberto Emparan[1,*] and Harvey S. Reall[2]
[1] *Theory Division, CERN, CH-1211 Geneva 23, Switzerland*
[2] *Physics Department, Queen Mary College, Mile End Road, London E1 4NS, United Kingdom*
(Received 8 November 2001; published 21 February 2002)

Thus in higher dimensions, black hole uniqueness does not hold and horizon topology need not be spherical.

This caused a great surge of activity in the study of higher dimensional black holes.

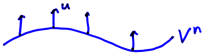
- *Question:* What horizon topologies are allowed in higher dimensions? What restrictions are there?

Marginally Outer Trapped Surfaces

Want to describe a generalization of Hawking's theorem to higher dimensions. This will be based on properties of *marginally outer trapped surfaces*.

- By an *initial data set* in a spacetime (M^{n+1}, g) we mean a triple (V^n, h, K), where V^n is a spacelike hypersurface, h is the induced metric and K is the second fundamental form: For vectors $X, Y \in T_pV$,

$$K(X, Y) = h(\nabla_X u, Y)$$



- Given an initial data set (V, h, K), the spacetime dominant energy condition implies,

$$\mu \geq |J|$$

  along V, where $\mu =$ *local energy density* $= G(u, u)$, and $J =$ *local momentum density* $=$ 1-form $G(u, \cdot)$ on V, where $G = Ric - \frac{1}{2}Rg$.
- Using the Gauss-Codazzi equations, μ and J can be expressed solely in terms of the initial data

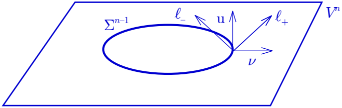
$$\mu = \frac{1}{2}\left(S + (\mathrm{tr}\, K)^2 - |K|^2\right),$$
$$J = \mathrm{div} K - d(\mathrm{tr}\, K).$$

  These are the *Einstein constraint* equations. In the *time-symmetric* case $(K = 0)$, the DEC reduces to $S \geq 0$.

Marginally outer trapped surfaces and the topology of black holes

Consider an initial data set (V^n, h, K) in a spacetime (M^{n+1}, g), $n \geq 3$.

Let Σ^{n-1} be a closed 2-sided hypersurface in V^n. Σ admits a smooth unit normal field ν in V.



$$\ell_+ = u + \nu \qquad \text{f.d. outward null normal}$$
$$\ell_- = u - \nu \qquad \text{f.d. inward null normal}$$

Null second fundamental forms: χ_+, χ_-

$$\chi_\pm(X, Y) = g(\nabla_X \ell_\pm, Y) \qquad X, Y \in T_p\Sigma$$

Null expansion scalars: θ_+, θ_-

$$\theta_\pm = \mathrm{tr}_\gamma \chi_\pm = \gamma^{AB}(\chi_\pm)_{AB} = \mathrm{div}_\Sigma \ell_\pm$$

In terms of initial data (V^n, h, K),

$$\theta_\pm = \mathrm{tr}_\Sigma K \pm H$$

where H = mean curvature of Σ within V. (*Note:* In the 'time-symmetric' case $(K = 0)$, $\theta_+ = H$.)

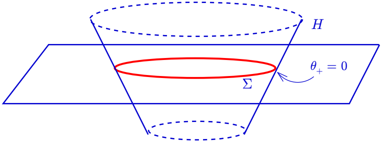
Recall, Σ is a trapped surface if both θ_+ and θ_- are negative. Focusing attention on the outward null normal only:

- If $\theta_+ < 0$ - we say Σ is outer trapped.
- If $\theta_+ = 0$ - we say Σ is a marginally outer trapped surface (MOTS).

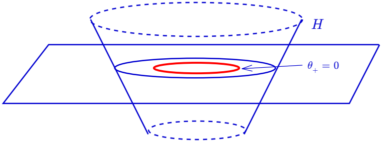
(*Note:* In the time symmetric case a MOTS is simply a minimal surface.)

MOTSs arise naturally in several situations.

- In stationary black hole spacetimes - cross sections of the event horizon are MOTS.



- In dynamical black hole spacetimes - MOTS typically occur inside the event horizon:



Stability of MOTS

MOTSs admit a notion of stability based on variations of the null expansion (Andersson, Mars and Simon (2005, 2008)).

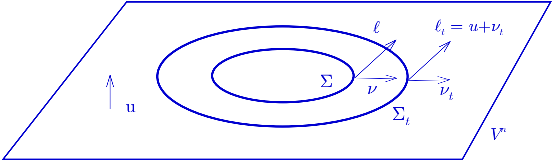
Let Σ be a MOTS in an initial data set (V, h, K) with outward normal ν. Consider normal variations of Σ in V, i.e., variations $t \to \Sigma_t$ of $\Sigma = \Sigma_0$ with variation vector field

$$\mathcal{V} = \left.\frac{\partial}{\partial t}\right|_{t=0} = \phi\nu, \quad \phi \in C^\infty(\Sigma)\,.$$

Let

$$\theta(t) = \text{ the null expansion of } \Sigma_t,$$

with respect to $\ell_t = u + \nu_t$, where ν_t is the unit normal field to Σ_t in V.



A computation shows,

$$\left.\frac{\partial\theta}{\partial t}\right|_{t=0} = L(\phi)\,,$$

where $L : C^\infty(\Sigma) \to C^\infty(\Sigma)$ is given by,

$$L(\phi) = -\triangle\phi + 2\langle X, \nabla\phi\rangle + \left(Q + \operatorname{div} X - |X|^2\right)\phi\,, \tag{7}$$

$$Q = \frac{1}{2}S_\Sigma - (\mu + J(\nu)) - \frac{1}{2}|\chi|^2\,, \tag{8}$$

In the time-symmetric case ($K = 0$), $\theta = H$, $X = 0$, and L reduces to the classical stability operator of minimal surface theory.

In analogy with minimal surface case, we refer to L as the *MOTS stability operator*. Note, however, that L is *not* in general self-adjoint.

Lemma 4.2 (Andersson, Mars, and Simon)

Among eigenvalues with smallest real part, there is a real eigenvalue $\lambda_1(L)$, called the principal eigenvalue. The associated eigenfunction ϕ, $L(\phi) = \lambda_1\phi$, is unique up to a multipicative constant, and can be chosen to be strictly positive.

In analogy with the minimal surface case, *we say that a MOTS Σ is* stable *provided $\lambda_1(L) \geq 0$.*

Heuristically, a stable MOTS is *infinitesimally* outermost.

- There is a basic criterion for a MOTS to be stable. We say a MOTS Σ is *weakly outermost* provided there are no outer trapped ($\theta < 0$) surfaces outside of, and homologous, to Σ.

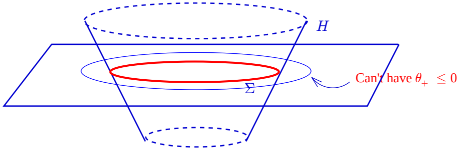
  Fact: A weakly outermost MOTS is stable.

  Proof: Suppose to the contrary, $\lambda_1 < 0$. Consider the variation $t \to \Sigma_t$ of Σ with variation vector field $\mathcal{V} = \phi\nu$, where ϕ is a positive eigenfunction associated to $\lambda_1 = \lambda_1(L)$. Then,

  $$\left.\frac{\partial\theta}{\partial t}\right|_{t=0} = L(\phi) = \lambda_1\phi < 0$$

  Since $\theta(0) = 0$, this implies $\theta(t) < 0$ for small $t > 0$, and hence there are outer trapped surfaces outside of, and homologous to, Σ.

- **Fact:** Cross sections of the event horizon in AF stationary black hole spacetimes obeying the DEC are (weakly) outermost MOTSs.



- More generally, (weakly) outermost MOTSs can arise as the boundary of the "trapped region" (Andersson and Metzger, Eichmair).

A Generalization of Hawking's Black Hole Topology Theorem

Theorem 4.3 (G. and Schoen [14])

Let (V, h, K) be an n-dimensional initial data set, $n \geq 3$, satisfying the dominant energy condition (DEC), $\mu \geq |J|$. If Σ is a stable MOTS in V then (apart from certain exceptional circumstances) Σ must be of positive Yamabe type, i.e. must admit a metric of positive scalar curvature.

Remarks.

- The theorem may be viewed as a spacetime analogue of a fundamental result of Schoen and Yau concerning stable minimal hypersurfaces in manifolds of positive scalar curvature.
- The 'exceptional circumstances' are ruled out if, for example, the DEC holds strictly at some point of Σ or Σ is not Ricci flat.
- Σ being of positive Yamabe type implies many well-known restrictions on the topology (see e.g.; see e.g. [17, Chapter 7]).

We consider here two basic examples, and for simplicity we assume Σ is orientable.

Case 1. $\dim \Sigma = 2$ ($\dim M = 3 + 1$): In this case, Σ being of positive Yamabe type means that Σ admits a metric of positive Gaussian curvature. Hence, by the Gauss-Bonnet theorem, Σ is topologically a 2-sphere, and we recover Hawking's theorem.

Case 2. $\dim \Sigma = 3$ ($\dim M = 4 + 1$).

Theorem. (Gromov-Lawson, Schoen-Yau) If Σ is a closed orientable 3-manifold of positive Yamabe type then Σ must be diffeomorphic to:

- a spherical space, or
- $S^2 \times S^1$, or
- a connected sum of the above two types.

Remark: By the prime decomposition theorem, Σ must be a connected sum of (1) spherical spaces, (2) $S^2 \times S^1$'s, and (3) $K(\pi, 1)$ manifolds. But since Σ is positive Yamabe, it cannot contain any $K(\pi, 1)$'s in its prime decomposition.

Thus, the basic horizon topologies in $\dim \Sigma = 3$ case are S^3 and $S^2 \times S^1$.

(Kundari and Lucietti have recently constructed an AF black hole spacetime satisfying the DEC with $\mathbb{RP}^3$ horizon topology, arXiv 2014.)

Proof of Theorem 4.3: Need to produce a metric of positive scalar curvature on Σ.

- Formally setting $X = 0$ in the MOTS stability operator L, we obtain the "symmetrized" operator
$$L_0(\phi) = -\triangle\phi + Q\,\phi\,.$$
where $Q = \frac{1}{2}S_\Sigma - (\mu + J(\nu)) - \frac{1}{2}|\chi|^2$.
- **Key Fact:** $\lambda_1(L_0) \geq \lambda_1(L)$.
- Since, by assumption, $\lambda_1(L) \geq 0$, we have $\lambda_1(L_0) \geq 0$. (This, in particular, implies the *MOTS stability inequality*,
$$\int_\Sigma |\nabla\psi|^2 + Q\,\psi^2 dA \geq 0\,, \quad \forall\psi \in C^\infty(\Sigma)$$
In essence, the situation has been reduced to the Riemannian case.)
- Now make the conformal change: $\tilde{\gamma} = \phi^{\frac{2}{n-2}}\gamma$, where ϕ is a positive eigenfunction corresponding to $\lambda_1(L_0)$:
$$L_0(\phi) = -\triangle\phi + Q\phi = \lambda_1(L_0)\phi\,.$$
A standard computation shows,
$$\tilde{S}_\Sigma = \phi^{-\frac{n}{n-2}}(-2\triangle\phi + S_\Sigma\phi + \frac{n-1}{n-2}\frac{|\nabla\phi|^2}{\phi})$$
$$= \phi^{-\frac{2}{n-2}}(2\lambda_1(L_0) + 2(\mu + J(\nu)) + |\chi|^2 + \frac{n-1}{n-2}\frac{|\nabla\phi|^2}{\phi^2}) \geq 0$$

Proof of the Key Fact. Let ϕ be a positive eigenfunction associated to $\lambda_1(L)$, $L(\phi) = \lambda_1(L)\phi$. Completing the square on the right hand side of (7)) gives,

$$-\triangle\phi + (Q + \mathrm{div}\, X)\,\phi + \phi|\nabla \ln \phi|^2 - \phi|X - \nabla \ln \phi|^2 = \lambda_1(L)\phi$$

where $Q := \frac{1}{2}S - (\mu + J(\nu)) - \frac{1}{2}|\chi|^2$. Letting $u = \ln \phi$, we obtain,

$$-\triangle u + Q + \mathrm{div}\, X - |X - \nabla u|^2 = \lambda_1(L)\,.$$

Absorbing the Laplacian term $\triangle u = \mathrm{div}\,(\nabla u)$ into the divergence term gives,

$$Q + \mathrm{div}\,(X - \nabla u) - |X - \nabla u|^2 = \lambda_1(L).$$

Setting $Y = X - \nabla u$, we arrive at,

$$-Q + |Y|^2 + \lambda_1(L) = \mathrm{div}\, Y\,.$$

Given any $\psi \in C^\infty(\Sigma)$, we multiply through by ψ^2 and derive,

$$\begin{aligned}
-\psi^2 Q + \psi^2|Y|^2 + \psi^2\lambda_1(L) &= \psi^2 \mathrm{div}\, Y \\
&= \mathrm{div}\,(\psi^2 Y) - 2\psi\langle\nabla\psi, Y\rangle \\
&\leq \mathrm{div}\,(\psi^2 Y) + 2|\psi||\nabla\psi||Y| \\
&\leq \mathrm{div}\,(\psi^2 Y) + |\nabla\psi|^2 + \psi^2|Y|^2\,.
\end{aligned}$$

Integrating the above inequality yields,

$$\lambda_1(L) \leq \frac{\int_\Sigma |\nabla\psi|^2 + Q\psi^2}{\int_\Sigma \psi^2} \quad \text{for all } \psi \in C^\infty(\Sigma),\ \psi \not\equiv 0\,.$$

$$\implies \lambda_1(L) \leq \lambda_1(L_0).$$

One drawback of Theorem 4.3 is that it allows certain possibilities that one would like to rule out. E.g., the theorem does not rule out the possibility of a vacuum black hole spacetime with toroidal horizon topology. In a subsequent paper such possibilities were ruled out.

Theorem 4.4 (G. [12])

Let (V, h, K) be an n-dimensional, $n \geq 3$, initial data set in a spacetime obeying the DEC. Suppose Σ is a weakly outermost MOTS in V. If Σ is not of positive Yamabe type then an outer neighborhood $U \approx [0, \epsilon) \times \Sigma$ of Σ is foliated by MOTS.

Thus, if Σ is *outermost* (i.e, if there are no outer trapped, or marginally outer trapped, surfaces outside of and homologous to Σ) then Σ must be of positive Yamabe type *without exception*.

This rigidity result belongs to a *family* of related rigidity results; see e.g. [13] and references therein.

Topics

The geometry and topology of the exterior region

In principle, the topology of space could be quite complicated. In fact GR does not put any restrictions on the topology of space.

Theorem 5.1 (Isenberg-Mazzeo-Pollack (2003))

Let N be any compact n-dimensional manifold, and $p \in N$ arbitrary. Then $N \setminus \{p\}$ admits an AF initial data set satisfying the vacuum constraint equations.

However, according to the principle of *topological censorship*, the topology of the region outside of all black holes (and white holes) should be simple. The rationale is roughly as follows:

- The Theorem of Gannon [15] and Lee [20] discussed in Section 3 suggests that nontrivial topology induces gravitational collapse.
- *Weak cosmic censorship conjecture:* In the standard collapse scenario, the process of gravitational collapse leads to the formation of an event horizon which shields the singularities from view.
- *Topological censorship:* The nontrivial topology that induced collapse should end up behind hidden the event horizon, and the region outside the black hole should have simple topology.

The geometry and topology of the exterior region

This circle of ideas was formalized by the Topological Censorship Theorem of Friedman, Schleich and Witt [10], which says, in physical terms, that observers who remain outside the event horizon are unable to probe nontrivial topology.

Their theorem applies to asymptotically flat spacetimes, i.e. spacetimes that have an asymptotic structure similar to that of Minkowski space in the sense of admitting a regular past and future null infinity $\mathscr{I}^-$ and $\mathscr{I}^+$, respectively; cf. Wald [23].

$$\mathrm{DOC} = I^-(\mathscr{I}^+) \cap I^+(\mathscr{I}^-)$$

In [11] the following strengthened version was obtained:

Theorem 5.2

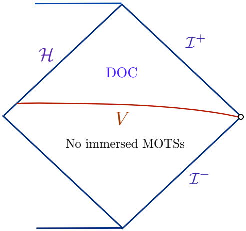
Let M be an asymptotically flat spacetime, and suppose that the domain of outer communications (the region outside of all black holes and white holes) is globally hyperbolic and satisfies the NEC. Then the DOC is simply connected.

These and other results supporting the notion of topological censorship are *spacetime* results; they involve assumptions that are essentially global in time. It is a natural but difficult question to determine whether a given initial data set will give rise to a spacetime that satisfies these conditions.

The geometry and topology of the exterior region

The aim of more recent work with Michael Eichmair and Dan Pollack was to obtain a result supportive of the principle of topological censorship *at the pure initial data level*, thereby circumventing these difficult issues of global evolution.

Consider the following schematic Penrose-type diagram of a black hole spacetime.



- Should think of the initial data manifold V as representing an asymptotically flat spacelike slice in the domain of outer communications (DOC) whose boundary ∂V corresponds to a cross section of the event horizon.
- At the initial data level, we represent this cross section by a MOTS.
- We assume there are no *immersed* MOTSs in the 'exterior region' $V \setminus \partial V$.

The geometry and topology of the exterior region

Theorem 5.3 (Eichmair, G., Pollack [8])

Let (V, h, K) be a 3-dimensional asymptotically flat (AF) initial data set such that V is a manifold-with-boundary, whose boundary ∂V is a compact MOTS. If there are no (immersed) MOTS in $V \setminus \partial V$, then V is diffeomorphic to $\mathbb{R}^3$ minus a finite number of open balls.

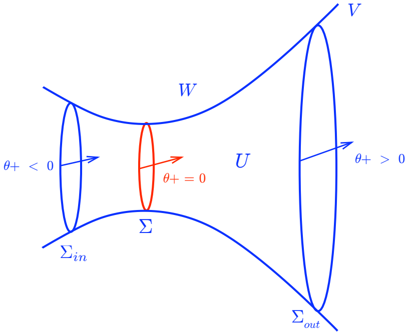
AF: V outside a compact set is diffeomorphic to $\mathbb{R}^3 \setminus B_1(\mathbf{0})$, such that $h \to \delta$ and $K \to 0$ at suitable rates as $r \to \infty$.

Remarks

- The proof makes use of powerful existence results for MOTSs (Schoen, Andersson and Metzger, Eichmair).
- The proof also makes use of an important consequence of geometrization, namely that the fundamental group of every compact 3-manifold is residually finite.
- Dominant energy condition not required!
- MOTSs detect nontrivial topology (reminiscent of how minimal surfaces have been used in Riemannian geometry to detect nontrivial topology, cf., Meeks-Simon-Yau).
- In the 'no horizon' case ($\partial V = \emptyset$) the conclusion is that V is diffeomorphic to $\mathbb{R}^3$.

Theorem 5.4 (Schoen, Andersson and Metzger, Eichmair)

Let W^n be a connected compact manifold-with-boundary in an initial data set (V^n, h, K), $3 \leq n \leq 7$. Suppose, $\partial W = \Sigma_{in} \cup \Sigma_{out}$, such that Σ_{in} is outer trapped ($\theta_+ < 0$) and Σ_{out} is outer untrapped ($\theta_+ > 0$). Then there exists a smooth compact MOTS in W that separates Σ_{in} from Σ_{out}.

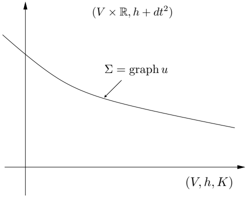


Remarks on the proof of existence:

- In their proof of the positive mass theorem in the general (non-time symmetric case), Schoen and Yau studied in detail the existence and regularity properties of Jang's equation, which they interpreted as a prescribed mean curvature equation of sorts,

$$H_\Sigma + \mathrm{tr}_{\Sigma} \bar{K} = 0\,,$$

where H_Σ = the mean curvature of $\Sigma = \mathrm{graph}\, u$ in $V \times \mathbb{R}$.



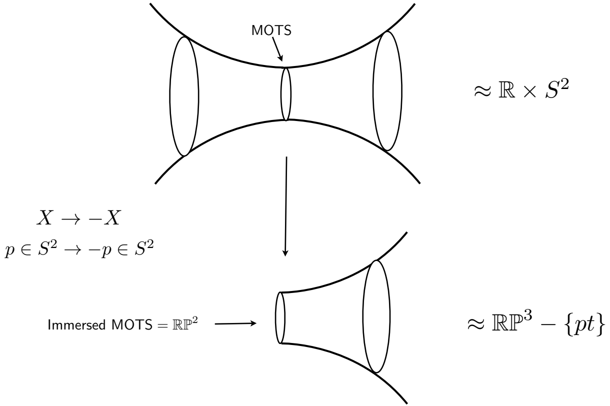
- They showed that the only obstruction to global existence is the presence of MOTSs in the initial data.
- As Schoen described at a conference in Miami, one can turn this 'drawback' into a 'feature': One can establish existence by inducing blow-up of Jang's equation.
- Complete proofs were given by Andersson and Metzger (dim $= 3$), Eichmair (dim ≥ 3); cf. the excellent survey article by Andersson, Eichmair, Metzger [2].

Immersed MOTS

Definition 5.1

A subset $\Sigma \subset V$ is an immersed MOTS in an initial data set (V, h, K) if there exists a finite cover $\tilde{V}$ of V with covering map $p : \tilde{V} \to V$ and a MOTS $\tilde{\Sigma}$ in $(\tilde{V}, p^* h, p^* K)$ such that $p(\tilde{\Sigma}) = \Sigma$.

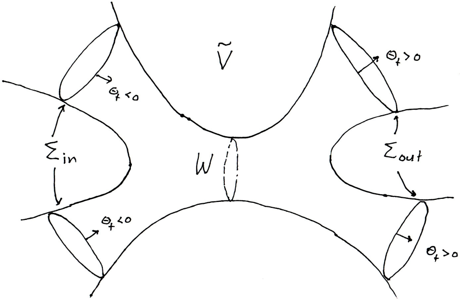
A simple example is given by the so called $\mathbb{RP}^3$ geon. Consider the $T = 0$ slice in the extended Schwarzschild spacetime (Flamm paraboloid):



Lemma 5.5

Let (V^n, h, K) be an n-dimensional, $3 \le n \le 7$, AF initial data set. If V admits a finite nontrivial cover then V contains an immersed MOTS.

Proof: Let $\tilde{V}$ be a finite nontrivial cover of V. Then $\tilde{V}$ (with data pulled back via the covering map) will be AF with more than one, but finitely many ends, as illustrated below.



Hence $\tilde{V}$ contains a MOTS. Then the projection via the covering map is an immersed MOTS in V. □

Theorem 5.3 (no horizon case)

Let (V^3, h, K) be a 3-dimensional AF initial data set. If there are no immersed MOTS in V then V is diffeomorphic to $\mathbb{R}^3$.

Proof:

- By AFness, may write V as $V = \mathbb{R}^3 \# N$ where N is a closed 3-manifold.
- By the residual finiteness of $\pi_1(N)$, if $\pi_1(N) \neq 0$, then N admits a finite nontrivial cover. If that's the case then V also admits a finite nontrivial cover.
- But then, by the Lemma, V contains an immersed MOTS $\rightarrow\leftarrow$.
- Hence, $\pi_1(N) = 0$ and so $N \approx S^3$ by Poincaré. $\therefore V \approx \mathbb{R}^3 \# S^3 \approx \mathbb{R}^3$. $\square$

Consult [8] to see how the argument is modified to deal with the horizon case.

Higher dimensions.

By similar techniques the following is shown in [8]

Theorem 5.6 (Eichmair-G-Pollack)

Let (V, h, K) be an n-dimensional, $3 \leq n \leq 7$, AF initial data set. If V does not contain any immersed MOTS.then V has vanishing first Betti number, $b_1(V) = 0$.

Proof: $b_1(V) \neq 0 \implies V$ contains a nonseparating hypersurface $\implies V$ admits a finite nontrivial cover $\implies V$ contains an immersed MOTS by the Lemma.

The geometry and topology of the exterior region

In very recent work with Andersson, Dahl and Pollack (arXiv, 2015), we have taken an entirely different approach to the study of the topology of the exterior region. We have shown roughly the following.

Theorem 5.7 (Andersson, Dahl, G., Pollack [1])

Let (V^n, h, K) be an n-dimensional initial data set, $3 \le n \le 7$, satisfying the DEC, with MOTS boundary $\Sigma = \partial V$, and assume there are no MOTS in the 'exterior region' $V \setminus \partial V$.

Then the compactification V' admits a metric h' of positive scalar curvature, such that the induced metric on Σ is conformal to that induced by h, and such that h' is a product metric in a neighborhood of Σ.

Remarks:

- In this set-up, one can then apply certain index theory obstructions, and minimal surface theory obstructions, to obtain restrictions on the topology of V.
- This result may be viewed as an extension of Theorem 4.3 on the topology of higher dimensional black holes, whereby the metric of positive scalar curvature metric on the boundary can be extended to the one-point compactification of the region exterior to the boundary.
- Much machinery is needed: Existence results for Jang's equation, Eichmair's regularity and compactness theory for MOTSs, Eichmair-Huang-Lee-Schoen density result, ...

References

[1] L. Andersson, M. Dahl, G. J. Galloway, and D. Pollack, *On the geometry and topology of initial data sets with horizons*, 2015, arXiv:1508.01896.

[2] L. Andersson, M. Eichmair, and J. Metzger, *Jang's equation and its applications to marginally trapped surfaces*, Complex analysis and dynamical systems IV. Part 2, Contemp. Math., vol. 554, Amer. Math. Soc., Providence, RI, 2011, pp. 13–45.

[3] A. N. Bernal and M. Sánchez, *Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes*, Comm. Math. Phys. **257** (2005), no. 1, 43–50.

[4] ______, *Globally hyperbolic spacetimes can be defined as 'causal' instead of 'strongly causal'*, Classical Quantum Gravity **24** (2007), no. 3, 745–749.

[5] D. Christodoulou, *The formation of black holes in general relativity*, Geometry and analysis. No. 1, Adv. Lect. Math. (ALM), vol. 17, Int. Press, Somerville, MA, 2011, pp. 247–283.

[6] P. T. Chruściel, E. Delay, G. J. Galloway, and R. Howard, *Regularity of horizons and the area theorem*, Ann. Henri Poincaré **2** (2001), no. 1, 109–178.

[7] P. T. Chruściel, G. J. Galloway, and D. Solis, *Topological censorship for Kaluza-Klein space-times*, Ann. Henri Poincaré **10** (2009), no. 5, 893–912.

[8] M. Eichmair, G. J. Galloway, and D. Pollack, *Topological censorship from the initial data point of view*, J. Differential Geom. **95** (2013), no. 3, 389–405.

[9] R. Emparan and H. S. Reall, *A rotating black ring solution in five dimensions*, Phys. Rev. Lett. **88** (2002), no. 10, 101101, 4.

[10] J. L. Friedman, K. Schleich, and D. M. Witt, *Topological censorship*, Phys. Rev. Lett. **71** (1993), no. 10, 1486–1489.

[11] G. J. Galloway, *On the topology of the domain of outer communication*, Classical Quantum Gravity **12** (1995), no. 10, L99–L101.

[12] ______, *Rigidity of marginally trapped surfaces and the topology of black holes*, Comm. Anal. Geom. **16** (2008), no. 1, 217–229.

[13] G. J. Galloway and A. Mendes, *Rigidity of marginally outer trapped 2-spheres*, 2015, arXiv:1506.00611, to appear in CAG.

[14] G. J. Galloway and R. Schoen, *A generalization of Hawking's black hole topology theorem to higher dimensions*, Comm. Math. Phys. **266** (2006), no. 2, 571–576.

[15] D. Gannon, *Singularities in nonsimply connected space-times*, J. Mathematical Phys. **16** (1975), no. 12, 2364–2367.

[16] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*, Cambridge University Press, London, 1973, Cambridge Monographs on Mathematical Physics, No. 1.

[17] G. Horowitz (ed.), *Black holes in higher dimensions*, Cambridge University Press, London, 2012.

[18] A. Ionescu and S. Klainerman, *Rigidity results in general relativity: a review*, 2015, arXiv:1501.01587.

[19] S. Klainerman, J. Luk, and I. Rodnianski, *A fully anisotropic mechanism for formation of trapped surfaces in vacuum*, 2013, arXiv:1302.5951v1.

[20] C. W. Lee, *A restriction on the topology of Cauchy surfaces in general relativity*, Comm. Math. Phys. **51** (1976), no. 2, 157–162.

[21] R. C. Myers and M. J. Perry, *Black holes in higher-dimensional space-times*, Ann. Physics **172** (1986), no. 2, 304–347.

[22] Barrett O'Neill, *Semi-Riemannian geometry*, Pure and Applied Mathematics, vol. 103, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983, With applications to relativity.

[23] Robert M. Wald, *General relativity*, University of Chicago Press, Chicago, IL, 1984.