Global Spacetime Structure*

John Byron Manchak
manchak@uw.edu

Abstract

Here, we outline the basic structure of relativistic spacetime and record a number of facts. We then consider a distinction between local and global spacetime properties and provide important examples of each. We also examine two clusters of global properties and question which of them should be regarded as physically reasonable. The properties concern “singularities” and “time travel” and are therefore of some philosophical interest.

1 Introduction

The study of global spacetime structure is a study of the more foundational aspects of general relativity. One steps away from the details of the theory and instead examines the qualitative features of spacetime (e.g. its topology and causal structure).

We divide the following into three main sections. In the first, we outline the basic structure of relativistic spacetime and record a number of facts. In the second, we consider a distinction between local and global spacetime properties and provide important examples of each. In the third, we examine two clusters of global properties and question which of them should be regarded as physically reasonable. The properties concern “singularities” and “time travel” and are therefore of some philosophical interest.

*I am grateful to Bob Batterman, Erik Curiel, and David Malament for comments on a previous draft.
2 Relativistic Spacetime

We take a (relativistic) spacetime to be a pair \((M, g_{ab})\). Here \(M\) is a smooth, connected, \(n\)-dimensional \((n \geq 2)\) manifold without boundary. The metric \(g_{ab}\) is a smooth, non-degenerate, pseudo-Riemannian metric of Lorentz signature \((+,-,...,-)\) on \(M\).\(^1\)

2.1 Manifold and Metric

Let \((M, g_{ab})\) be a spacetime. The manifold \(M\) captures the topology of the universe. Each point in the \(n\)-dimensional manifold \(M\) represents a possible event in spacetime. Our experience tells us that any event can be characterized by \(n\) numbers (one temporal and \(n - 1\) spatial coordinates). Naturally, then, the local structure of \(M\) is identical to \(\mathbb{R}^n\). But globally, \(M\) need not have the same structure. Indeed, \(M\) can have a variety of possible topologies.

In addition to \(\mathbb{R}^n\), the sphere \(S^n\) is certainly familiar to us. We can construct a number of other manifolds by taking Cartesian products of \(\mathbb{R}^n\) and \(S^n\). For example, the 2-cylinder is just \(\mathbb{R}^1 \times S^1\) while the 2-torus is \(S^1 \times S^1\) (see Figure 1). Any manifold with a closed proper subset of points removed also counts as a manifold. For example, \(S^n - \{p\}\) is a manifold where \(p\) is any point in \(S^n\).

Figure 1: The cylinder \(\mathbb{R}^1 \times S^1\) and torus \(S^1 \times S^1\).

We say a manifold \(M\) is Hausdorff if, given any distinct points \(p, p' \in M\), one can find open sets \(O\) and \(O'\) such that \(p \in O\), \(p' \in O'\), and \(O \cap O' = \emptyset\). Physically, Hausdorff manifolds ensure that spacetime events are distinct. In what follows, we assume that manifolds are Hausdorff.\(^2\)

---

\(^1\)In what follows, the reader is encouraged to consult Hawking and Ellis (1973), Geroch and Horowitz (1979), Wald (1984), Joshi (1993), and Malament (2011).

\(^2\)See Earman (2008) for a discussion of this condition.
We say a manifold is *compact* if every sequence of its points has an accumulation point. So, for example, $S^n$ and $S^n \times S^m$ are compact while $\mathbb{R}^n$ and $\mathbb{R}^n \times S^m$ are not. It can be shown that every non-compact manifold admits a Lorentzian metric. But there are some compact manifolds which do not. One example is the manifold $S^4$. Thus, assuming spacetime is four dimensional, we may deduce that the shape of our universe is not a sphere. One can also show that, in four dimensions, if a compact manifold does admit a Lorentzian metric (e.g. $S^1 \times S^3$), it is not simply connected. (A manifold is *simply connected* if any closed curve through any point can be continuously deformed into any other closed curve at the same point.)

We say two manifolds $M$ and $M'$ are *diffeomorphic* if there is a bijection $\varphi : M \to M'$ such that $\varphi$ and $\varphi^{-1}$ are smooth. Diffeomorphic manifolds have identical manifold structure and can differ only in their underlying elements.

The Lorentzian metric $g_{ab}$ captures the geometry of the universe. Each point $p \in M$ has an associated tangent space $M_p$. The metric $g_{ab}$ assigns a length to each vector in $M_p$. We say a vector $\xi^a$ is *timelike* if $g_{ab}\xi^a\xi^b > 0$, *null* if $g_{ab}\xi^a\xi^b = 0$, and *spacelike* if $g_{ab}\xi^a\xi^b < 0$. Clearly, the null vectors create a double cone structure; timelike vectors are inside the cone while spacelike vectors are outside (see Figure 2). In general, the metric structure can vary over $M$ as long as it does so smoothly. But it certainly need not vary and indeed most of the examples considered below will have a metric structure which remains constant (i.e. a *flat* metric).

![Figure 2: Timelike, null, and spacelike vectors fall (respectively) inside, on, and outside the double cone structure.](image)

For some interval $I \subseteq \mathbb{R}$, a smooth curve $\gamma : I \to M$ is *timelike* if its tangent vector $\xi^a$ at each point in $\gamma[I]$ is timelike. Similarly, a curve is *null* (respectively, *spacelike*) if its tangent vector at each point is null (respectively, *null*).
spacelike). A curve is \textit{causal} if its tangent vector at each point is either null or timelike. Physically, the worldlines of massive particles are images of timelike curves while the worldlines of photons are images of null curves. We say a curve $\gamma : I \to M$ is not \textit{maximal} if there is another curve $\gamma' : I' \to M$ such that $I$ is a proper subset of $I'$ and $\gamma(s) = \gamma'(s)$ for all $s \in I$.

We say a spacetime $(M, g_{ab})$ is \textit{temporally orientable} if there exists a continuous timelike vector field on $M$. In a temporally orientable spacetime, a future direction can be chosen for each double cone structure in way that involves no discontinuities. A spacetime which is not temporally orientable can be easily constructed by taking the underlying manifold to be the Möbius strip. In what follows, we will assume that spacetimes are temporally orientable and that a future direction has been chosen.\footnote{See Earman (2002) for a discussion of this condition.}

Naturally, a timelike curve is \textit{future-directed} (respectively, \textit{past-directed}) if all its tangent vectors point in the future (respectively, past) direction. A causal curve is \textit{future-directed} (respectively, \textit{past-directed}) if all its tangent vectors either point in the future (respectively, past) direction or vanish.

Two spacetimes $(M, g_{ab})$ and $(M', g'_{ab})$ are \textit{isometric} if there is a diffeomorphism $\varphi : M \to M'$ such that $\varphi_*(g_{ab}) = g'_{ab}$. Here, $\varphi_*$ is a map which uses $\varphi$ to “move” arbitrary tensors from $M$ to $M'$. Physically, isometric spacetimes have identical properties. We say a spacetime $(M', g'_{ab})$ is a (proper) \textit{extension} of $(M, g_{ab})$ if there is a proper subset $N$ of $M'$ such that $(M, g_{ab})$ and $(N, g'_{ab|N})$ are isometric. We say a spacetime is \textit{maximal} if it has no proper extension. One can show that every spacetime which is not maximal has a maximal extension.

Finally, two spacetimes $(M, g_{ab})$ and $(M', g'_{ab})$ are \textit{locally isometric} if, for each point $p \in M$, there is an open neighborhood $O$ of $p$ and an open subset $O'$ of $M'$ such that $(O, g_{ab|O})$ and $(O', g'_{ab|O'})$ are isometric, and, correspondingly, with the roles of $(M, g_{ab})$ and $(M', g'_{ab})$ interchanged. Although locally isometric spacetimes can have different global properties, their local properties are identical. Consider, for example, the spacetimes $(M, g_{ab})$ and $(M', g'_{ab})$ where $M = S^1 \times S^1$, $p \in M$, $M' = M \setminus \{p\}$, and $g_{ab}$ and $g'_{ab}$ are flat. The two are not isometric but are locally isometric. Therefore, they share the same local properties but have differing global structures (e.g. the first is compact while the second isn’t). One can show that for every spacetime $(M, g_{ab})$, there is a spacetime $(M', g'_{ab})$ such that the two are not isometric but are locally isometric.
2.2 Influence and Dependence

Here, we lay the foundation for the more detailed discussion of causal structure in later sections. Consider the spacetime \((M, g_{ab})\). We define the two-place relations \(\ll\) and \(<\) on the points in \(M\) as follows: we write \(p \ll q\) (respectively, \(p < q\)) if there exists a future-directed timelike (respectively, causal) curve from \(p\) to \(q\). For any point \(p \in M\), we define the timelike future (domain of influence) of \(p\), as the set \(I^+(p) \equiv \{q : p \ll q\}\). Similarly, the causal future (domain of influence) of \(p\) is the set \(J^+(p) \equiv \{q : p < q\}\).

The causal (respectively, timelike) future of \(p\) represents the region of spacetime which can be possibly influenced by particles (respectively, massive particles) at \(p\). The timelike and causal pasts of \(p\), denoted \(I^-(p)\) and \(J^-(p)\), are defined analogously. Finally, given any set \(S \subseteq M\), we define \(I^+[S]\) to be the set \(\cup \{I^+(p) : p \in S\}\). The sets \(I^-[S]\) and \(J^+[S]\), and \(J^-[S]\) are defined analogously. We shall now list a number of properties of timelike and causal pasts and futures.

For all \(p \in M\), the sets \(I^+(p)\) and \(I^-(p)\) are open. Therefore, so are \(I^+[S]\) and \(I^-[S]\) for all \(S \subseteq M\). However, the sets \(J^+(p)\), \(J^-(p)\), \(J^+[S]\) and \(J^-[S]\) are not, in general, either open or closed. Consider Minkowski spacetime\(^4\) and remove one point from the manifold. Clearly, some causal pasts and futures will be neither open nor closed.

By definition, \(I^+(p) \subseteq J^+(p)\) and \(I^-(p) \subseteq J^-(p)\). And it is clear that if \(p \in I^+(q)\), then \(q \in I^-(p)\) and also that if \(p \in I^-(q)\), then \(q \in I^+(p)\). Analogous results hold for causal pasts and futures. We can also show that if either (i) \(p \in I^+(q)\) and \(q \in J^+(r)\) or (ii) \(p \in J^+(q)\) and \(q \in I^+(r)\), then \(p \in J^+(r)\). Analogous results hold for the timelike and causal pasts. From this it follows that \(I^+(p) = J^+(p)\), \(I^-(p) = J^-(p)\), \(\hat{I}^+(p) = \hat{J}^+(p)\), and \(\hat{I}^-(p) = \hat{J}^-(p)\).\(^5\)

Because future-directed causal curves can have vanishing tangent vectors, it follows that for all \(p\), we have \(p \in J^+(p)\) and \(p \in J^-(p)\). Of course, a similar result does not hold generally for timelike futures and pasts. But there do exist some spacetimes such that, for some \(p \in M\), \(p \in J^+(p)\) (and therefore \(p \in I^-(p)\)). Gödel spacetime is one famous example (Gödel 1949).

We say the chronology violating region of a spacetime \((M, g_{ab})\) is the

---

\(^4\)Minkowski spacetime \((M, g_{ab})\) is such that \(M = \mathbb{R}^n\), \(g_{ab}\) is flat, and there exist no incomplete geodesics (defined below). See Hawking and Ellis (1973).

\(^5\)In what follows, for any set \(S\), the sets \(\overline{S}\), \(\hat{S}\), and \(\text{int}(S)\) denote the closure, boundary, and interior of \(S\) respectively.
(necessarily open) set \( \{ p \in M : p \in I^+(p) \} \). We say a timelike curve \( \gamma : I \rightarrow M \) is \textit{closed} if there are distinct points \( s, s' \in I \) such that \( \gamma(s) = \gamma(s') \). Clearly, a spacetime contains a closed timelike curve if and only if it has a non-empty chronology violating region. One can show that, for all spacetimes \((M, g_{ab})\), if \( M \) is compact, the chronology violating region is not empty (Geroch 1967). The converse is false. Take any compact spacetime and remove one point from the underlying manifold. The resulting spacetime will contain closed timelike curves and also fail to be compact.

Figure 3: \textit{Cylindrical Minkowski spacetime containing a closed causal curve (e.g. the dotted line) but no closed timelike curves.}

We define a causal curve \( \gamma : I \rightarrow M \) to be \textit{closed} if there are distinct points \( s, s' \in I \) such that \( \gamma(s) = \gamma(s') \) and \( \gamma \) has no vanishing tangent vectors. It is immediate that closed timelike curves are necessarily closed causal curves. But one can find spacetimes which contain the latter but not the former. Consider, for example, Minkowski spacetime \((M, g_{ab})\) which has been “rolled up” along one axis in such a way that some null curves but no timelike curves are permitted to loop around \( M \) (see Figure 3). Other conditions relating to “almost” closed causal curves will be considered in the next section.

Finally, we say the spacetimes \((M, g_{ab})\) and \((M, g'_{ab})\) are \textit{conformally related} if there is a smooth, strictly positive function \( \Omega : M \rightarrow \mathbb{R} \) such that \( g'_{ab} = \Omega^2 g_{ab} \) (the function \( \Omega \) is called a \textit{conformal factor}). Clearly, if \((M, g_{ab})\) and \((M, g'_{ab})\) are conformally related, then for all points \( p, q \in M \), \( p \in I^+(q) \) in \((M, g_{ab})\) if and only if \( p \in I^+(q) \) in \((M, g'_{ab})\). Analogous results hold for timelike pasts and causal futures and pasts. Thus, the causal structures of conformally related spacetimes are identical.
A point \( p \in M \) is a **future endpoint** of a future-directed causal curve \( \gamma : I \to M \) if, for every neighborhood \( O \) of \( p \), there exists a point \( s' \in I \) such that \( \gamma(s) \in O \) for all \( s > s' \). A **past endpoint** is defined analogously. For any set \( S \subseteq M \), we define the **future domain of dependence of** \( S \), denoted \( D^+(S) \), to be the set of points \( p \in M \) such that every causal curve with future endpoint \( p \) and no past endpoint intersects \( S \). The **past domain of dependence of** \( S \), denoted \( D^-(S) \), is defined analogously. The entire **domain of dependence of** \( S \), denoted \( D(S) \), is just the set \( D^-(S) \cup D^+(S) \). If “nothing can travel faster than light”, there is a sense in which the physical situation at every point in \( D(S) \) depends entirely upon the physical situation on \( S \).

Clearly, we have \( S \subseteq D^+(S) \subseteq J^+[S] \) and \( S \subseteq D^-(S) \subseteq J^-[S] \). Given any point \( p \in D^+(S) \), and any point \( q \in I^+[S] \cap I^-(p) \), we know that \( q \in D^+(S) \). An analogous result holds for \( D^-(S) \). One can verify that, in general, \( D(S) \) is neither open nor closed. Consider Minkowski spacetime \((M, g_{ab})\). If \( S = \{p\} \) for any point \( p \in M \), we have \( D(S) = S \) which is not open. If \( S = I^+(p) \cap I^-(q) \) for any points \( p \in M \) and \( q \in I^+(p) \), we have \( D(S) = S \) which is not closed.

A set \( S \subseteq M \) is a **spacelike surface** if \( S \) is a submanifold of dimension \( n - 1 \) such that every curve in \( S \) is spacelike. We say a set \( S \subseteq M \) is **achronal** if \( I^+[S] \cap S = \emptyset \). One can show that for an arbitrary set \( S \), \( I^+[S] \) is achronal. In what follows, let \( S \) be a closed, achronal set. We have \( D^+(S) \cap I^-[S] = D^-(S) \cap I^+[S] = \emptyset \). We also have \( \text{int}(D^+(S)) = I^-[D^+(S)] \cap I^+[S] \) and the analogous result for \( D^-(S) \). Finally, we have \( \text{int}(D(S)) = I^-[D^+(S)] \cap I^+[D^-(S)] = I^+\left[D^-(S) \cap I^-[D^+(S)]\right] \).

We say a closed, achronal set \( S \) is a **Cauchy surface** if \( D(S) = M \). Physically, conditions on a Cauchy surface \( S \) (necessarily a submanifold of \( M \) of dimension \( n - 1 \)) determine conditions throughout spacetime (Choquet-Bruhat and Geroch 1969). Clearly, if \( S \) is a Cauchy surface, any causal curve without past or future endpoint must intersect \( S \), \( I^+[S] \), and \( I^-[S] \). One can verify that Minkowski spacetime admits a Cauchy surface.

We define the **future Cauchy horizon** of \( S \), denoted \( H^+(S) \), as the set \( D^+(S) \cap I^-[D^+(S)] \). The **past Cauchy horizon** of \( S \) is defined analogously. One can verify that \( H^+(S) \) and \( H^-(S) \) are closed and achronal. The **Cauchy horizon** of \( S \), denoted \( H(S) \), is the set \( H^+(S) \cup H^-(S) \). We have \( H(S) = \hat{D}(S) \) and therefore \( H(S) \) is closed. Also, a non-empty, closed, achronal set \( S \) is a Cauchy surface if and only if \( H(S) = \emptyset \).

The **edge** of a closed, achronal set \( S \subseteq M \) is the set of points \( p \in S \) such that every open neighborhood \( O \) of \( p \) contains a point \( q \in I^+(p) \), a point
Figure 4: *Minkowski spacetime with one point removed contains a slice $S$ but no Cauchy surface. The region above the dotted line is not part of $D(S)$."

$r \in I^{-}(p)$, and a timelike curve from $r$ to $q$ which does not intersect $S$. A closed, achronal set $S \subset M$ is a slice if it is without edge. It follows that every Cauchy surface is a slice. The converse is false. Consider Minkowski spacetime with one point removed from the manifold. It certainly admits a slice but no Cauchy surface (see Figure 4). Of course, not every spacetime admits a slice. For a counterexample, consider any spacetime which has a chronology violating region identical to its manifold.

3 Spacetime Properties

We say a property $P$ on a spacetime is *local* if, given any two locally isometric spacetimes $(M, g_{ab})$ and $(M', g'_{ab})$, $(M, g_{ab})$ has $P$ if and only if $(M', g'_{ab})$ has $P$. A property is *global* if it is not local. Below, we will introduce and classify a number of spacetime properties of interest.

3.1 Local Properties

The most important local spacetime property is that of being a “solution” to Einstein’s equation. There are a number of ways one can understand this property and we shall investigate each of them in what follows.

Let $(M, g_{ab})$ be a spacetime. Associated with the metric $g_{ab}$ is a unique (torsion-free) derivative operator $\nabla_{a}$ such that $\nabla_{a}g_{bc} = 0$. Given a smooth curve $\gamma : I \rightarrow M$ with tangent field $\xi^{a}$, we say a vector $\eta^{a}$, defined at every point in the range of $\gamma$, is *parallelly transported* along $\gamma$ if $\xi^{b}\nabla_{b}\eta^{a} = 0$ (see Figure 5). We say a smooth curve $\gamma : I \rightarrow \mathbb{R}$ is a *geodesic* (i.e. non-
accelerating) if its tangent field $\xi^a$ is such that $\xi^b \nabla_b \xi^a = 0$. Given any point $p \in M$, there is some neighborhood $O$ of $p$ such that any two points $q, r \in O$ can be connected by a unique geodesic contained entirely in $O$. Such a neighborhood is said to be **convex normal**.

The derivative operator $\nabla_a$ can be used to define the **Riemann curvature tensor**. It is the unique tensor $R^a_{bcd}$ such that for all $\xi^a$, $R^a_{bcd} \xi^b = -2 \nabla_c \nabla_d \xi^a$. A metric $g_{ab}$ on $M$ is flat if and only if its associated Riemann curvature tensor $R^a_{bcd}$ vanishes everywhere on $M$. The tensors $R^a_{bcd}$ and $R_{abcd}$ have a number of useful symmetries: $R^a_{b(cd)} = 0$, $R^a_{[bcd]} = 0$, $\nabla_{[a} R^a_{b|c|d]} = 0$, $R_{ab(cd)} = 0$, $R_{a[bcd]} = 0$, $R_{(ab)cd} = 0$, and $R_{abcd} = R_{cdab}$.

We define the **Ricci tensor** $R_{ab}$ to be $R^c_{abc}$ and the **scalar curvature** $R$ to be $R^a_a$. The **Einstein tensor** $G_{ab}$ is then defined as $R_{ab} - \frac{1}{2} R g_{ab}$. It plays a central role in what follows. One can verify that $\nabla_a G_{ab} = 0$.

We suppose that the entire matter content of the universe can be characterized by smooth tensor fields on $M$. For example, a source-free electromagnetic field is characterized by an anti-symmetric tensor $F_{ab}$ on $M$ which satisfies Maxwell’s equations: $\nabla_{[a} F_{bc]} = 0$, $\nabla^a F_{ab} = 0$. Other forms of matter, such as perfect fluids and Klein-Gordon fields, are characterized by other smooth tensor fields on $M$.

Associated with each matter field is a smooth, symmetric **energy-momentum tensor** $T_{ab}$ on $M$. For example, the energy-momentum tensor $T_{ab}$ associated with an electromagnetic field $F_{ab}$ is $F_{an} F^n_b + \frac{1}{4} g_{ab} (F^{nm} F_{nm})$. Note that $T_{ab}$

---

6In what follows, square brackets denote anti-symmetrization. Parentheses denote symmetrization. See Malament (2011).
is a function not only of the matter field itself but also of the metric. Other matter fields, such as those mentioned above, will have their own energy-momentum tensors $T_{ab}$.

Fix a point $p \in M$. The quantity $T_{ab}\xi^a\xi^b$ at $p$ represents the energy density of matter as given by an observer with tangent $\xi^a$ at $p$. The quantity $T^a_b\xi^b - T_{ab}\xi^a\xi^b$ at $p$ represents the spatial momentum density as given by the same observer at $p$. We require that any energy-momentum tensor satisfy the conservation condition: $\nabla^aT_{ab} = 0$. Physically, this ensures that energy-momentum is locally conserved.

Finally, we come to Einstein's equation: $G_{ab} = 8\pi T_{ab}$. It relates the curvature of spacetime with the matter content of the universe. In four dimensions, Einstein's equation can be expressed as $R_{ab} = 8\pi(T_{ab} - \frac{1}{2}Tg_{ab})$ where $T = T_a^a$.

Of course, any spacetime $(M, g_{ab})$ can be thought of as a trivial solution to Einstein's equation if $T_{ab}$ is simply defined to be $\frac{1}{8\pi}G_{ab}$. Note that $T_{ab}$ automatically satisfies the conservation condition since $\nabla^aG_{ab} = 0$. But, in general, the energy momentum tensor defined in this way will not be associated with any known matter field. However, if the $T_{ab}$ so defined is also the energy momentum tensor associated with a known matter field (or the sum of two or more energy momentum tensors associated with known matter fields) the spacetime is an exact solution. We say an exact solution is also a vacuum solution if $T_{ab} = 0$. And, in four dimensions, one can use the alternate version of Einstein's equation to show that $T_{ab} = 0$ if and only if $R_{ab} = 0$.

Between trivial and exact solutions, there are the constraint solutions. These are spacetimes whose associated energy-momentum tensors (defined via Einstein's equation) satisfy one or more conditions of interest. Here, we outline three. We say $T_{ab}$ satisfies the weak energy condition if, for any future-directed unit timelike vector $\xi^a$ at any point in $M$, the energy density $T_{ab}\xi^a\xi^b$ is not negative.

We say $T_{ab}$ satisfies the strong energy condition if, for any future-directed unit timelike vector $\xi^a$ at any point in $M$, the quantity $(T_{ab} - \frac{1}{2}Tg_{ab})\xi^a\xi^b$ is not negative. The strong energy condition can be interpreted as the requirement that a certain effective energy density is not negative. Note that, in four dimensions, the strong energy condition is satisfied if and only if the

\[^7\text{Here, we drop the cosmological constant term } -\Lambda g_{ab} \text{ sometimes added to the left side of the equation for some } \Lambda \in \mathbb{R}. \text{ For more on this term, see Earman (2001).}\]
(timelike) convergence condition, $R_{ab}\xi^a\xi^b \geq 0$, is also satisfied. This latter condition can be understood to assert that gravitation is attractive in nature.

Finally, we say $T_{ab}$ satisfies the dominant energy condition if, for any future-directed unit timelike vector $\xi^a$ at any point in $M$, the vector $T^a_{\ b}\xi^b$ is causal and future-directed. This last condition can be interpreted as the requirement that matter cannot travel faster than light. Indeed, if $T_{ab}$ vanishes on some closed, achronal set $S \subset M$ and satisfies the dominant energy and conservation conditions, then $T_{ab}$ vanishes on all of $D(S)$ (Hawking and Ellis 1973). Clearly, the dominant energy condition implies (but is not implied by) the weak energy condition.

One can show that being a trivial, exact, or vacuum solution of Einstein’s equation is a local spacetime property. In addition, being a constraint solution is also a local spacetime property if the constraint under consideration is one of the three energy conditions considered here.

3.2 Global Properties

A large number of important global properties concern either “causal structure” or “singularities”. Here we investigate them.

There is a hierarchy of conditions relating to the causal structure of spacetime. Each condition corresponds to a global spacetime property (the property of satisfying the condition). We say a spacetime satisfies the chronology condition if it contains no closed timelike curves (equivalently, $p \notin I^+(p)$ for all $p \in M$). A spacetime satisfies the causality condition if there are no closed causal curves (equivalently, $J^+(p) \cap J^-(p) = \{p\}$ for all $p \in M$). As mentioned previously, causality implies chronology but the implication does not run in the other direction (see Figure 3). The next few conditions serve to rule out “almost” closed causal curves.

We say a spacetime $(M, g_{ab})$ satisfies the future distinguishability condition if there do not exist distinct points $p, q \in M$ such that $I^+(p) = I^+(q)$. The past distinguishability condition is defined analogously. One can show that a spacetime $(M, g_{ab})$ satisfies the future (respectively, past) distinguishability condition if and only if, for all points $p \in M$ and every open set $O$ containing $p$, there is an open set $V \subset O$ also containing $p$ such that no future (respectively, past) directed causal curve that starts at $p$ and leaves $V$ ever

\footnote{Although we only consider a small handful here, there are an infinite number of conditions in the causal hierarchy (Carter 1971).}
returns to $V$. We say a spacetime satisfies the *distinguishability* condition if it satisfies both the past and future distinguishability conditions.

Future or past distinguishability implies causality. But the converse is not true. Of course, distinguishability implies past (or future) distinguishability. But one can certainly find spacetimes which satisfy future (respectively, past) distinguishability but not past (respectively, future) distinguishability (Hawking and Ellis 1973).

Consider two distinguishing spacetimes $(M, g_{ab})$ and $(M', g'_{ab})$ and a bijection $\varphi : M \rightarrow M'$ such that for all $p, q \in M$, $p \in I^+(q)$ if and only if $\varphi(p) \in I^+(\varphi(q))$. One can show (Malament 1977) that $\varphi$ is a diffeomorphism and $\varphi^*(g_{ab}) = \Omega^2 g'_{ab}$ for some conformal factor $\Omega : M' \rightarrow \mathbb{R}$. Thus, if the causal structure of spacetime is sufficiently well-behaved, that structure alone determines the shape of the universe as well as the metric structure up to a conformal factor.

We say a spacetime satisfies the *strong causality* condition if, for all points $p \in M$ and every open set $O$ containing $p$, there is an open set $V \subset O$ also containing $p$ such that no causal curve intersects $V$ more than once. If a spacetime $(M, g_{ab})$ satisfies strong causality, then, for every compact set $K \subset M$, a causal curve $\gamma : I \rightarrow K$ must have future and past endpoints in $K$. Thus, in a strongly causal spacetime, an inextendible causal curve cannot be “imprisoned” in a compact set. Clearly, strong causality implies distinguishability. One can show that the implication does not run in the other direction (Hawking and Ellis 1973).

A spacetime $(M, g_{ab})$ satisfies the *stable causality* condition if there is a timelike vector field $\xi^a$ on $M$ such that the spacetime $(M, g_{ab} + \xi_a \xi^b)$ satisfies the chronology condition. Physically, even if the light cones are “opened” by a small amount at each point, the spacetime remains free of closed timelike curves. We say a spacetime $(M, g_{ab})$ admits a *global time function* if there is a smooth function $t : M \rightarrow \mathbb{R}$ such that, for any distinct points $p, q \in M$, if $p \in J^+(q)$, then $t(p) > t(q)$. The function assigns a “time” to every point in $M$ such that it increases along every (non-trivial) future-directed causal curve. An important result is that a spacetime admits a global time function if and only if it satisfies stable causality (Hawking 1969). One can also show that stable causality implies strong causality but the converse is false (see Figure 6).

The remaining causality conditions not only require that there be no almost closed causal curves but, in addition, that there be limitations on the kinds of “gaps” in spacetime (Hawking and Sachs 1974).
We say a spacetime \((M, g_{ab})\) satisfies the causal continuity condition if it satisfies distinguishability and, for all \(p, q \in M\), \(I^+(p) \subseteq I^+(q)\) if and only if \(I^-(q) \subseteq I^-(p)\). Physically, causal continuity ensures that points which are close to one another do not have wildly different timelike futures and pasts. One can show that causal continuity implies stable causality. The converse is not true. A counterexample can be constructed by taking Minkowski spacetime and excising from the manifold a closed proper subset with non-empty interior. The resulting spacetime satisfies stable causality but not causally continuity.

A spacetime \((M, g_{ab})\) satisfies the causal simplicity condition if it satisfies distinguishability and, in addition, for all \(p \in M\), the sets \(J^+(p)\) and \(J^-(p)\) are closed. One can show that causal simplicity implies causal continuity. The converse is false since Minkowski spacetime with a point removed from the manifold satisfies causal continuity but not causal simplicity.

Finally, we say a spacetime \((M, g_{ab})\) satisfies global hyperbolicity if it satisfies strong causality and, in addition, for all \(p, q \in M\), the set \(J^+(p) \cap J^-(q)\) is compact. A fundamental result is that a spacetime satisfies global hyperbolicity if and only if it admits a Cauchy surface (Geroch 1970b). In addition, one can show that the manifold of any spacetime which satisfies global hyperbolicity must have the topology of \(\mathbb{R} \times \Sigma\) for any Cauchy surface \(\Sigma\). Global hyperbolicity implies causal simplicity but the converse is not true. Anti-de Sitter spacetime is one counterexample (Hawking and Ellis 1973).

In sum, we have the following implications (none of which run in the other direction): global hyperbolicity \(\Rightarrow\) causal simplicity \(\Rightarrow\) causal continuity \(\Rightarrow\) stable causality \(\Rightarrow\) strong causality \(\Rightarrow\) distinguishability \(\Rightarrow\) future (or past) distinguishability \(\Rightarrow\) causality \(\Rightarrow\) chronology.

There are a number of senses in which a spacetime may be said to contain
a “singularity”. Here, we restrict attention to the most important one: geodesic incompleteness. We say a geodesic \( \gamma : I \to M \) is \textit{incomplete} if it is maximal and such that \( I \neq \mathbb{R} \). We say a future-directed maximal timelike or null geodesic \( \gamma : I \to M \) is \textit{future incomplete} (respectively, \textit{past incomplete}) if there is a \( r \in \mathbb{R} \) such that \( r > s \) for all \( s \in I \). A \textit{past incomplete} geodesic is defined analogously.

Naturally, a spacetime is \textit{timelike geodesically incomplete} if it contains a timelike incomplete geodesic. In a timelike geodesically incomplete spacetime, it is possible for a non-accelerating massive particle to experience only a finite amount of time. We can define \textit{spacelike} and \textit{null geodesic incompleteness} analogously. Finally, we say that a spacetime is \textit{geodesically incomplete} if it is either timelike, spacelike, or null geodesically incomplete.

If a spacetime has an extension, it is geodesically incomplete. The converse is false. Consider Minkowski spacetime \((M, g_{ab})\) and let \( M' \) be the manifold \( M - \{p\} \) for any \( p \in M \). Let \( \Omega : M' \to \mathbb{R} \) be a conformal factor which approaches zero as the missing point \( p \) is approached. The resulting spacetime \((M', \Omega g_{ab}|M')\) is maximal but contains timelike, spacelike, and null incomplete geodesics. Other maximal spacetimes exist which are geodesically incomplete and have a flat metric. In other words, one can have singularities without any spacetime curvature at all. Since there are certainly flat spacetimes which are geodesically complete (e.g. Minkowski spacetime), it follows that geodesic incompleteness is a global property. We mention in passing that the property of being maximal is also global.

Finally, one can show that timelike, spacelike, and null incompleteness are independent conditions in the sense that there are spacetimes which are incomplete in any one of the three types and complete in the other two (Geroch 1968). Additionally, one can show that compact spacetimes are not necessarily geodesically complete (Misner 1963). These two results suggest that geodesic incompleteness fails to mesh completely with our notion of a “hole” in spacetime.

---

9See Ellis and Schmidt (1977), Geroch, Liang, and Wald (1982), Clarke (1993), and Curiel (1999) for details.

10Here is one example. Remove a point from \( \mathbb{R}^2 \) and take the universal covering space. Let the resulting spacetime manifold have a flat metric.
4 Which Properties are Reasonable?

So far, we have provided examples of a number of spacetime properties. In this section, we ask: Which properties are “physically reasonable”?

It is usually taken for granted that “the normal physical laws we determine in our spacetime vicinity are applicable at all other spacetime points” (Ellis 1975). This assumption allows us to stipulate that the local property of being a solution to Einstein’s equation is a physically reasonable one. And often this means that we take the energy conditions as necessarily satisfied. However, some have argued that even the energy conditions can be violated in some physically reasonable spacetimes (Vollick 1997).

One global property which is usually taken to be physically reasonable is that spacetime be maximal. Metaphysical considerations seem to drive the assumption. One asks (Geroch 1970a), “Why, after all, would Nature stop building our universe...when She could just as well have carried on?” Of course, such reasoning can be questioned (Earman 1995).

What about the global properties concerning singularities and causal structure? Which of them are to be considered physically reasonable?

4.1 Singularities

Much of the work in global structure has concerned singularities. The task has been to show, using fairly conservative assumptions, that all physically reasonable spacetimes must be (null or timelike) geodesically incomplete. The project has produced a number of theorems of this type. Here, we examine an influential one due to Hawking and Penrose (1970).

Three preliminary conditions are crucial and each have been taken to be satisfied by all (or almost all) physically reasonable spacetimes. We shall temporarily adopt these background assumptions in what follows. The first is chronology (no closed timelike curves). The second is the convergence condition ($R_{ab}\xi^a\xi^a \geq 0$ for all unit timelike vectors $\xi^a$). Recall that the convergence condition is satisfied in four dimensions if and only if the strong energy condition is. In this section, we will restrict attention to four dimensional spacetimes. The third is the generic condition – that each causal geodesic with tangent $\xi^a$ contains a point at which $\xi_{[a}R_{b]cd[e}\xi_f]\xi^c\xi^d \neq 0$. Physically, the generic condition requires that somewhere along each causal curve a certain effective curvature is encountered. Although highly symmetric spacetimes may not satisfy the generic condition (e.g. Minkowski spacetime) it is
thought to be satisfied by all sufficiently “generic” ones. Now, consider the following statement.

(S) Any spacetime which satisfies chronology, the convergence condition, the generic condition, and ________, must be timelike or null geodesically incomplete.

We seek to fill in the blank with physically reasonable “boundary” conditions which make (S) true. Hawking and Penrose (1970) considered three of them (see also Earman 1999).

First, if the boundary condition is the requirement that there exist a compact slice, (S) is true. So, a “spatially closed” universe is singular if it is physically reasonable. One can show that the existence of a compact slice is a necessary condition for predicting future spacetime events (Manchak 2008). Thus, we have the somewhat counterintuitive result that prediction is possible in a physically reasonable spacetime only if singularities are present.\(^\text{11}\)

Second, (S) is true if the boundary condition is the requirement that there exist a trapped surface. A trapped surface is a two-dimensional compact spacelike surface \(T\) such that both sets of “ingoing” and “outgoing” future-directed null geodesics orthogonal to \(T\) have negative expansion at \(T\).\(^\text{12}\) Physically, whenever a sufficiently large amount of matter is contained in a small enough region of spacetime, a trapped surface forms (Schoen and Yau 1983).

Third, (S) is true if the boundary condition is the requirement that there is a point \(p \in M\) such that the expansion along every future (or past) directed null geodesic through \(p\) is somewhere negative. Physically, a spacetime which satisfies this condition contains a contracting region in the causal future (or past) of a point. It is thought that the observable portion of our own universe contains such a region (Ellis 2007).

Additional examples of boundary conditions which make (S) true could be multiplied (Senovilla 1998). And instead of boundary conditions, one can also find causal conditions which make (S) true. We mention one here. It turns out that (S) is true if the causal condition is the requirement that stable causality is not satisfied (Minguzzi 2009). Thus, physically reasonable

\(^{11}\)For a related discussion, see Hogarth (1997).

\(^{12}\)The (scalar) expansion of a congruence of null geodesics is a bit complicated to define (see Wald 1984). But one can get some some idea of the quantity by noting that the expansion of a congruence of timelike geodesics with unit tangent field \(\xi^a\) is \(\nabla_a \xi^a\).
spacetimes (which are assumed to be causally well behaved in the sense that they satisfy chronology) are singular if they are not too causally well behaved. One naturally wonders if it is possible for physically reasonable spacetimes to avoid singularities if the chronology condition is dropped. But this seems unlikely (Tipler 1977, Kriele 1990).

A large number of physically reasonable spacetimes (including our own) seem to satisfy at least one of the above mentioned boundary conditions and hence contain singularities. And the worry has been that these singularities can be observed directly – that they are “naked” in some sense. So, one would like to show that all (or almost all) physically reasonable spacetimes do not contain naked singularities. This is the “cosmic censorship” hypothesis. There are a number of ways to formulate the hypothesis (Joshi 1993, Penrose 1999). Here, we consider one.

Figure 7: Minkowski spacetime with one point removed is nakedly singular. The future incomplete geodesic $\gamma$, contained in the timelike past of $p$, approaches the missing point.

We do not wish to count a “big bang” singularity as naked and therefore restrict attention to future (rather than past) incomplete timelike or null geodesics. We say a spacetime $(M, g_{ab})$ is nakedly singular if there is a point $p \in M$ and a future incomplete timelike or null geodesic $\gamma : I \rightarrow M$ such that the range of $\gamma$ is contained in $I^-(p)$ (see Figure 7).

One can show that a nakedly singular spacetime does not admit a Cauchy surface (Geroch and Horowitz 1979). Thus, if all physically reasonable spacetimes are globally hyperbolic, then the cosmic censorship hypothesis is true. And Penrose (1969, 1979) has suggested that one might be able to show the antecedent of this conditional. The idea would be to show that spacetimes which fail to be globally hyperbolic are unstable under certain types of perturbations. However, such a claim is difficult to express precisely (Geroch
And although some evidence does seem to indicate that instabilities are present in non-globally hyperbolic spacetimes (Chandrasekhar and Hartle 1982), still other evidence suggests otherwise (Morris, Thorne, and Yurtsever 1988).

There is also an epistemological predicament at issue. An observer never can have the evidential resources to rule out the possibility that his or her spacetime is not globally hyperbolic—even under any assumptions concerning local spacetime structure (Manchak 2011b). And how could we ever know that all physically reasonable spacetimes are globally hyperbolic if we cannot even be confident that our own spacetime is?

4.2 Time Travel

If the cosmic censorship hypothesis is false, there are physically reasonable spacetimes which do not satisfy global hyperbolicity. Might there be some physically reasonable spacetimes which do not even satisfy chronology? We investigate the question here.

One way to rule out a number of chronology-violating spacetimes concerns self-consistency constraints on matter fields of various types. Here, we examine source free Klein-Gordon fields. Let \((M, g_{ab})\) be a spacetime. We say an open set \(U \subset M\) is *causally regular* if, for every function \(\varphi : U \rightarrow \mathbb{R}\) which satisfies \(\nabla^a \nabla_a \varphi = 0\), there is a function \(\varphi' : M \rightarrow \mathbb{R}\) such that \(\nabla^a \nabla_a \varphi' = 0\) and \(\varphi'|_U = \varphi\). We say \((M, g_{ab})\) is *causally benign* if, for every \(p \in M\) and every open set \(U\) containing \(p\), there is an open set \(U' \subset U\) containing \(p\) which is causally regular.

It has been argued that a spacetime which is not causally benign is not physically reasonable. We certainly know that every globally hyperbolic spacetime is causally benign. But although some chronology violating spacetimes are not causally benign, a number of others are (Yurtsever 1990, Friedman 2004).

Given the existence of causally benign yet chronology violating spacetimes, another area of research seems fruitful to pursue. One wonders if chronology violating region can, in some sense, be “created” by rearranging the distribution and flow of matter (Stein 1970). In other words, can a physically reasonable spacetime contain a “time machine” of sorts? Here, we examine one way of formalizing the question given by Earman, Smeenk, and
Wüthrich (2009).\textsuperscript{13}

First, in order to count as a time machine, a spacetime \((M, g_{ab})\) must contain a spacelike slice \(S \subset M\) representing a “time” before the time machine is switched on. Second, the spacetime must also have a chronology violating region \(V\) after the machine is turned on. So we require \(V \subset J^+[S]\). Finally, in order to capture the idea that a time machine must “create” a chronology violating region, every physically reasonable maximal extension of \(\text{int}(D(S))\) must contain a chronology violating region \(V'\).\textsuperscript{14} Consider the following statement.

\begin{itemize}
  \item \(\text{T}\) There is a spacetime \((M, g_{ab})\) with a spacelike slice \(S \subset M\) and a chronology violating region \(V \subset J^+[S]\) such that every maximal extension of \(\text{int}(D(S))\) which satisfies \(\) contains some chronology violating region \(V'\).
\end{itemize}

We seek to fill in the blank with physically reasonable “potency” conditions which make (T) true. And we know from counterexamples constructed by Krasnikov (2002) that (T) will be false unless there \textit{is} a potency condition and this condition limits spacetime “holes” in some sense.

But Hawking (1992) has suggested that limiting holes may not be enough. Indeed, he conjectured that all physically reasonable spacetimes are “protected” from chronology violations and provided some evidence for the claim. We say \(H^+(S)\) is \textit{compactly generated} if all past directed null geodesics through \(H^+(S)\) enter and remain in some compact set. Any spacetime with a slice \(S\) such that \(H^+(S)\) is non-empty and compactly generated does not satisfy strong causality. And Hawking showed there is no spacetime which satisfies the weak energy condition which has a non-compact slice \(S\) such that \(H^+(S)\) is non-empty and compactly generated.

But some have argued that insisting on a compactly generated Cauchy horizon rules out some physically reasonable spacetimes (Ori 1993, Krasnikov 1999). And of course, a slice \(S\) need not be non-compact to be physically reasonable. Thus, Hawking’s chronology protection conjecture remains an open question.

Are there any potency conditions which make (T) true? We say a spacetime \((M, g_{ab})\) is \textit{hole-free} if, for any spacelike surface \(S\) in \(M\) there is no

\textsuperscript{13}See also Earman and Wüthrich (2010) and Smeenk and Wüthrich (2011).

\textsuperscript{14}Here we abuse the notation somewhat. Properly, we require that every physically reasonable maximal extension of \((\text{int}(D(S)), g_{ab|\text{int}(D(S))})\) must contain a chronology violating region \(V'\).
isometric embedding $\theta : D(S) \to M'$ into another spacetime $(M', g'_{ab})$ such that $\theta(D(S)) \neq D(\theta(S))$. Physically, hole-freeness ensures that, for any spacelike surface $S$, the domain of dependence $D(S)$ is “as large as it can be”. And one can show that any spacetime with one point removed from the underlying manifold fails to be hole-free. It has been argued that all physically reasonable spacetimes are hole-free (Clarke 1976, Geroch 1977). And it turns out that (T) is true if the potency condition is hole-freeness (Manchak 2009b). The two-dimensional spacetime of Misner (1967) can be used to prove the result (see Figure 8).

Figure 8: Misner spacetime. Every maximal, hole-free extension of $\text{int}(D(S))$ (the region below the dotted line) contains some chronology violating region.

However, hole-freeness may not be a physically reasonable potency condition after all. Indeed, some maximal, globally hyperbolic models, including Minkowski spacetime, are not hole-free (Manchak 2009a, Krasnikov 2009). But, another more reasonable “no holes” potency condition can be used to make (T) true: the demand that, for all $p \in M$, $J^+(p)$ and $J^-(p)$ are closed (Manchak 2011a). Call this condition causal closedness and recall that causal closedness is used, along with distinguishability, to define causal simplicity.

Not only is causal closedness satisfied by all globally hyperbolic models, including Minkowski spacetime, but it is also satisfied by many chronology violating spacetimes as well (e.g. Gödel spacetime, Misner spacetime). In this sense, then, it is a more appropriate condition than hole-freeness. But is causal closedness satisfied by all physically reasonable spacetimes? The question is open. So too is the question of which other potency conditions make (T) true.
5 Conclusion

Here, we have outlined the basic structure of relativistic spacetime. As we have seen, general relativity allows for a wide variety of global spacetime properties – some of them quite unusual. And one wonders which of these properties are physically reasonable.

Early work focused on singularities. Initially, a number of results established that all physically reasonable spacetimes are geodesically incomplete. Next, the relationship between these singularities and determinism was investigated: Can a physically reasonable (and therefore geodesically incomplete) spacetime fail to be globally hyperbolic? The question remains open.

Recently, focus has shifted somewhat toward acausality: Can physically reasonable spacetimes contain closed timelike curves? If so, can these closed timelike curves be “created” in some sense by rearranging the distribution and flow of matter? Again, these questions remain open.

References


