Causality, Black Holes, Prediction, and Counterfactuals in General Relativity

Martin Lesourd

Linacre College, University of Oxford, UK

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

May 2019

Acknowledgements

I am deeply grateful to both of my supervisors - Professors Harvey Brown and Christopher Timpson. Harvey's belief, insight, and support has been instrumental throughout this journey, and Chris' guidance has profoundly impacted and lifted my understanding of the ideas presented.

I express my sincerest gratitude to Professors John Manchak and Erik Curiel with whom I have had the opportunity and pleasure of discussing many of the ideas in the thesis. There is no doubt that I wouldn't have been able to complete let alone start this work without having been inspired by their articles. John Manchak's papers form much of the basis of part III of the thesis, and part II.1 was essentially inspired by Erik Curiel. I would also like to thank them for their respective invitations to visit UC Irvine and Harvard's Black Hole Initiative, during which I was given the opportunity to present some of these ideas. I would also like to thank Professor Weatherall for supporting my visit at Irvine.

I would like to thank Professor Paul Tod of the Oxford Mathematics department for his support in my mathematical endeavors and for the numerous enlightening conversations on the subject of mathematical relativity. I would also like to thank Dr Jan Sbierksi for discussions which never failed to leave me with an impression of clarity.

I would also like to thank Dr Adam Caulton for his very useful feedback on some of the ideas in part III of the thesis.

I would also like to thank Tushar Menon, Patrick Duerr, and James Read for many illucidating conversations.

I would of course like to express my sincerest appreciation and acknowledgement of the financial support that has allowed me to carry out this research. In

particular, I thank dearly Michael and Alice Crampin of the Ruth and Nevill Mott Scholarship of Linacre College, and the AHRC.

Finally, and perhaps most of all, I would like to thank my family for their unconditional love and support, and to the many friends that have been there for me all along with the way. It is certain that I would never have been able to start or complete this project without them.

Abstract

This thesis treats foundational questions in classical general relativity. The emphasis is towards proving new results in Lorentzian geometry which bear on these questions. For the most part, the results either improve or build upon the work of previous authors, including Geroch, Hawking, Kodama, Malament, Manchak, Newman, Penrose, Tipler, and Wald. In some cases I propose strengthened versions of known theorems, and in others I prove new results that are meant to better expose the meaning of known results. The subjects are broadly divided into three parts, II, III, and IV.

Part II focuses on black hole phenomena and is composed of three independent sections treating, respectively, Hawking's area theorem, black hole singularities in the presence of closed timelike curves, and the causal structure of evaporating black holes in view of the Kodama-Geroch-Wald theorem.

Part III is focused on the epistemological status of observers in general relativistic spacetimes and the meaning of Penrose's enigmatic conjecture of strong cosmic censorship. Here, the independent works of Malament and Manchak provide the crucial starting point.

Part IV is more discussive and less about proving new results. The essential idea is to look for a Lewisian-inspired analysis of counterfactuals that can be formulated about spacetimes. The outcome is a set of perspectives along with a number of naturally arising and mathematically precise questions which could bear on the

discussion.

Contents

I. Introduction

I.1. An Outline of the Thesis

I.2. Mathematical Preliminaries

II. Black Holes

II.1. A Stronger Area Theorem

II.2. Black Hole Singularities with Achronologies

II.3. The Causal Structure of Evaporating Black Holes

III. Reasoning on the Basis of Past Lightcones

IV. Counterfactuals and Laws for Spacetimes

V. Conclusions and Future Outlook

Publication Information

The results in II have individually been the subject of three sole-authored papers published in *Classical and Quantum Gravity* and *General Relativity and Gravitation*. They have also been presented at seminars, conferences, and workshops, including for instance UC Irvine, Harvard University, Perimeter Institute, and the 2018 International Conference in Lorentzian geometry. Part III is in the process of being submitted for review, and Part IV is also likely to be in the coming future.

I. Introduction

I.1. An Outline of the Thesis

The clearest and most empirically successful description of gravity and spacetime is provided by classical spacetime theories. Among such theories, the focus here will be on its most widely studied and arguably most elegant instance: general relativity (GR). The subjects treated are foundational in nature and are divided into the following four parts:

I. an outline of the thesis and some mathematical preliminaries;

II. a collection of three independent results in black hole theory;

III. a study of the epistemic status of observers in general relativistic spacetimes;

IV. a discussion of counterfactuals in the context of spacetimes and GR.

A binding theme throughout is the study of the physical and conceptual significance of various classic mathematical results in the field, by philosophers, theoretical physicists, and mathematicians. In II and III, the aim is to describe the significance of these results, and then to formulate new ones improving or building upon them. The goal in IV isn't to produce any particular theorem, but rather to propose a set of perspectives concerning counterfactuals and laws in this setting. I shall often speak about *our own* spacetime as if it were a Lorentzian manifold, but it is of course understood that classical theories such as GR can provide no more than an approximate description of the real world.

The connection between subjects treated in II is explicit since they all concern foundational questions in black hole theory; and as for III and IV, though the former is broadly speaking epistemological and the latter more metaphysical, there is a connection in that both deal explicitly with the global structure of spacetime and the potential relevance of GR's foremost open theoretical problem — the conjecture of *strong cosmic censorship* — originally proposed by Penrose in 1979 [32].

In I, I shall introduce a number of important ideas, definitions, and results. I shall seek to explain certain aspects of their significance so as to end up with

something more informative than a mere list of definitions. For purposes later on, I shall also include a very brief section on the Partial Differential Equations (PDEs) aspect of the subject. Though this side of the subject isn't usually emphasized by philosophers, it is now the main subject of mathematical general relativity.

In II my focus will be directed entirely on black holes and various questions concerning our understanding of their behaviour. Black holes came into our scientific consciousness as predictions of classical general relativity and they have since played a key role in some of the most important theoretical advances of the latter part of the twentieth century, with many of the most important discoveries made some decades ago (individually and or jointly) by Geroch, Hawking, Penrose, and Wald. From the time of those discoveries until today, black holes have remained at the forefront of astrophysics, quantum gravity, and mathematical general relativity. They have also been of interest to a number of philosophers of physics, including though not limited to Belot, Curiel, Earman, Huggett, Manchak, Maudlin, Reutsche, Timpson, Wallace, Weatherall, and W*ü*thrich.[1]

More specifically, II contains the following three indepedent new results in black hole theory.

II.1. **A Stronger Area Theorem.**

In 1972 Hawking [13] proved what is now considered to be a fundamental result in classical GR. He showed that under certain global and local geometrical assumptions, the total area of the black holes is non-decreasing with time. To prove his theorem he used the null convergence condition, which is a local geometrical inequality satisfied by null contractions of the Ricci tensor of the spacetime, and which, by the Einstein field equations, is equivalent to the non-negativity of null contractions of the stress-energy tensor. Later in 1974 he performed a set of calculations showing that quantum field theoretic effects can cause black holes to shrink [5]. Given that his classical

[1]Curiel's recent *Stanford Encyclopedia* article offers an introductory overview of various important questions of interest to both physicists and philosophers [4].

and semi-classical descriptions are at odds with one another, it is usually understood that quantum field theory must violate one or more of the input assumptions of his area theorem. Since studies have shown that the null energy condition is violated by quantum fields, it is usually inferred that its violation is responsible for the area theorem's semi-classical failure. In view of this reasoning, it becomes of interest to know whether his area theorem can be proved with weaker assumptions, which in particular could permit for violations of the null energy condition. Here, I prove and describe a result of this kind.

II.2. **Black Hole Singularities with Achronologies.**

In 1970 Hawking and Penrose [8] proved what has long been regarded to be *the* classic and fundamental singularity theorem of GR. Their theorem shows that spacetimes are geodesically incomplete — or *singular* — under various physically reasonable conditions. Though undoubtedly deserving of its reputation, there are two drawbacks to their theorem: it says nothing about the location or type of the incomplete geodesics, and it assumes chronology — i.e., the absence of closed timelike curves — which we know to be violated in some of the most important solutions of the field equations. One example is the Kerr-Newman class of vacuum black hole solutions, whose interiors are achronological and singular, and whose exteriors are thought to represent the asymptotic state of astrophysical black holes. Here, I prove and describe a theorem that is better suited to such spacetimes than that of Hawking and Penrose. This new theorem permits locating incomplete *null* geodesics within the black hole interior, and allows for certain types of chronology violation therein.

II.3. **The Causal Structure of Evaporating Black Holes.**

In Hawking's semi-classical picture of black hole evaporation [5], [6], [7], it is usually imagined that full evaporation leaves behind a naked singularity, and thus that the resulting spacetime cannot admit a Cauchy surface. Classically,

therefore, evaporating black hole spacetimes are thought to be inhospitable arenas for prediction in the sense of initial data.

Based on this suspected predictability loss, Hawking proposed the so-called information paradox. The name 'paradox' is given in virtue of the apparent inconsistency between the transition from a pure initial state to a final mixed state (which black hole evaporation suggests) and the quantum mechanical principle of unitary evolution (which prohibits such transitions).

At the classical and semi-classical level, one can ask whether black hole evaporation and the global causal structural properties suggested by it imply a loss of predictability in the sense of initial data. If such a thing were true, it could be used as a basis for the more difficult question of semi-classical information loss. This particular question was studied by Kodama in 1979 [10] and then later by Wald in 1984 [33], and the joint work of these authors is a theorem in Lorentzian geometry which *seems* to imply that the causal structure of evaporated black hole spacetimes leads to a loss in predictability.

Here, I show that the Kodama-Geroch-Wald theorem (as it has come to be known) isn't sufficient to rule out globally hyperbolic spacetimes, i.e., there are globally hyperbolic spacetimes that satisfy the conditions of their theorem. In view of this, I formulate and describe two new theorems in Lorentzian geometry which show that evaporated black hole spacetimes (under a common construal thereof) are indeed inhospitable to prediction in the sense of initial data.[2] I shall discuss some of the consequences of this result, and I shall mention some recent work on the subject by the philosophers Maudlin, Manchak, Wallace, and Weatherall.

In III, I turn to an entirely different question and consider the epistemic status

[2]The result also resolves a conjecture of Earman from 1995 in chapter 3 of [8].

of observers in general relativistic spacetimes. That is, I ask to what extent can observers make predictions and / or draw conclusions about the global structure of their spacetime on the basis of their past lightcones. This question was first investigated in 1977 by Earman, Geroch, Glymour and Malament in a collection of articles [10]. More recently, Manchak [19], [22] has shown significantly more general results, which moreover seem to imply a strong sense of epistemic underdetermination.

My first aim in III is to better understand the meaning and the significance of Manchak's results. In doing so, I describe certain parallels between the above question and the traditional problem of induction, where by the latter I mean the usual problematic associated with extrapolating beyond some sample of observations by means of inductive arguments. As I shall seek to show, despite the terms 'induction' and 'spacetime' seeming quite distant from one another, questions about observers in spacetime can have, when cast in a certain light, a certain inductive quality about them, and pursuing this connection can lead to a better grasp of the spacetime context.

As is often the case, theorems are best understood by studying their proofs. In this respect, since Manchak's results are based on cut and paste spacetimes, it is of interest to ask whether they may be bypassed by requiring that spacetimes satisfy various seemingly physically reasonable conditions. Here, I prove various results of this kind, and describe their significance. Among other results, I make the elementary observation that when a spacetime is globally hyperbolic or causally simple, Manchak's [22] remarkable generalization of Malament's ideas is based on constructing a spacetime which fails to be hole-free and causally continuous (facts whose consequences I discuss).

In considering the notion of 'physically reasonable' spacetimes, I will be brought to consider the relevance of Penrose's conjecture of strong cosmic censorship, whose original formulation was that physically reasonable spacetimes are globally hyperbolic [32]. The potential relevance of the conjecture is that if it were true in some suitable way, then perhaps we could use it to restrict the class of admissible space-

times and thus improve the observer's epistemic status. I shall explain the extent to which this view is legitimate. I shall argue that although the view is lacking in certain respects, recent progress has meant that it is not without hope.

In IV, I turn to more metaphysical questions and consider under broadly Lewisian guidelines how to assess the truth value of counterfactuals that we can formulate about spacetimes, or parts of them. Naturally, I will also be drawn to consider Lewis' view on the role and nature of physical laws, which in this context are the laws of spacetime theories and GR in particular. The motivation to study these questions is partly to better understand spacetime counterfactuals (since these form an important part of our intuitions about spacetimes), but also to see how GR and its laws may be understood within a broad version of Lewis' Humean view. In addition, since the primary object of study in GR are spacetimes, and since these are mathematical representations of worlds (or universes), the parallel between '(possible) worlds' and '(possible) spacetimes' is already more or less present from the outset.

As for the aforementioned 'Lewisian guidelines', I will focus on the following parts of his thought:

- the counterfactual truth conditions given in *Counterfactuals* [17],
- the broadly Humean description of laws also described therein,
- his refinement given in *Counterfactual Dependence and Time's Arrow* [18].

Though his Humean Supervenience view integrates much of these ideas, I will avoid discussing it here; and I will also avoid questions about modal realism and the ontological status of possible worlds.

The emphasis in IV isn't on proving any one particular theorem (or collection thereof), but rather towards a number of perspectives as to how one might think of spacetime counterfactuals and GR laws in this broadly Humean setting, and, more tangibly, towards a list of naturally arising mathematically precise questions that could bear on the discussion.

My conclusion is overall an optimistic one. Namely, the broad aspects of Lewis' views *can* be translated into a form that is suitable for spacetimes, and, moreover, in doing so we are led to a number of further mathematical questions with definite answers that we don't yet know.

I.2. Mathematical Preliminaries

One can think of the mathematical content of general relativity as the study of a specific kind of differential geometric pair — a manifold and a Lorentzian metric on it — and the study of a certain set of partial differential equations (PDEs) obeyed by the manifold's metric. Most classical spacetime theories follow this template, and as such the subject is broadly speaking divided in two, one component dealing with the basic geometries of the manifold, and the other concerned with a detailed study of the field equations, which are a system of elliptic and hyperbolic coupled non-linear PDEs.

This first component may be grouped under the title of 'Lorentzian geometry', and in a sense it provides the background structure. The ideas involve basic differential geometric and topological definitions and the theorems are mostly straightforward and geometric in character. The scope of applicability is wide in the sense that these notions will apply to a whole host of classical theories modelled on the template of PDEs on a Lorentzian manifold.

The second, which here will be described with GR specifically in mind, is the study of the field equations. This aspect is where the PDEs side of the subject come to the fore, and the mathematics is essentially a blend of geometry and PDE theory. The theorems are much more technical and are often less easy to interpret than those above, which tend to be more elementary and geometric.

In very brief fashion and focusing only on what is necessary for what follows, I shall describe these two complementary aspects. Naturally, I assume some familiarity with elementary notions in differential geometry, general relativity, and topology. With regards to the first aspect, the book of Malament [27] and the brief yet beautiful article of Manchak [30] are complementary. As for the second aspect, since I will only describe the PDEs side of the subject at a very broad level and the reader is encouraged to consult some of the suggested references for more details. My reason for mentioning the PDE aspect of the subject is because it dominates the current mathematical study of general relativity.

Far more detailed and complete accounts of what is to follow can be found in the various classic references. For smooth manifolds and aspects of differential geometry there is the very readable and comprehensive textbook [23], for Riemannian geometry [34], for Lorentzian geometry [2], [15], for general relativity [4], [33], and for an introduction to the PDE aspects of general relativity [4], [13], [37], [29]. An easy going introduction to global causal structure is given in [15], and one that is well suited for philosophers is [30].

1. Elementary Lorentzian geometry

Everything that is to follow will take place in the context of a spacetime, defined as follows.

Definition. *A spacetime (M, g) is a connected C^∞ Hausdorff manifold of dimension two or greater which has countable basis, a Lorentzian metric g of signature $(-, +, +, ...)$, and a time and space orientation.*

Taking spacetimes to be time and space orientable isn't necessary. The amendments to do analysis on non-orientable spacetimes aren't always straightforward, though sometimes we can simply pass to the oriented covering spacetime as in chapter 3 of [2].

These basic topological choices for M are chosen because analysis on M would quickly become intractable without them. Moreover, the idea of living in a spacetime with boundary and or in a disconnected spacetime is difficult to make sense of. As for the property of Hausdorffness, it plays an important role in a number of applications and although it could be problematically restrictive, the kind of routine geometric analysis that is carried out on (M, g) would be virtually impossible without it.

The lightcone structure is inherited on the tangent space T_pM at every point $p \in M$. A vector $v \in T_pM$ is called spacelike, null, or timelike, when, respectively, $g(v, v) > 0, = 0, < 0$. This leads to the definition of a causal curve in M.

Definition. *A piecewise smooth curve γ is a piecewise smooth map $\gamma : \mathbb{R} \to M$. Its character at any point $q \in \gamma$ is determined by the causal character of its tangent vector; that is, a curve is timelike at p if its tangent vector at p is timelike.*

Given the above definition we define the chronological / causal past of a point p by $I^-(p)$ / $J^-(p)$, as the set of points q which can be reached by piecewise smooth timelike / causal curves with future endpoint p. It is a theorem (cf. chapter 3 of [2]) that $I^{+(-)}(S)$ is an open set for any $S \subset M$. The same cannot be said of $J^{+(-)}(S)$, which may be neither open nor closed, depending on the causal structure of M.

An observer O in a spacetime M will be represented by a timelike curve γ which we will take to have future endpoint. At any given time O will be represented by a point on γ.

The domain of dependence $D(S)$ of a set S is denoted by $D(S) = D^+(S) \cup D^-(S)$, where $D^+(S)$ is defined as the set of points $\{q\} \subset M$ such that any past directed causal curve from q intersects S. The idea here is that signals propagate within the lightcone, and thus that $D^+(S)$ tracks those sets of points which depend on what was registered on S.

2. The causal hierarchy of spacetimes

The causal hierarchy (or causal ladder) is a list of definitions which capture the causal structure of (M, g). The standard contemporary reference is [31]. The point of the causal hierarchy is to classify the various causal properties that spacetimes can have. Spacetimes with inhabitual causal structure can arise in many different ways. Sometimes they are trivially constructed: take n-dimensional Minkowski spacetime and compactify the $\mathbb{R}$ factor in the timelike direction. Other times however they are more serious and unavoidable features of certain solutions to the field equations. As we shall see, the Kerr-Newman class of vacuum black hole solutions contain closed timelike curves in their interiors. The infamous Misner, Gödel, and Taub-NUT spacetimes provide other examples of acausal solutions. The next few definitions

are but a subset of the usual total set of definitions, and they will appear in order of strength, i.e., latter definitions imply former ones but not the other way around.

Definition. *A spacetime (M, g) is non-totally vicious if there doesn't pass a closed timelike curve through every point in M. It can be shown that a totally vicious spacetime is one in which there is a point $p \in M$ such that $M \subset I^-(p)$. When such a point exists, the spacetime is said to be causally bizarre. Totally-vicious and causally bizarre aren't equivalent however, Misner spacetime is causally bizarre but not totally-vicious.*

An important example of a totally vicious spacetime other than the above one is Gödel spacetime; see the classic reference [14] for a description of its structure.

Definition. *A spacetime (M, g) is chronological when it has no closed timelike curves. It is causal when it has no closed causal curves.*

Causality obviously implies chronology but a spacetime can be acausal though not achronological. As we shall see in more detail in II.2., recent progress in Lorentzian geometry has led to a new definition for the boundary of chronology violating region, which incidentally is necessarily an open set.

Definition. *A spacetime (M, g) is non-totally imprisoning when there are no future or past inextendible causal curves contained within a compact set $K \subset M$.*

Just above non-total imprisoning in the hierarchy is the distinguishing property.

Definition. *A spacetime is distinguishing iff it is future and past distinguishing. It is future distinguishing iff $I^+(p) = I^+(q) \Rightarrow p = q$, and analogously in the past direction.*

Just above distinguishing is the property of being strongly causal, which is a pointwise definition.

Definition. *A spacetime is strongly causal iff it is strongly causal at all $p \in M$, and a spacetime is strongly causal at p if the following holds. Given any neighborhood U of p, there exists a neighborhood $V \subset U$, $p \in V$, such that any future-directed (and hence also past-directed) causal curve $\gamma : I \to M$ with endpoints in V is entirely contained in U. The term $I^+(p;U)$ means the chronological future of $I^+(p) \cap U$ where is a neighborhood of p and a smaller spacetime centred on p.*

The idea here is that a causal curve does not leave an arbitrary small neighborhood and come back to it. That is, no point p of M lies on an *almost* closed causal curve (which would otherwise approach p upon leaving some arbitrarily small neighborhood of p).

Definition. *A spacetime (M, g) is stably causal iff there exists a g' such that $g < g'$ and g' is causal. Here (and elsewhere) '$g < g'$' signifies that the lightcones of g are everywhere strictly narrower than those of g'.*

Stable causality also has alternative definitions, and the condition turns out to be equivalent to that of admitting a time function, i.e., a continuous function $t : M \to \mathbb{R}$ which is strictly increasing on any future-directed causal curve γ.

Definition. *A spacetime is causally continuous iff it is reflecting and distinguishing. A spacetime is reflecting iff it is future and past reflecting, and a spacetime is past reflecting at q iff $q \in \overline{I^+(p)} \Rightarrow p \in \overline{I^-(q)}$.*

There are alternative definitions of causal continuity, and indeed these will be relevant when we come to consider the subject of evaporating black holes in II.3.

Definition. *A spacetime is causally simple iff it is causal and $J^-(+)(p)$ is closed for all $p \in M$.*

Causal simplicity is just one below global hyperbolicity, which is the final and strongest definition in the hierarchy.

Definition. *A spacetime is globally hyperbolic iff it is causal and* $J^+(p) \cap J^-(q)$ *is compact for all* $p, q \in M$*.*

For reasons we shall discuss below, global hyperbolicity is probably the most important definition in the causal hierarchy. In virtue of being the strongest, it also has the most consequences. The concept has a number of equivalent definitions, each revealing slightly different aspects of the concept.

The above definition refers directly to the lightcone structure of (M, g), but due to a theorem of Geroch [13], the above sense of global hyperbolicity is equivalent to the spacetime admitting a Cauchy surface — a notion that will be defined below. An alternative to Geroch's results was put forward by O'Neill [15], and indeed it is his version of this result that we now follow. This approach differs only slightly from some of the more classic references [33] and [14].

Definition. *A set is achronal (acausal) iff no two members of it can be connected by a timelike (causal) curve.*

Definition. *A subset* S *of a* n*-dimensional topological manifold* M *is a topological hypersurface provided that for each* $p \in S$ *there is a neighborhood* $\mathcal{U}$ *of* p *in* M *and a homeomorphism* ϕ *of* $\mathcal{U}$ *onto an open set in* $\mathbb{R}^n$ *such that* $\phi(\mathcal{U} \cap S) = \phi(\mathcal{U}) \cap \Pi$*, where* Π *is a hyperplane in* $\mathbb{R}^n$*.*

Definition. *Given an achronal set* A*, the edge of* A*, denoted* $edge(A)$*, consists of all points* $p \in \overline{A}$ *such that every neighborhood* U *of* p *contains a timelike curve from* $I^-(p; U)$ *to* $I^+(p; U)$ *which does not meet* A*.*

Proposition. *An achronal set* A *is a closed topological hypersurface iff* A *contains no edge points.*

Definition. *A Cauchy hypersurface in* M *is a subset* S *that is met exactly once by every inextendible timelike curve in* M*.*

Note that the above definition is merely about how S sits in the spacetime. It is not yet shown that a Cauchy hypersurface is indeed a topological hypersurface.

Proposition. *A Cauchy hypersurface S is a closed achronal topological hypersurface and is met by every inextendible causal curve.*

Theorem. *A spacetime (M, g) is globally hyperbolic iff it admits a Cauchy hypersurface.*

In Wald's classic textbook [33], he starts by defining a *slice* — an achronal and edgeless set S — and he then says that a spacetime is globally hyperbolic if $M = D(S)$. If indeed $M = D(S)$ then he says that S is a Cauchy surface for M. It is not too difficult to show that this sense of Cauchy surface is indeed a Cauchy hypersurface in the sense of O'Neill.

There is yet another definition of global hyperbolicity due to Leray [24] whereby one demands that the space of causal curves relating any two points in the spacetime be compact under a suitable topology.

An important consequence of global hyperbolicity, first proven by Geroch [13], is that it forces the spacetime to have topology $\mathbb{R} \times S$ where S is a Cauchy (hyper)surface. As a consequence, for any globally hyperbolic spacetime, its Cauchy (hyper)surfaces all have the same topology.

In fact, a stronger more recent result [2] shows that the spacetime (M, g) splits *smoothly* so that $(M, g) = (\mathbb{R} \times S, -f(t)dt^2 + h(t))$ where $f(t)$ is smooth and $h(t)$ is a possibly evolving Riemannian metric on S

Theorem. *A globally hyperbolic spacetime (M, g) is isometric to a product manifold $(\mathbb{R} \times S, -f(t)dt^2 + h(t))$ where $f(t)$ is a smooth function whose level sets are smooth spacelike Cauchy hypersurfaces with induced Riemannian metric $h(t)$.*

The topological splitting had been known since Geroch [13], but the smoothness of $f(t)$ had been open until a few years ago [2].

3. Aspects of the PDE sector

There are two fundamental theorems in general relativity which ground the modern viewpoint on the theory. In short, these theorems embody the fact that

general relativity is well posed as an initial value formulation. Here, the sense of well posedness is that of Hadamard, i.e., whether solutions exist, whether they are unique, and moreover whether the solutions depend continuously on the initial data in some appropriate sense. Note that in GR, the uniqueness here will only be 'up to isometry', and that strictly speaking as Hadamard well posedness is not satisfied.

All that follows derives from the field equations, originally proposed by Einstein in 1915, which are schematically written as follows

$$\mathrm{Ric}(g) - \frac{1}{2}\,\mathrm{Scal}(g)g \equiv G(g) = T[\phi, ...]$$

Here, the cosmological constant term $+\Lambda g$ has been omitted, the Einstein tensor $G(g)$ of g is defined in terms of the Ricci and the scalar curvature tensor of g, and the stress energy tensor associated with a collection of matter fields $(\phi, ...)$ is denoted $T[\phi, ...]$.

These equations are daunting in many ways but for an $(n+1)$-dimensional spacetime (M, g) they may be split into an initial value formulation, whereby one thinks of an n-dimensional spacelike hypersurface N embedded into M. In order for N to embed into a spacetime (M, g) that satisfies the field equations, the submanifold itself has to satisfy a number of equations. These correspond to the Gauss and Codazzi equations for the embedding of N into M, and the relevant data is composed of Riemannian metric h (induced onto N by g and the embedding $N \hookrightarrow M$) and the second fundamental form k on N, a symmetric two tensor describing how N varies locally in M. These equations are known as the constraint equations, and letting u be the unique (up to normalisation) timelike vector normal to N, they take the following form

$$\mathrm{Scal}(h) - (\mathrm{trace}_h\, k)^2 + |k|^2 = \mu \equiv T(u, u)$$

$$\mathrm{div}_h(k - (\mathrm{trace}_h(k)h) = J \equiv T(u, \cdot)$$

The evolutionary part is a little more cumbersome to write and so we avoid it here.

3.1 Local well posedness

As mentioned, the constraint equations are necessary conditions for a spacelike hypersurface to embed into a spacetime. Due to a fundamental theorem of Choquet Bruhat [3] in 1954, it turns out that the constraint equations are also *sufficient* to guarantee the local existence of a spacetime (M, g) into which the initial data set (N, h, k) embeds, and acts as a Cauchy surface for that spacetime. That is, given any triple (N, h, k) satisfying the above constraint equations along with some suitable energy condition, there exists a unique spacetime (M, g) into which (N, h, k) isometrically embeds. For any spacetime (M, g) arising in this way, we say that it is a local Cauchy development of (N, g, k).

This theorem is in a sense the first fundamental PDE theorem of general relativity, and it has been refined over the years. In the original work of Choquet Bruhat, a key role was played by the so-called 'wave gauge'. Picking this gauge amounts to choosing a certain co-ordinate system for the field equations, and expressed as such, the field equations take on a certain form which in turn allows one to use existence and uniqueness theorems from PDE theory. This reformulation of the field equations was to some extent the key observation of Choquet Bruhat.

In order to prove the existence and uniqueness theorems, one assumes that the initial data h, k satisfy various suitable regularity and differentiability requirements. Over the years, these have been made gradually weaker.

The nature and reason for these assumptions vary but the ultimate point is that one needs these fields to be sufficiently well behaved so that the relevant techniques from PDE theory can be used. These general PDE theorems concern more general classes of equations than merely the field equations of GR (with or without matter sources). That isn't to say however that the field equations of GR are undistinguished among other hyperbolic-elliptic PDE systems. Quite the contrary, the field equations of GR benefit from a number of properties which other superficially

sembling systems of equations do not. This is a subtle point however, and in any case we shan't need to understand it further here.

The final piece of Hadamard well posedness is continuity. That is, whether the local spacetime (M, g) around (N, h, k) (which the above theorem guarantees) depends continuously the field values h, k on N. The sense of continuity is expressed with respect to a certain topology, which in turn depends on what spaces the fields h, k are assumed to live in. Skipping the details, it turns out that the Cauchy development or local spacetime (M, g) does indeed depend continuously on the initial data; cf. Part III of [37] for a precise expression of this fact.

Continuity is important for it allows knowledge of the initial data set to be approximative in some sense. Moreover, continuity of the spacetime development guarantees that two initial data sets that are sufficiently similar to one another (in the relevant sense), produce local Cauchy developments which are also in a certain sense similar.

3.2 Maximal globally hyperbolic developments

Among the most fundamental theorem of GR is the existence and uniqueness of a maximal globally hyperbolic Cauchy development associated with any initial data set. This is a spacetime that is both unique up to isometry and continuously determined by the initial data set. The theorem is stated below, though before we make a note on the role of energy conditions.

The theorem can be proved in vacuum or one can assume the dominant energy condition, which is the requirement that for all timelike vectors u

$$T(u, u) \geq 0$$

and that $T(u, \cdot)$ (with indices suitably raised) is a non-spacelike vector. This condition is sufficient but not necessary. Weaker energy conditions also suffice but these don't seem to play an important role elsewhere, and their physical motivation

doesn't seem to be quite as strong as the dominant energy condition, and so we avoid delving further into this; see chapter 4 of [14] for a little more information on this issue, and [41] for a more contemporary account of the dominant energy condition.

We can now state the theorem as follows, where in particular the notion of a Sobolev space will be defined below.

Theorem. *Let (N, h, k) be any initial data set that satisfies the constraint equations, the dominant energy condition, and suppose that h, k is smooth. Then there exists a globally hyperbolic spacetime (M, g) with Cauchy surface N, unique up to isometry, that satisfies the vacuum Einstein field equations, and that is maximal in the sense that if (N, h, k) can be isometrically embedded into a globally hyperbolic spacetime (M', g'), then (M', g') isometrically embeds into (M, g).*

The version of this theorem involving matter fields is a subtle point that we will not go into here. The proof of this theorem is based on the previous local well posedness result. The original argument by Choquet Bruhat and Geroch [2] in 1969 assumed the initial data to be smooth and it involved some Lorentzian geometry as well as the axiom of choice. A fuller account of their proof was proposed by Ringstrom in [37], but unfortunately his argument contained a mistake. This was corrected by Ringstrom in his later comprehensive and authoritative treatise on cosmology and the Cauchy problem [29]. Chruściel [6] has also given a proof of the theorem under much lower regularity of the initial data. His result require the tools of lower regularity Lorentzian geometry, originally developed by Chruściel and Grant [7].

All of these theorems' proofs rely on the axiom of choice in a rather direct way. Owing to the enigmatic status of this axiom, it had been of interest to find a more constructive proof without it. This was recently achieved by Sbierski [39] who has been able to give a direct proof of it assuming the initial data is smooth along with a significantly weakened version of the axiom of choice. It may be that Sbierski's

argument can be generalized to deal with lower regularity initial data, and indeed that would constitute the most satisfying version of the above theorem.

3.3 The space of initial data

In order to prove theorems about initial data sets and spacetimes that satisfy the Einstein field equations, the fields that appear in the equations must live in the right functional spaces. When they live in the right spaces, the machinery of PDE theory can be unleashed onto the field equations, and it is that which in turn leads to the fundamental theorems that were briefly described above. At the most general level, these functional spaces are Banach spaces — i.e., complete normed vector spaces — whose norms give a notion of relative size and distances within the space. I shall now describe two important types of spaces, though for simplicity I shall only consider functions on $\mathbb{R}^n$, rather than on differentiable manifolds. The first types of spaces are in a sense more elementary, whereas the latter are a little closer to what is actually used in mathematical relativity. Note that what follows is highly condensed and that far more exhaustive and detailed accounts exist in the aforementioned references.

C^k spaces. We first consider $C^k(\mathbb{R}^n)$ spaces, i.e., the spaces of k-times differentiable functions on $\mathbb{R}^n$. Restrict now to $C^k_b(\mathbb{R}^n)$, i.e., the elements of $C^k(\mathbb{R}^n)$ which, together with all their derivatives of order less than or equal to k, are bounded in all of $\mathbb{R}^n$. Then $C^k_b(\mathbb{R}^n)$ has a natural norm given by

$$\|u\|_{C^k_b} = \sup_{\overline{x}\in\mathbb{R}^n} \sum_{|\alpha|\leq k} |\partial^\alpha u(\overline{x})|$$

Note here the norm tracks all the derivatives of u. Consequently if we are asked to produce a physical interpretation that u and v be close in that norm, we look for quantities that somehow track and treat with equal importance all the derivatives. This is quite different to the idea that the value of two functions on $\mathbb{R}^n$

be close in the sense of their absolute values being within some ϵ of each other. Consider for instance a function representing the temperature or density of some fluid. Two such functions being close to one another in the ϵ sense aforementioned has a simple interpretation: the temperature or density profiles are close. This ignores how these profiles are varying. The above norm for $C^k_b(\mathbb{R}^n)$ is in a sense finer in the sense that it tracks and treats with equal importance all derivatives of u.

As it turns out, spaces of the above type are not those which appear in the modern existence and uniqueness results of general relativity. For that purpose we need slightly more refined notions which we illustrate with the following simplified example.

H^k spaces. To define $H^k(\mathbb{R}^n)$ spaces, we first define $L^p(\mathbb{R})$ spaces.

Definition. *For $1 \leq p < \infty$, $L^p(\mathbb{R}^n)$ is defined as the set of all measurable functions u from $\mathbb{R}^n$ to $\mathbb{R}$ such that*

$$||u|| = (\int_{\mathbb{R}^n} |u|^p dx)^{1/p} < \infty$$

We now define the notion of weakly differentiable functions on $\mathbb{R}^n$; for some of the terminology used below refer to [37].

Definition. *A function $u \in L^1(\mathbb{R}^n)$ is k times weakly differentiable if for every multiindex α with $\mid \alpha \mid < k$ there is a function $u_\alpha \in L^1_{loc}(\mathbb{R}^n)$ with the property that*

$$\int_{\mathbb{R}^n} u \partial^\alpha \phi dx = (-1)^\alpha \int_{\mathbb{R}^n} u_\alpha \phi dx$$

for all $\phi \in C^\infty_0(\mathbb{R}^n)$. The functions u_α are referred to as the weak derivatives of u. Here, $L^1_{loc}(\mathbb{R}^n)$ denotes the space of functions that are L^1 on every compact set on which the function is defined.

We can now define the Sobolev spaces $W^{k,p}(\mathbb{R}^n)$ and $H^k(\mathbb{R}^n)$, the latter of

which corresponds to the special case $p = 2$.

Definition. *Let $1 \leq p < \infty$ and let $\mathcal{W}^{k,p}(\mathbb{R}^n)$ denote the set of k times weakly differentiable functions such that all the weak derivatives are in $L^p(\mathbb{R}^n)$. Let $W^{k,p}(\mathbb{R}^n)$ denote the set of equivalence classes of elements in $\mathcal{W}^{k,p}(\mathbb{R}^n)$, two elements being equivalent if the set on which differ has measure zero. For an element u of $\mathcal{W}^{k,p}(\mathbb{R}^n)$ we define the following norm*

$$||u||_{W^{k,p}} = (\Sigma_{|\alpha| \leq k} \int_{\mathbb{R}^n} |\partial^\alpha u|^p dx)^{1/p}$$

Although these definitions look rather technical, there is a story behind their physical interpretation. To see why there are integrals of powers of derivatives of functions, consider the simplest case of the wave equation $\Box u = 0$ in 2-dimensional Minkowski space $(\mathbb{R}^2, \eta)$. In such contexts, we take the energy associated with such a solution to be the integral of squares of first order derivatives of u, i.e., $\int |\partial_t u|^2 + |\partial_x u|^2$. In virtue of $\Box u = 0$, this quantity is conserved in the evolution of u. For wave equations that differ from $\Box u = 0$, the above 'energy' is not in general conserved. However, for equations sufficiently similar to '$\Box u = 0$', it can be shown that departures from conservation remain small. The sense of smallness is naturally tracked by the above kinds of norms.

In fact, this phenomenon highlights one of the reasons why H^k type spaces end up appropriate for the Cauchy problem of GR. The reason being related to the fact that the field equations of GR (expressed in a suitable co-ordinate system) are quasi-linear hyperbolic PDEs (i.e., a more general type of wave equation). H^k type spaces are thus more suitable than C^k_b type spaces because they are sensitive to the wave character of the field equations.

In passing, we note that another important reason why H^k type spaces are suitable in GR (and for wave equations more generally) is that the associated degree of regularity is preserved in the evolution. In fact, it is a standard phenomenon that wave equations can degrade the regularity of the initial data. To get around this

problem, and in particular to show statements about the continuity of solutions from initial data, one needs to include weakly differentiable functions and an associated norm that allows comparison between these functions.

4. The physical significance of the lightcone

General relativity and alternative theories of gravity are basically based on the same template: a Lorentzian manifold on which are satisfied some set of field equations. The physical significance of the geometry is loosely speaking inherited from the fact that g captures relative notions of distances and time intervals. Although there are various questions of interpretation that one can pose, it is worth noting that the physical significance of the geometrical concepts so far defined (eg., lightcones, curves, domains of dependence, etc.) depend on the underlying field equations. The reason for this can be broadly understood as follows.

Any system of hyperbolic partial differential equations naturally leads to the posing of a Cauchy problem, and, consequently, to the notion of characteristic hypersurfaces. These characteristic hypersurfaces can be defined precisely but we shall avoid the digression since they won't be of any real use to us here. For our purposes, it suffices to think of them as entities defined with respect to a system of hyperbolic equations and an initial data set which track the realm of influence that disturbances in the initial data set can have in the development of that data set.

When the development of some initial data set takes place in a Lorentzian manifold (M, g), there is also a background spacetime metric g associated with the Lorentzian manifold. For a system of hyperbolic PDEs evolving in a Lorentzian manifold, these characteristic hypersurfaces (which are sensitive to the details of the field equations) will differ in general to the null hypersurfaces of the spacetime metric. Indeed, such is true for many alternative theories of gravity. In vacuum general relativity, characteristic hypersurfaces and null hypersurfaces coincide and to some extent this is what accounts for the fundamentality of the above geometric definitions, for instance like that of the domain of dependence $D(S)$. In case of fail-

ure of this coincidence, it is still possible to define a domain of dependence concept that is intrinsic to the PDE system. In general however, this concept won't match with that defined purely in terms of the lightcone structure.

The effect of this possible discrepancy between lightcone structure and PDE structure is best visualized in the context of alternative theories of gravity. As described and studied in [35], the propagation of gravitational disturbances (in particular gravitational waves) needn't obey what we ordinarily think of as limits imposed by the lightcone structure. This subtlety was approached from a more foundational angle by Geroch [14], and later with somewhat different motivations by Earman [9].

5. The non-globally hyperbolic Cauchy problem

Throughout the thesis I shall often use the fact that globally hyperbolicity along with some energy conditions is sufficient for a form of spacetime determinism (in the sense of developing initial data). Though it won't make a difference to the main arguments that I will be making, it ought to be said that these conditions are sufficient but not always necessary. Indeed, there are examples of non-globally hyperbolic spacetimes in which a Cauchy problem is well posed in the sense of Hadamard.

With that said, it is hard to know how seriously we ought to take cases of non-globally hyperbolic determinism. The results which are available are few and far between, and moreover they either parasitic on the globally hyperbolic case, or they rely on fairly specific set-ups otherwise absent in the context of globally hyperbolic spacetimes.

For an example of the former, consider the case of temporally non-orientable (and thus non-globally hyperbolic) spacetimes, considered for instance in [8]. There, one constructs a class of such spacetimes and one shows that determinism (in the relevant sense) holds. The argument however is entirely based on the globally hyperbolic case. It proceeds by considering the oriented double cover of the non-

orientable spacetime, which itself is globally hyperbolic, and then it projects the well posedness of the covering spacetime (using the natural covering map) onto the non-orientable spacetime.

The next best examples of non-globally hyperbolic determinism are the chronology violating, asymptotically flat, stationary, four dimensional spacetimes constructed in [11]. In [11], the authors prove what look like respectable existence and uniqueness theorems for the massless wave equation existing in these pathological spacetime backgrounds. In both of these the data is prescribed on the spacetime's conformal boundary 'scri' (i.e., $\mathcal{J}$) and the solution is shown to exist uniquely in the interior. Continous dependence is not shown in general though it is likely to hold under various further assumptions on the initial data. In these examples, it is assumed that the spacetime is topologically $\mathbb{R} \times \Sigma$ and that each cross section Σ is flat outside a compact set. This assumption is important because from the rigidity part of the positive mass theorem, under the dominant energy condition, any asymptotically flat spacelike hypersurface which is flat outside a compact set is in fact everywhere flat. Since Σ as it is used in [11] is non-flat, no such Σ can arise as the end to a spacelike hypersurface with boundary for which the dominant energy condition is satisfied. Thinking general relativistically, the T_{ab} required for these examples badly violates the dominant energy condition. The situation in [11] is actually even more non-general relativistic in that the spacetime is fixed and it is the massless wave equation whose Cauchy problem is studied. In particular, closed timelike curves are not shown to form dynamically. Rather, they are assumed to exist and the spacetime is then used as a fixed arena to study certain PDEs.

The lesson to be drawn from these examples is that although non-globally hyperbolic spacetimes can be arenas for respectable Cauchy problems for the massless wave equation, they by no means stand on equal footing as those globally hyperbolic spacetimes. It is nevertheless worth being aware of the fact global hyperbolicity isn't strictly speaking necessary for Hadamard well posedness and the initial data sense of determinism so far described. In his 1995 book [8] Earman describes a set of ques-

tions on the subject of predictability and determinism in the context of non-globally hyperbolicity which to my knowledge remain unanswered.

References

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

[2] Bernal, A., and Sánchez, M., Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. Commun. Math. Phys. 257, 43–50 (2005).

[3] Choquet-Bruhat, Y., Théoréme d'existence pour certains systémes d'équations aux dérivées partielles non linéaires. Acta Math. 88 (1952), 141–225.

[4] Choquet-Bruhat, Y., General Relativity and the Einstein Equations, Oxford Science Publication: 2009

[5] Choquet-Bruhat Y., Geroch, R., (1969), Global aspects of the Cauchy problem in general relativity Commun. Math. Phys. 14 329–35

[6] Chruściel, P., (2011), On maximal globally hyperbolic vacuum space-times, Journal of Fixed Point Theory and Applications 14 (2013), pp. 325-353

[7] Chruściel, P., Grant, J., On Lorentzian causality with continuous metrics, Classical and Quantum Gravity (2012) 29 145001

[8] Curiel, E., Singularities and Black Holes, The Stanford Encyclopedia of Philosophy (Spring 2019 Edition)

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

[10] Earman, J., No superluminal propagation for classical relativistic and relativistic quantum fields, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, Volume 48, Part B, 2014, Pages 102-108.

[11] Friedman, J., The Cauchy Problem on Spacetimes That Are Not Globally Hyperbolic, in The Einstein Equations and the Large Scale Behavior of Gravitational Fields, 50 Years of the Cauchy Problem in General Relativity, 2004, Birkhauser Basel, Editors: Chrusciel, Friedrich

[12] Foundations of SpaceTime Theories, ed. J. Earman, C. Glymour, and J. Stachel, 61-80. Minnesota Studies in the Philosophy of Science 7. Minneapolis: University of Minnesota Press (1977)

[13] Geroch, R., Domain of dependence. J. Math. Phys. 11, 437–449 (1970)

[14] Geroch, R., 2011, Faster than light? In: Plaue, M., Rendall, A., Scherfner, M. (Eds.), Advances in Lorentzian Geometry. American Mathematical Society, Providence, RI, pp. 59–80.

[15] Geroch, R., Horowitz, G. (1979). Global Structure of Spacetimes, in S. W. Hawking and W. Israel (eds), General Relativity: An Einstein Centenary Survey. Cambridge: Cambridge University Press, 212-293.

[16] Hawking, S. W., Black holes in general relativity. Comm. Math. Phys. 25(2), 152–166 (1972)

[17] Hawking, S. W., Black hole explosions. Nature 248 (5443), 30-31 (1974)

[18] Hawking, S. W., Particle creation by black holes. Communications in mathematical physics 43 (3), 199-220 (1975)

[19] Hawking, S. W., Breakdown in predictability in gravitational collapse. Physical Review D 14 (10), 2460, (1976)

[20] Hawking, S. W. and Ellis, G. F. R.: The Large Scale Structure of Space-Time. Cambridge: Cambridge University Press (1973)

[21] Hawking, S. W ., Penrose, R., Proc. Roy. Soc. London A 3 14, 529, (1970)

[22] Kodama, H., Inevitability of a naked singularity associated with the black hole evaporation. Progress in Theoretical Physics 62, 1434-1435, (1979)

[23] Lee, J., Introduction to Smooth Manifolds, Graduate Texts in Mathematics, 2003

[24] Leray, J., Hyperbolic differential equations, Lecture notes, Princeton University, 1952

[25] Lewis, D., Counterfactuals, Oxford: Blackwell Publishers and Cambridge (1973)

[26] Lewis, D., Counterfactual Dependence and Time's Arrow, Noûs, 13: 455–476 (1979)

[27] Malament, D., Topics in the Foundations of General Relativity and Newtonian Gravitation Theory, University of Chicago Press 2012

[28] Manchak, J., Is prediction possible in general relativity? Foundations of Physics, 38(4): 317-21, (2008)

[29] Manchak, J. Can We Know the Global Structure of Spacetime? Studies in History and Philosophy of Modern Physics 40:53–56, (2009)

[30] Manchak, J., Global Spacetime Structure, (2011), The Oxford handbook of philosophy of physics

[31] Minguzzi, E., Sánchez, M., (2008), The causal hierarchy of spacetimes Recent Developments in Pseudo-Riemannian Geometry (ESI Lect. Math. Phys.) (Zürich: Eur. Math. Soc.) pp 299–358

[32] O'Neill, B.: Semi-Riemannian Geometry. San Diego: Academic Press (1983)

[33] Penrose, R., Singularities and Time Asymmetry, In General Relativity. An Einstein. Centenary Survey (ed. S.W. Hawking and W. Israel.) Cambridge University Press, (1979)

[34] Petersen, P., Riemannian Geometry, 2nd edition, Graduate Texts in Mathematics 2006

[35] Reall, H., Tanahashi, N., Way, B., Causality and Hyperbolicity of Lovelock Theories, Class.Quant.Grav. 31 (2014) 205005

[36] Rendall, A., Partial Differential Equations in General Relativity, (2008) New York: Oxford University Press

[37] Ringstrȯm, The Cauchy Problem in General Relativity, European Mathematical Society, 2009

[38] Ringstrȯm, H., On the topology and future stability of the universe, Oxford University Press, Oxford, 2013, Oxford Mathematical Monographs

[39] Sbierski, J., On the existence of a Maximal Cauchy Development for the Einstein Equations - a Dezornification, Annales Henri Poincare 17 (2016) no.2, 301-329

[40] Wald, R. M., Black holes, singularities and predictability. In: Christensen, S. M. (Ed.), Quantum theory of gravity. Essays in honor of the 60th birthday of Bryce S. DeWitt. CRC Press, Boca Raton, FL, pp. 160-168, (1984)

[41] Wong, W. W-Y., Regular hyperbolicity, dominant energy condition and causality for Lagrangian theory of maps, Class.Quant.Grav. 28: 215008, 2011

II. Black Holes

II.1. A Stronger Area Theorem

1. Introduction

Many classic theorems of relativity are obtained by positing a number of local conditions on the geometry of spacetime. These geometric conditions are inequalities imposed, by fiat, on certain contractions of the Einstein or Ricci tensor. With the use of Einstein's equations, these geometric conditions become energy conditions that are supposed to faithfully capture the energetic properties of whatever matter fields happen to be in the spacetime.

In the actual mathematical arguments which will be presented, I will only write the geometric form of these conditions, so as to emphasize their logic, even if it is of course understand that they are motivated by the Einstein field equations, and the properties of physically reasonable matter fields.

When thought of as conditions on the actual matter fields as opposed to simply local geometric conditions, it has become known that these local energy conditions are violated by a number of classical matter models and, moreover, that violations are ubiquitous in the context of quantum field theory in both flat and curved spacetime.[3] As such, many if not all of the classic theorems aforementioned rely on assumptions that are not satisfied in physically relevant contexts. As such, it is natural to ask whether, for certain theorems of interest, they can be strengthened somehow.

My purpose here is two-fold. First, to show that Hawking's area theorem can be strengthened as such, and, with the semi-classical context in mind, to interpret the result presented. The definitions used from now on will follow the standard text [33].

It is instructive, before delving into the area theorem, to consider the well known singularity theorems. There is a common template to these theorems. Their assumptions usually include an energy condition, a restriction on the causal properties of the spacetime, and an initial or boundary condition, and their conclusions almost

[3] See [4] for a foundational perspective on energy conditions.

always involve no more than the failure of non-spacelike geodesic completeness. Moreover, these theorems invariably rely on arguments involving Raychaudhuri's equation, which is an ODE that describes the expansion of geodesic congruences. In the four dimensional null irrotational case it reads

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \sigma^2 - \mathrm{Ric}(k,k)$$

with λ the affine parameter along the geodesic, σ the shear, θ the expansion and $\mathrm{Ric}(k,k)$ the Ricci tensor twice contracted with a null vector k tangent to the geodesic. Assuming the null convergence condition, $\mathrm{Ric}(k,k) \geq 0$, or, with Einstein's equations in four dimensions, the null energy condition (NEC), $T(k,k) \geq 0$, it follows that if the expansion θ satisfies $\theta(\lambda_0) < 0$, then $\theta \to -\infty$ within finite affine parameter $\lambda \in (\lambda_0, \infty)$. This behavior is sometimes referred to as geodesic focusing.

The onset of geodesic focusing signals the failure of certain geodesics to satisfy certain properties which would otherwise be associated with them. In the null case, a null geodesic focusing to the future of a point signals the geodesic's failure to remain on the boundary of the causal future of that point, and in the timelike case, focusing signals the geodesic's failure to maximize proper time.

Many proofs of singularity theorems work by setting up a contradiction under the assumption that all null or timelike geodesics are complete. The rough template is as follows. Assume that all null (or timelike) geodesics are complete and deduce, under the energy and boundary or initial conditions, the onset of geodesic focusing. Then, combining causal restriction and initial or boundary condition, show that the focusing produced leads to a contradiction. Deduce, therefore, that not all null (timelike) geodesics can be complete.

Tipler [14] was among the first to show that singularity theorems may be strengthened by way of weaker energy conditions. He defined these weaker energy conditions as non-local restrictions on the integral, along certain types of null

or timelike geodesics, of various contractions of the Ricci or Einstein tensor. He showed that these conditions were sufficient to cause focusing, and, thus, he was able to strengthen certain singularity theorems without major amendments to the original arguments.

His observation was developed by many authors, and, in time, there arose a number of various weaker energy conditions, now falling under the umbrella term 'average energy conditions'. One example that continues to generate interest is the *average null energy condition*.[4] This is, roughly speaking, the requirement that

$$\int_\gamma \mathrm{Ric}(k,k)d\lambda \geq 0$$

for some suitable class of null geodesics $\{\gamma\}$ with k a null vector tangent to the geodesic.

Though mathematically weaker, the physical interpretation of these average energy conditions is still murky at best. Over the course of a long list of studies, it has been found that many of the average energy conditions allowing for theorem-strengthening are violated by classical matter models, and, less straightforwardly, in the context of quantum field theory in curved spacetime.[5] Thus, many of the classic theorems of relativity still fail to cover various classical and semi-classical models of interest. Efforts to better this situation are ongoing.

In the quantum field theoretic context, there is growing evidence that the extent of the violation of certain energy conditions is, in some sense, restricted. There is a whole body of work dedicated to making this precise. The idea is to produce certain kinds of inequalities that represent the spatiotemporal constraints that (contractions of) renormalized stress energy tensors in various contexts of quantum field theory in curved spacetime obey. These inequalities are known as *quantum energy inequalities*, and, in the best of cases, they have been used either to prove the av-

[4]See [4] for a foundationally oriented introduction to such average energy conditions.

[5]See [4], [5], [15] for various expressions of this fact.

erage in certain circumstances or to constrain the properties of certain spacetime scenarios otherwise associated with violations of energy conditions. For example, there is Ford and Roman's study of the properties of traversable wormholes [9], Fewster, Olum, Pfenning's quantum energy inequalities proof of the average null energy condition under certain circumstances [7], and Kontou and Olum's quantum energy inequalities proof under different circumstances [12].

Though quantum energy inequalities have undoubtedly led to important insights, they have not yet permitted the extension of certain classic relativity theorems to the semi-classical context. Fewster and Galloway [6] recently proposed a study aimed in this direction. They formulated versions of Hawking's cosmological and Penrose's collapse singularity theorems based on energy conditions allowing for, respectively, violations of the strong and null energy conditions. Though inspired by quantum energy inequalities methods, the conditions that stand in as energy conditions in their singularity theorems are not, strictly speaking, quantum energy inequalities. On this point, it is worth noting that the relationship between quantum energy inequalities and energy conditions remains not particularly well understood, with only a few definitive results available. One example is [9] where Fewster and Roman show that the existence of quantum energy inequalities is not necessary for the average null energy condition to be satisfied.

Galloway and Fewster's arguments rely on new results establishing sufficient conditions for focusing. Their strengthening of Penrose and Hawking's theorems proceed without changing the spirit of the original arguments. They also consider a specific matter model to illustrate how the null energy condition may be violated whilst satisfying their weakened energy conditions. The specific lemma that we shall use from their work is as follows.

Lemma. *Consider the initial value problem for* $z(t)$

$$\dot{z} = \frac{z^2}{s} + r$$

where $r(t)$ is continuous on $[0,\infty)$, $z(0) = z_0$ and $s > 0$. If there exists $c \geq 0$ such that

$$z_o - \frac{c}{2} + \lim_{T\to\infty} \inf \int_0^T e^{-2ct/s} r(t)dt > 0$$

then the initial value problem has no solution on $[0,\infty)$, where 'no solution' means $z(t) \to \infty$ as $t \to t_*^- < \infty$.

Here, $\lim_{T\to\infty} \inf \int^T$ is shorthand for $\lim_{T\to\infty} \inf_{T'\geq T} \int^{T'}$.

2. A Stronger Theorem

Hawking's area theorem [13] is often considered to be one of the most important results in black hole theory. It describes a fundamental property of dynamical black holes, it underlies black hole thermodynamics and it bounds the amount of radiation that can be emitted upon black hole collisions. Consider Ashketar and Krishnan's remarks in their recent well cited review on generalized black holes [1].

Ashtekar and Krishnan. "For fully dynamical black holes, apart from the 'topological censorship' results which restrict the horizon topology [...], there has essentially been only one major result in exact general relativity. This is the celebrated area theorem proved by Hawking in the early seventies [...]: If matter satisfies the null energy condition, the area of the black hole event horizon can never decrease. This theorem has been extremely influential because of its similarity with the second law of thermodynamics."

The precise statement of the area theorem depends on a number of definitions for which there are a number of possible choices. The definitions and the proof used in theorem below are identical to those used by Wald [33]. This choice is made for reasons of expediency. Note, however, that the following arguments apply to other available formulations of the area theorem.

Inspired by the above lemma, let us now define the following new average energy condition.

Definition. *A spacetime (M, g) satisfies the damped averaged null energy condition, if along each future complete affinely parametrized null geodesic $\gamma : [0, \infty) \to M$, there exists a non-negative constant $c \geq 0$ such that*

$$\lim_{T\to\infty} \inf \int_0^T e^{-ct} \operatorname{Ric}(\gamma', \gamma')dt - \frac{c}{2} > 0$$

where γ' is a tangent vector for γ.

As it turns out, the damped average null energy condition is sufficient for the following stronger version of Hawking's classic result.

Theorem. *Let (M, g) be a strongly asymptotically predictable spacetime that satisfies the damped average null energy condition. Let Σ_1 and Σ_2 be spacelike Cauchy surfaces for the globally hyperbolic region $\tilde{V}$ with $\Sigma_2 \subset I^+(\Sigma_1)$ and let $\mathcal{H}_{1(2)} = H \cap \Sigma_{1(2)}$, where H denotes the event horizon, i.e., the boundary of the black hole region of (M, g). Then the area of $\mathcal{H}_2$ cannot be less than that of $\mathcal{H}_1$.*

Remark 1. The definition of 'strongly asymptotically predictable' spacetime follows Wald [33]. Namely, for a spacetime (M, g) that conformally embeds into an unphysical spacetime $(\tilde{M}, \tilde{g})$, (M, g) is said to be strongly asymptotically predictable if in $(\tilde{M}, \tilde{g})$ there is an open region $\tilde{V} \subset \tilde{M}$ with $\overline{M \cap J^-(\mathcal{J}^+)} \subset \tilde{V}$ such that $(\tilde{V}, \tilde{g})$ is globally hyperbolic. See chapter 11 of [33] for the relevant notion of conformal embedding and the so-called unphysical spacetime $(\tilde{M}, \tilde{g})$.

Proof. The argument is a straightforward application of the above lemma with the original arguments by Hawking, which are as in [33] in this formulation. We start by showing that the expansion θ of the null generators of H is everywhere non-negative.

Suppose $\theta < 0$ at some $p \in H$. Let Σ be a spacelike Cauchy surface for $\tilde{V}$

passing through p and consider the two-surface $\mathcal{H} = H \cap \Sigma$. Since $\theta < 0$ at p, we can deform $\mathcal{H}$ outward in a neighborhood of p to obtain a surface $\mathcal{H}'$ on Σ which enters $J^-(\mathcal{J}^+$ and has $\theta < 0$ everywhere in $J^-(\mathcal{J}^+)$. However, by the same argument as in proposition 12.2.2 of [33], this leads to a contradiction as follows. Let $K \subset \Sigma$ be the closed region lying between $\mathcal{H}$ and $\mathcal{H}'$, and let $q \in \mathcal{J}^+$ and with $q \in \partial J^+(K)$. Then the null geodesic generator of $J^+(K)$ on which q lies must meet $\mathcal{H}'$ orthogonally.

However, this is impossible since, by $\theta < 0$ on $\mathcal{H}'$, the null generator of $J^+(K)$, eg., γ, will have a conjugate point before reaching q. This follows from consideration of Raychaudhuri's equation for irrotational null congruences, and identifying with the equation in the lemma. That is, take $s = 2$, $z = \theta(t)$, $r(t) = \mathrm{Ric}(\gamma', \gamma') + 2\sigma^2$ where $\theta(t)$ is the null expansion relevant to the null geodesic γ. We now have an initial value problem as in the lemma with $z(0) = -\theta > 0$. By the damped average null energy condition and the non-negativity of σ^2, it follows that the conditions of the lemma are satisfied, and so there occurs a focal point within finite affine parameter from $\gamma(0)$.

It now follows that $\theta \geq 0$. By standard properties of event horizons (described in Chapter 12 of [33]), each $p \in \mathcal{H}_1$ lies on a future inextendible null geodesic, η, contained in H. Since Σ_2 is a Cauchy surface, η must intersect Σ_2 at a point $q \in \mathcal{H}_2$. Thus we obtain a natural map from $\mathcal{H}_1$ into a portion of $\mathcal{H}_2$. Since $\theta \geq 0$, the area of the portion of $\mathcal{H}_2$ given by the image of $\mathcal{H}_1$ under this map must be at least as large as the area of $\mathcal{H}_1$. In addition, since the map need not be onto, the area of $\mathcal{H}_2$ may be even larger. Thus, the area of $\mathcal{H}_2$ cannot be smaller than that of $\mathcal{H}_1$. $\square$

The main reason for choosing the damped average null energy condition stems from the fact that Fewster and Galloway's work is both state of the art and neatly amenable to classical field models that violate the null energy condition. In section 6 of their paper [6], they construct such a model which, by inessential modifications, is straightforwardly applicable to the above theorem.

It is worth noting that a recent study [3] has shown there to be unstated or

undesirable assumptions with regards to the differentiability of the generators of the horizon, in both Hawking [13] and Wald's [33] formulation of the area theorem. In that same study, a much more sophisticated area theorem has been proved that overcomes these deficiencies. We observe that the strengthening offered here also generalizes immediately to their superior area theorem.

To see the meaning of the damped average null energy condition, observe that the exponential damping in the integrand shows that boundary effects are crucial in determining whether area non-decrease obtains. Provided there occurs positive contributions to the null contractions of the Ricci tensor near the horizon, the condition may be satisfied even if negative contributions persist for arbitrarily long segments of the null geodesic away from the horizon. That is: the area may be nondecreasing even if there are large violations to the standard average null energy condition for semi-complete geodesics.

Note of course that strictly speaking of course, this interpretation won't be valid if the relevant contractions of the Ricci tensor are such that the contributions grow sufficiently fast in regions away from the black hole so as to overcome the damping factor, eg., if the profile of the Ricci tensor is of the form $e^{e^{ct}}$ then the integrand is of the form $e^{-ct+e^{ct}}$ and of course in that case contributions far away matter enormously so the overall energy condition. That being said, such energy profiles are so incredibly unlike the typical setting of asymptotically flat black hole solutions that we can safely disregard them in the current interpretation.

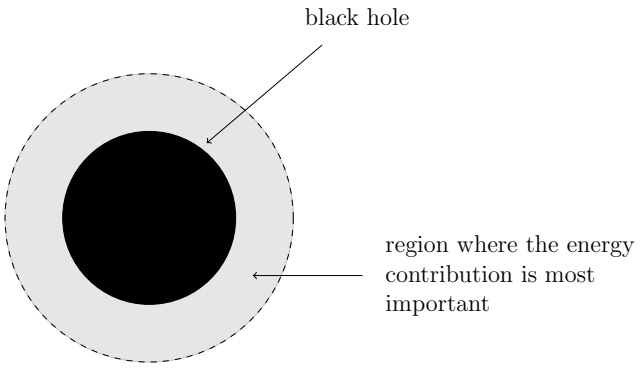


Figure II.1. In virtue of the exponential damping of the integrand's contribution, it is only that which lies within some vicinity of the event horizon that contributes to the black hole's non-decrease behaviour.

It is perhaps also worth noting that the current arguments can be applied to a number of other standard theorems about black holes. This includes results describing the location of trapped surfaces, marginally trapped surfaces, apparent horizons, and theorems about topological censorship; see [33] for a description of these results. I have focused on the area theorem here due to the obvious tension it generates with the idea of Hawking radiation.

As a final point, we also note that the current picture dovetails nicely with the generalized black holes framework. In that framework, area non-decrease results are obtained for trapping, isolated and dynamical horizons upon assuming that the null energy condition holds within a neighborhood of the relevant horizon. Notice, then, that the damped average null energy condition will be satisfied in such cases.

3. Discussion

There is now a certain conventional wisdom regarding the tension between Hawking radiation and the area theorem. Consider, for instance, the remarks made by Visser [15] and, respectively, Bousso and Engelhart [2].

Visser. "The very fact that Hawking evaporation occurs at all violates the area increase theorem...for classical black holes. This implies that the quantum process underlying the Hawking evaporation process must also induce a violation of one or more of the input assumptions used in proving the classical area increase theorem. The only input assumption that seems vulnerable to quantum violation is the assumed applicability of the null energy condition."

Bousso and Engelhart "Hawking's theorem holds in spacetimes obeying the null curvature condition, $Ric(k,k) \geq 0$ for any null vector k. This will be the case if the Einstein equations are obeyed with a stress tensor satisfying the NEC, $T(k,k) \geq 0$. The NEC is satisfied by ordinary classical matter, but it is violated by valid quantum states (e.g., in the Standard Model). In particular, the null energy condition fails in a neighborhood of a black hole horizon when Hawking radiation is emitted. Indeed, the area of the event horizon of an evaporating black hole decreases, violating the Hawking area law."

The first point to make in view of the above comments is that a precise and rigorous understanding of the area theorem's semi-classical failure remains a point of discussion. Moreover, since it is now widely accepted that black holes radiate, there are at least two attitudes one can take: either the standard classical geometrical arguments aren't applicable and we must await further progress, or, those arguments apply and it is only the energy conditions component of these arguments that needs generalization to the quantum field theory in curved spacetime context.

In the former case, the theorem presented here provides no particularly new information.

In the latter, there are at least two further possibilities. One is that there may be conflicting results, whereby a model satisfies an energy condition that mandate area non-decrease result (similar perhaps to the one above), despite being one for

which evaporation is expected. In that case, there would seem to be an interesting contradiction at work.

The other possibility is that everything remains consistent. That is, models which satisfy energy conditions that permit for area non-decrease end up being precisely those in which radition occurs, but is outweighed by the energy absorbed in the form of infalling matter, so that, in the end, evaporation doesn't actually occur.

Results in either direction are in demand.

My own view is that neither case is backed by precise and rigorous arguments, and as such that we're still some way from understanding the semi-classical failure of the area theorem. Progress is likely to come from a better understanding of the energy conditions that are suitable in quantum field theory in curved spacetime, and the effects of back-reaction.

References

[1] Ashtekar, A., Kirshanan, B., Living.Rev.Rel. 7:10, (2004)

[2] Bousso, R., Engelhart, N., Phys. Rev. Lett. 115, 081301, (2015)

[3] Chruściel, P., Delay, E., Galloway, G., Howard, R., Regularity of Horizons and the Area Theorem, Annales Henri Poincaré, 109-178, (2001)

[4] Curiel, E., A primer on energy conditions, arXiv:1405.0403 [physics.hist-ph]

[5] Fewster, C.J., Energy inequalities in quantum field theory XIVth International Congress on Mathematical Physics ed Zambrini J C (Singapore World Scientific), arXiv:math-ph/0501073

[6] Fewster, C.J., Galloway, G., Class. Quant. Grav. 28 125009, (2011)

[7] Fewster, C.J., Olum, K. D., Pfenning, M. J.: Phys. Rev. D 75, 025007 (2007)

[8] Fewster, C., Roman, T., Phys. Rev. D 67, 044003, (2003)

[9] Ford, L.H., Roman, T.A.: Phys. Rev. D 53, 5496–5507, (1996)

[10] Hawking, S.W., Black holes in general relativity. Comm. Math. Phys. 25(2), 152–166, (1972)

[11] Jaramillo, J., An introduction to local Black Hole horizons in 3+1 approach to general relativity. arXiv:1108.2408 [gr-qc]

[12] Kontou, E. A., Olum, K.D.: Phys. Rev. D 92, 124009, (2015)

[13] Roman, T.A., Phys. Rev. D 37, 546–548, (1988)

[14] Tipler, J.F., Phys. Rev. D 17, 2521, (1978)

[15] Visser, M., Lorentzian wormholes: from Einstein to Hawking AIP Series in Computational and Applied Mathematical Physics, American Institute of Physics (1996)

[16] Wald, R., General Relativity, University of Chicago Press, (1984)

II.2. Black Hole Singularities with Achronologies

1. Introduction

Singularity theorems are a cornerstone of general relativity and that of Hawking and Penrose [8] is arguably the most important example; see chapter 9 of [20] for a discussion.

Theorem (Hawking-Penrose [8])**.** *Let (M, g) be a spacetime satisfying the timelike and null generic conditions and the strong energy condition. Suppose that (M, g) either admits, a compact achronal set without edge, a trapped surface, or a point $p \in M$ such that the expansion of the future (or past) directed null geodesics emanating from p becomes negative along each geodesic in this congruence. Then, if (M, g) is chronological, there is an incomplete non-spacelike geodesic in M.*

Note that while this theorem relies on the assumption of chronology, there are a number of explicit solutions to the Einstein field equations that contain closed timelike curves, and moreover, that some of these are still singular (i.e., non-spacelike geodesic incomplete). The most important example is the Kerr-Newman class of vacuum black hole solutions, the interiors of which contains closed timelike curves and incomplete null geodesics, and the exteriors of which is widely believed to be the asymptotic state of astrophysical black holes; see chapter 5 of [14] for more details on these solutions.

In view of this, soon after Hawking and Penrose's theorem, Tipler [20] studied whether one could get rid of the singularities of the above theorem by removing the assumption of chronology. And although he was able to prove a singularity theorem under various assumptions that allow for chronology violation, his theorem is rather restricted in its scope of applicability, and like Hawking and Penrose's, it doesn't provide any information on the location or the type of the relevant incomplete geodesics.

Some years later, Newman made a remarkable contribution [15]. He constructed an explicit spacetime solution which violates chronology, contains a trapped surface, is non-spacelike geodesically complete, and yet satisfies all the local conditions of

Hawking and Penrose's theorem. He therefore showed that singularities can indeed disappear if one totally removes chronology. Crucially, however, his example violates chronology in the strongest possible way, and as such it was of interest to explore whether there were intermediate versions of chronology violation that are sufficient for the Hawking Penrose theorem.[6]

Inspired by Newman's example, Kriele [9] sought to formulate a singularity theorem that relied on weaker forms of chronology violation. He was eventually able to generalize Hawking and Penrose's theorem so as to permit for certain types of chronology violation. Unfortunately, as is easily seen, his theorem involves a number of non-standard definitions and thus somewhat unwieldy.

Like Hawking and Penrose's theorem, Kriele's theorem also gave no information as to the location or type of the incomplete geodesics.

In view of this line of results, I shall now formulate a new singularity theorem designed specifically for black holes that achieves things that the aforementioned ones do not. I shall do so by positing various geometric conditions that permit identifying a region corresponding to a black hole interior region. I shall not make use of the notion of a conformal boundary for the spacetime. The conclusion obtained is that there are incomplete *null* geodesics in the black hole's interior, even if chronology is partially violated therein. There are two key new ingredients that I shall use and both are due to Minguzzi.

First, I shall express the permitted chronology violation using Minguzzi's [12] recently offered definition of the boundary of a chronology violating set. This has the advantage of leading to a statement which relies on clear geometric conditions. Indeed, it is easy to see that the one employed here is cleaner and more user-friendly than Kriele's. The second ingredient is a non-trivial lemma in Lorentzian geometry that Minguzzi proves in [11].

Another interesting difference with the results of Kriele and Hawking and Pen-

[6]His example is totally vicious; that is, there passes a closed timelike curve through every point.

rose, is that I will use neither the strong energy condition nor the generic conditions.

Note below that $p \leq q$ means $q \in J^+(p)$, and that by g', we mean the spacetime metric g of M restricted to the open set $O \subset M$, i.e., $g\,|_O$.

Theorem. *Let (M, g) be a spacetime satisfying the null energy condition. Let $V \subset M$ be an edgeless, spacelike and achronal hypersurface of co-dimension one. Let $S \subset V$ be a compact spacelike, achronal submanifold embedded in V such that S has non-empty boundary ∂S such that $\partial S = edge(S) = \Sigma$. Let N be a null hypersurface defined by $N \equiv \partial I^+(S) \backslash S$ such that the following properties are obeyed:*

(i) $N \cap S = \Sigma$ and for any point $p \in N$, any point $q \in \partial I^-(p) \cap N$ is the future endpoint of a future directed achronal null geodesic with past endpoint on Σ;

(ii) there is no causality violation on N and the relationship "$\leq$" is closed on N;

(iii) Σ is a C^0 surface such that for each point $\sigma \in \Sigma$ there is only one null generator η_σ of N emanating from σ,

(iv) the achronal null geodesic generators of N, $\{\eta\}$ are future inextendible.

Denote by O the open set defined by $O = I^+(S)$ and suppose that the following properties obtain:

(a) all future inextendible null geodesics in O meet a future trapped surface in O,

(b) either (O, g') satisfies chronology, or there is a non-empty chronology violating class $[r] \subset O$ with future and past boundary, denoted by $B_f([r])$ and $B_p([r])$, such that at least one of the following holds:

(b.i.) $B_f([r]) \cap O \neq \emptyset$,

(b.ii.) (O, g') is non-totally vicious and either $B_f([r])$ or $B_p([r])$ is compact.

Then there are future incomplete null geodesics in O.

The picture in mind for the theorem is as follows.

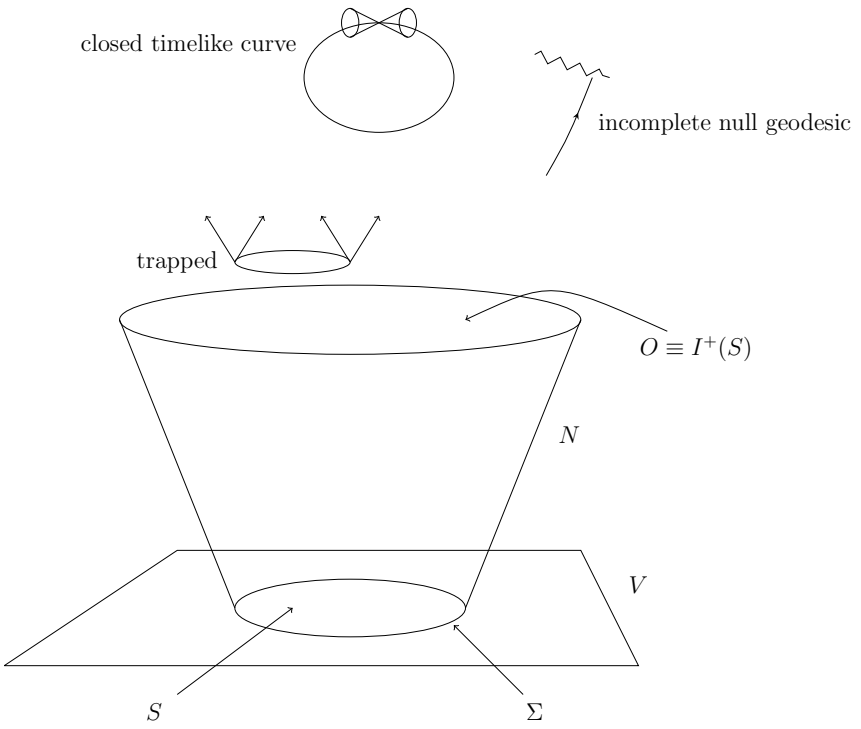


Figure II.2. V slices through the event horizon, represented by N, and S is the portion of V lying inside the black hole, and Σ is its boundary. The interior of the black hole is O, there are lots of trapped surfaces in O. Closed timelike curves may be present, but the interior will still contain null geodesic incompleteness under the theorem's assumptions.

The conditions of the theorem are set up so as to apply to black holes which are close to being stationary, where N models the event horizon and V is a slice through the spacetime which intersects N. The specific conditions on N labelled by (i),(ii), (iii), and (iv) are supposed to capture the physical idea that eventually, after things settle down, the geometry of spacetime near the horizon should not be so different from the stationary case. The proof of the theorem actually reveals that

some of these assumptions are unnecessary, and that all that one needs is that N be non-compact. Still, I've included them for the purpose of physical motivation.

With this picture in mind, S is then of codimension one, closed, spacelike submanifold corresponding to the part of V lying inside the black hole. Σ is the boundary of S, which thus lies on N.

Assumption (a) captures the idea that black hole interiors are filled with trapped surfaces. Note that this assumption can be weakened in a way that preserves the contradiction of the Penrose singularity theorem. In particular, rather than requiring that all null geodesics encounter trapped surface, one could instead require that such null geodesics in O encounter either a trapped surface or a weakly trapped surface in such a way that the generic condition holds for null geodesics emanating from this surface in the direction of zero expansion.

Assumption (b) permits various forms of chronology violation in the black hole interior. These have been chosen in view of the chronology violating and singular interior of Kerr, and they rely on the recently offered definitions and theorems of Minguzzi [11], [12] described in more detail below.

2. Chronology Violation

The convention adopted here follows [13] so that by a spacetime (M, g) we mean a C^r, $r \geq \{3\}$ connected, time-oriented Lorentzian manifold of dimension $d \geq 4$. The elements in M are as differentiable as permitted by M so C^{r-1} for g and C^{r-3} for the curvature tensor R.

By null (timelike) convergence condition we mean $R(k, k) \geq 0$ where k is any null (timelike) vector and $R(\cdot, \cdot)$ is the Ricci tensor.

By a *future trapped surface* we mean a closed smooth compact spacelike submanifold of M of co-dimension 2, with null expansion scalars $\theta_{+,-}$ associated to the future directed null normal vector fields $l_{+,-}$ satisfying $\theta_{+,-} < 0$. A *marginally trapped surface* (MTS) is the same geometrical object except that $\theta_{+,-} \leq 0$ and

one of these is zero, a *marginally outer trapped surface* (MOTS) has $\theta_+ = 0$ and no restriction on θ_-. We say that a null geodesic $\gamma : I \to M$ encounters a surface S if there is a point $p \in S$ such that $\gamma(a) = p$ for some $a \in I$.

By $x \ll y$ ($x \leq y$) we mean that there exists a future directed timelike (causal) curve from x to y. By a null line we mean an inextendible achronal causal curve. By a future (past) null ray we mean a half null line which is inextendible in the future (past) direction.

The concept of an edge is defined as follows. 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 $edge(S)$ the set of edge points of S.

By a totally vicious spacetime (M, g) we mean one where $I^{\pm}(p) = M$ for all $p \in M$.

We now recall the following standard result, cf., [1].

Proposition. *Let S be closed. Then each $p \in \partial I^+(S) \backslash S$ lies on a null geodesic contained in $\partial I^+(S)$, which either has a past endpoint on S, or else is past inextendible in M.*

We shall now recall some recent results of Minguzzi, though our account merely includes what is necessary. The interested reader is encouraged to consult the comprehensive description proposed in the original papers [11],[12].

Theorem (Minguzzi [11])**.** *Let (M, g) be any spacetime.*

(a) A chronological spacetime (M, g) without future null rays is globally hyperbolic.

(b) A non-totally vicious spacetime (M, g) without null rays is globally hyperbolic.

(c) A chronological spacetime (M, g) without null lines is stably causal.

(d) A non-chronological spacetime (M, g) without null lines is either totally vicious or has a non-empty chronology violating open set, possibly made up

of distinct sets $C \equiv \cup_\alpha C_\alpha$. The components of the closed boundaries $B_{\alpha k}$ of these sets $\partial C_\alpha \equiv \cup_k B_{\alpha k}$ are all disjoint and non-compact. Finally, non-totally vicious spacetimes without null lines are non-compact.

Another ingredient we shall use, also due to Minguzzi, is the recent proposal [12] for the definition of the boundary of a chronology violating region.[7] The chronology violating region is defined as the set of of points through which there passes a closed timelike curve. We write $x << x$ if there is a closed timelike curve starting at and returning to x. The relation $x \sim y$ if $x \ll y$ and $y \ll x$ is an equivalence relation in C, and, as is known from Carter [4], it splits the chronology violating region into (open) equivalence classes. We denote such a class in square brackets $[x] = I^+(x) \cap I^-(x)$. An important fact about chronology violating class is encoded in the following lemma.

Lemma (Minguzzi [12])**.** *Let $[r]$ be a chronology violating class. If $p \in \partial[r]$ then through p passes a future or past null ray contained in $\partial[r]$ and possibly both.*

Using this definition, a short argument leads to the following.

Definition. *Let $[r]$ be a chronology violating class. The set $R_f([r])$ is that subset of $[\partial r]$ which consists of the points q through passes a future null ray contained in $[\partial r]$. The set $R_p([r])$ is defined analogously.*

Lemma (Minguzzi [12])**.** *The sets $R_p([r])$ and $R_f([r])$ are closed and $\partial[r] = R_p([r]) \cup R_f([r])$.*

After showing that a number of relations hold between $R_f(p)([r])$, $I^{+(-)}([r])$, $\partial I^{+(-)}([r])$ and $\overline{[r]}$, Minguzzi considers the sets $B_{f(p)}([r]) \equiv \overline{[r]}/I^{-(+)}([r])$ and proves that they have the following properties.

[7] The boundary of a chronology violating set C had been considered in the literature [13], [19] but these studies were based on the non-trivial assumption that $M \backslash \overline{C}$ is globally hyperbolic (i.e. so that ∂C can be identified as a component of a Cauchy horizon). Part of the appeal of Minguzzi's analysis is that it is based only on properties that are intrinsic to chronology violation and that it makes no use of such assumptions.

Proposition (Minguzzi [12]). *The set $B_{p(f)}([r])$ is closed, achronal, and generated by past (future) null rays.*

Proposition (Minguzzi [12]). *The following identity holds $\partial[r] = B_p([r]) \cup B_f([r])$.*

Minguzzi finally settles with the following definition for the boundary of the chronology violating class.

Definition. *The sets $B_f([r])$ and $B_p([r])$ are respectively the future and the past boundaries of the chronology violating class $[r]$.*

3. Proof

Let (O, g') be a spacetime with $g' = g\,|_{M\cap O}$. We shall treat (O, g') as a spacetime in its own right and investigate the consequences in (M, g). Firstly, we observe that if (M, g) is future null complete then so is (O, g').

Step 1. If (O, g') is chronological and future null complete, then N must be compact.

By a standard focusing lemma, eg., proposition 9.3.9 in [20], (b) implies that if (O, g') is future null geodesically complete, then there can be no future null rays in O. A future null ray is achronal but by proposition 9.3.9, any directed null geodesic in O must meet a future trapped surface and focus. One might think that this is not enough to get rid of future null rays because we could consider the null geodesic whose endpoint is in the future of the point at which the null geodesic focused. Perhaps this new geodesic is a null ray. But this is straightforwardly seen to be impossible, for (b) applies to all null geodesics, and thus also applied to this geodesic which we have considered as a candidate ray.

It then follows, by a theorem in section 2, that either (O, g') admits a Cauchy surface or else violates chronology. By the Penrose's collapse singularity theorem

[17], given that (O, g') satisfies the null energy condition, contains a closed future trapped surface and is future null geodesically complete, its Cauchy surface must be compact.

Recall here that $I^+(S) = O$. By [2] we can take the compact Cauchy surface for (O, g') to be a spacelike slice $V \subset O$. Consider any point $q \in \partial O$. By definition of N and of S, any future timelike curve that intersects N and S must enter (O, g'). A standard result of globally hyperbolic spacetimes and Cauchy surfaces then dictates that all such timelike curves, if made future inextendible, will intersect V exactly once. This permits us to define a continuous map $\psi : \Sigma \cup \partial I^+(S) \to V$ by following integral curves of a smooth future directed timelike vector field in M. Note that $\Sigma \cup \partial I^+(S)$ and Σ are both closed. Continuity of ψ and compactness of Σ imply, together, that $\Sigma \cup \partial I^+(S)$ is compact. The set $\Sigma \cup \partial I^+(S)$ is also closed and by definition N is a closed subset of $\Sigma \cup \partial I^+(S)$ and is thus also compact.

Step 2. Assumptions in the theorem imply that N is homeomorphic to $\Sigma \times [0, \infty)$.

Here, we shall use a variation on an argument that was exploited in [3] with somewhat different assumptions.

We first make the following observation. By assumption, N is generated by achronal null geodesics $\{\eta\}$ which are future inextendible and which have past endpoints on Σ. Moreover, from our assumption on the geometry of N in the vicinity of S, it follows that for each point $\sigma \in \Sigma$, there is a unique future inextendible null geodesic generator η_σ emanating from σ.

Consider the set A defined by $A = \{a \in N : a \in \eta_\sigma, \sigma \in \Sigma\}$.

We shall show that A is both open and closed in N. It will then follow that $A = N$.

That A is open in N. For each $\sigma \in \Sigma$, let $Q(\sigma)$ denote the initial null tangent vector to η_σ. Our assumption that for each point $\sigma \in \Sigma$ there is only one null generator η_σ of N emanating from σ guarantees that $Q(\sigma)$ varies continuously with

σ. That is, $\sigma \to Q(\sigma)$ is a continuous null vector field along Σ.

Now consider the map defined by flowing along the generators of N. That is, we consider the map $\chi : U \subset \Sigma \times [0, \infty) \to N$ defined by $\chi(s, \lambda) = exp_\sigma(\lambda Q(\sigma)) = \eta_\sigma(s)$. The set U represents the largest set in $\Sigma \times [0, \infty)$ on which χ is defined, which may be strictly smaller. The continuity of $\sigma \to Q(\sigma)$ and the behavior of geodesics guarantees that χ is continuous. Since the generators of N cannot intersect, χ is one to one. Hence it follows from invariance of domain [14] that $A = \chi(U)$ is open.

That A is closed in N. This follows by the assumption noted in the previous paragraph and the compactness of S. In particular, let $\{p_n\}$ be a sequence of points in A such that $p_n \to p \in N$. For each n, p_n is on η_{q_n} for some $q_n \in \Sigma$. Since Σ is compact, there exists a convergent subsequence $q_n \to q \in \Sigma$. Since $p_n \geq q_m$, and the causal relationship "$\geq$" is closed by assumption, it follows that $p \geq q$. Since A is achronal, we must then have $p \in \eta_q$ and hence $q \in A$. Hence A is closed in N.

We now have $A = N$ and $\chi : U \subset \Sigma \times [0, \infty) \to N$ is a homeomorphism onto N. We fix a complete Riemannian metric h in M. For each $s \in \Sigma$, reparametrize η_σ with respect to arclength in the metric h. Since η_σ is future inextendible, the reparametrized curve $\tilde{\eta_s}$ will be defined on $[0, \infty)$. It follows that the map $\chi : \tilde{\Sigma} \times [0, \infty) \to H$ defined by $\tilde{\chi}(\sigma, \lambda) = \tilde{\eta_\sigma}(\lambda)$ is a homeomorphism since $\chi : U \to N$ is. Hence, N is homeomorphic to $\Sigma \times [0, \infty)$, with the homeomorphism given by the null generators (suitably parametrized).

Step 3. There is chronology violation in (O, g') as permitted in the main theorem.

The statements obtained in steps 1 and 2 can only be consistent if either (O, g') is future null incomplete or there is chronology violation in (O, g'). Assuming that (O, g') is future null complete, it follows that there is a non-empty chronology violating region $Q \subseteq O$, to which can be associated a chronology violating class, which we denote by $[r]$.

Using definitions and propositions of section 2, future null completeness along with assumption (b) imply that O contains no null rays, and, therefore, that $B_f([r]) \cap O = \emptyset$. This leads to condition (b)-i.

For (b)-ii, recall that lack of future null rays implies the lack of null lines, and thus, by a theorem in section 2, either (O, g') is totally vicious, or the boundaries of the chronology violating class is non-compact. Assuming non-total viciousness of (O, g'), compactness of either one of the boundaries provides the relevant contradiction. This completes the proof of the main theorem.

4. Discussion

The strongest assumption of the theorem is (a), i.e., that all future directed null geodesics future inextendible in O meet future trapped surfaces in O. This captures the idea that a black hole interior is a region of strong gravity and that the trapped surfaces are inside the event horizon. There are a number of classic theorems justifying the latter in various contexts of asymptotic predictability; for instance see chapter 9 of [14]. As for the former, it is difficult to provide a mathematically rigorous justification, other than alluding to existing black hole solutions.

On this point, there is also the classic result of Schoen and Yau [17] showing that initial data sets with sufficiently dense concentrations of matter contain *marginally outer trapped surfaces*, and in fact trapped surfaces. A marginally outer trapped surface is essentially the same geometric object as a trapped surface except that the null expansion is merely required to be zero in one null normal direction, rather than both negative. Further such existence results for trapped surfaces would be welcome of course. On this issue, we note the results in [5] showing that any initial data containing a marginally outer trapped surface which lies in an region satisfying a no - Killing Initial Data condition is arbitrarily close to another initial data in which there is a strictly outer trapped surface. This no - Killing Initial Data condition concerns the initial data and it is equivalent to the absence of Killing vectors

in the local spacetime development associated with that initial data. The condition has been shown to be generic under various conditions in the space of initial data. Generically, therefore, one expects there to be lots of trapped surfaces whenever marginally outer trapped surfaces (or indeed lots of concentration of matter) are present.

Ideally, we would like to possess a number of illustrative examples for the theorem. The Kerr-Newman class comes to mind since these solutions describe black hole interiors which contain both closed timelike curves and singularities. The issue with the Kerr-Newman class is that those solutions are usually considered in their maximal analytic form. In this form, it is possible to access another universe by going into the black hole interior. This feature makes the theorem inapplicable to these full extensions, because $O \equiv I^+(S)$ will now extend into different universes, which are regions going beyond the regions of applicability of the theorem. If we restrict the size of these extensions so as not to spill over into further universes, then the theorem would apply. However in that case, since the resulting spacetimes would be extendible, they would be geodesically incomplete anyway.[8]

Ultimately, what we would like the theorem to tell us is whether singularities occur in the course of realistic gravitational collapse. The applicability of Penrose's original theorem or indeed that by Hawking Penrose theorem isn't straightforward since it is a possibility which we cannot strictly speaking exclude that closed timelike curves could form. Though this may not occur, if it does then we'd like a theorem telling us whether or not to expect singularities. The theorem proved here achieves this goal, albeit under the relatively strong assumption that trapped surfaces are omnipresent in the black hole interior. As such, the physical applicability of the theorem remains, admittedly, a point of discussion. Whatever the case, it does show is that it is possible to remove, at least partially and in certain contexts, the assumption of chronology in Hawking and Penrose's theorem. As for the physics

[8]Indeed, it is basic result that 'extendibility $\Rightarrow$ geodesic incompleteness'; see for instance chapter 6 of [1].

of closed timelike curves and whether they do in fact form, the issue remains wide open.

References

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

[2] Bernal, A. N., Sánchez, M., Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes, Commun.Math.Phys. 257 (2005) 43-50, arXiv:gr-qc/0401112

[3] Browdy, S.F., Galloway, G.J., Topological censorship and the topology of black holes, Journal of Mathematical Physics 36, 4952 (1995); https://doi.org/10.1063/1.530930

[4] Carter, B., Causal structure in space-time, Gen. Relativ. Gravit., 1:349-391,1971

[5] Chruściel, P.T., Galloway, G., Outer trapped surfaces are dense near MOTSs, Class. Quantum Grav., 31 (2014) 045013

[6] Hawking, S. W., Chronology protection conjecture. Phys. Rev. D, 46:603611, 1992.

[7] Hawking, S. W ., and Ellis, G. F. R. (1973). The Large Scale Structure of Space-Time (Cambridge University Press, Cambridge).

[8] Hawking, S. W ., Penrose, R. (1970) . Proc. Roy. Soc. London A 3 14, 529.

[9] Kriele, M., The structure of chronology violating sets with compact closure. Class. Quantum Grav., 6:1607-1611, 1989

[10] Kriele, M., Foundations of General Relativity and Differential Geometry, 1999, Springer-Verlag Berlin Heidelberg.

[11] E. Minguzzi. Chronological spacetimes without lightlike lines are stably causal. Commun. Math. Phys., 288:801-819, 2009. arXiv:0806.0153

[12] Minguzzi, E., The boundary of the chronology violating set, Class. Quantum Grav. 33 (2016) 225004, arXiv:1603.08190 [gr-qc]

[13] Minguzzi, E., Sánchez, M., The causal hierarchy of spacetimes. In H. Baum and D. Alekseevsky (eds.), vol. Recent developments in pseudo-Riemannian geometry, ESI Lect. Math. Phys., (Eur. Math. Soc. Publ. House, Zurich, 2008), p. 299-358, ISBN=978-3-03719-051-7, arXiv:gr-qc/0609119

[14] Munkres, J., Elements of Algebraic topology. Addison-Wesley, Menlo Park, 1984.

[15] Newman, R. P. A. C., Black holes without singularities, Gen Relat Gravit (1989) 21: 981. https://doi.org/10.1007/BF00774083

[16] Penrose, R. (1965). Gravitational Collapse and Space-Time Singularities. Phys. Rev. Lett.14, 57.

[17] Schoen, R., Yau, S.T., The existence of a black hole due to condensation of matter, Commun.Math. Phys. (1983) 90: 575. https://doi.org/10.1007/BF01216187

[18] Senovilla, J., Singularity Theorems and Their Consequences, General Relativity and Gravitation, Vol.30, No. 5, pp.701–848, 1998, arXiv:1801.04912 [gr-qc]

[19] Thorne, K. S., Closed Timelike Curves, volume General Relativity and Gravitation, pages 295-315. Institute of Physics Publishing, Bristol, England, 1993.

[20] Tipler, F. J., Singularities and causality violation. Ann. Phys., 108:136, 1977.

[21] Wald, R., General Relativity. 1984. Chicago University Press.

II.3. The Causal Structure of Evaporating Black Holes

1. Introduction

Fundamental theoretical physics is split between two realms, general relativity and quantum field theory, both of which have been been spectacularly successful within their domains of applicability. After decades of effort however, these frameworks have yet to be unified in a way that doesn't divide opinion.

In this search for unification, black holes have played a key role. A watershed moment in the subject was Hawking's calculation that quantum mechanical effects force black holes to radiate away energy, and that this could cause them to shrink, and, eventually, disappear. Hawking's discovery was surprising at the time since it is totally at odds with the classical picture where, as we have just seen in II.1, under apparently reasonable assumptions, it is a rigorous *theorem* that black holes cannot shrink.

A couple of years after Hawking performed his calculations, he saw that they seemed to imply a certain conceptual puzzle, which has since been called the information paradox. This puzzle and its descendants have been a driving force in the search for a unified theory. Without delving deeply into the paradox, some basic attitudes may be stated as follows.

Classically, black hole evaporation appears to suggest a failure of global hyperbolicity and a consequent failure in predictability in the sense of initial data. In fact, a better understanding of this sentence is essentially the point of this current article. For a clear introduction to this idea, consider the article by Wald [20], or more recently the review paper [19].

Semi-classically, evaporation appears to produce a transition from a pure inital quantum state to a final mixed state, contra the principle of unitary evolution. There are a number of detailed description of this, starting with [18] and given a clear exposition in [21].

Whether and how the quantum version follows from the classical version is far from understood. Perhaps the only thing which is clear is that mathematically respectable formulations of quantum field theory on curved spacetime [21] have only

been achieved in the setting of globally hyperbolic spacetimes, and, furthermore, that it is not at all clear how to proceed when quantum fields are considered on causally pathological spacetimes.

Though these issues seem to be of fundamental importance, mathematically rigorous efforts directed towards better understanding of them are few and far between. The first attempt in this direction was by Kodama [10] in 1979. He sought to formulate a mathematically rigorous basis for the suspected loss in predictability by way of a theorem in causal structure. His theorem was then rewired and endorsed in an article by Wald [20], though Wald attributes the specific statement to Geroch.

Theorem (Kodama, Geroch, and Wald [10], [20])**.** *Let (M, g) be a spacetime containing achronal subsets S_1, S_2 where S_1, S_2 is edgeless and S_1 is connected. Then $S_2 \not\subset D^+(S_1)$ if the following holds:*

(i) $J^+(K) \cap S_2$ has compact closure, where $K = S_1 - D^-(S_2)$,

(ii) there is a point $p \in D^+(S_1) - J^+(S_2) \cup J^-(S_2)$.

The original motivation for this theorem (which we now call the KGW theorem) was to apply to it to evaporating black hole spacetimes, where S_1 and S_2 are imagined to lie, respectively, before and after the black hole has evaporated, and p is a point in the black hole interior which is causally inaccessible from S_2. In the asymptotically flat case, the spacetimes envisaged are as follows.

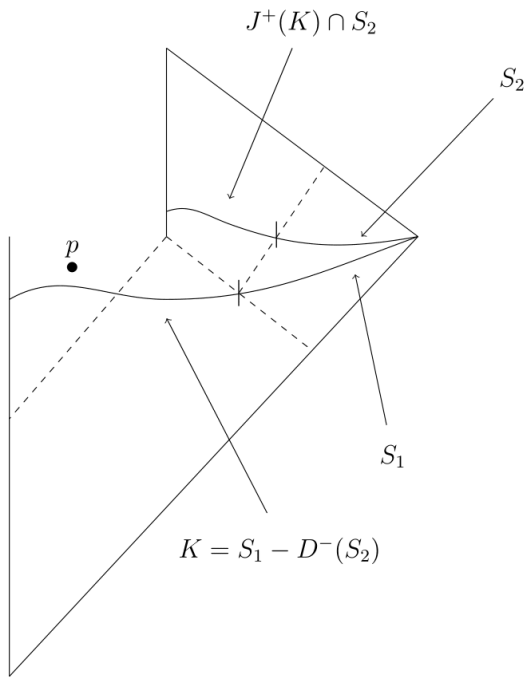


Figure II.3.1. This is the Penrose diagram commonly drawn to describe fully evaporated black hole spacetimes. As drawn, the diagram satisfies the conditions of the KGW theorem, and indeed there results a loss in predictability from S_1.

The KGW theorem appears to be the only mathematically rigorous study of causal structure in the context of evaporating black hole spacetimes. It is thus to be regarded as an important result in a topic where rigour is generally lacking. As it turns out however, the KGW theorem doesn't actually ground the idea that evaporating black hole spacetimes fail to be predictable. Rather, it shows that S_1 isn't a Cauchy surface for M and in particular doesn't rule out the possibility that (M, g) is globally hyperbolic from a different slice. The following trivial example shows that $S_2 \not\subset D^+(S_1)$ could merely be due to a poor choice of S_1 and S_2.

Example. Take two dimensional Minkowski spacetime $(\mathbb{R}^2, \eta)$ in Cartesian co-

ordinates. Remove everything but an open globally hyperbolic diamond with vertices given by $(x = \pm 1, t = 0)$ and $(x = 0, t = \pm 1)$. Remove from this diamond spacetime a further diamond with vertices given by $(x = 0, t = -1)$, $(x = \pm\frac{1}{2}, t = -\frac{1}{2})$, $(x = 0, t = 0)$. This spacetime is globally hyperbolic but the conditions of the KGW theorem are satisfied by suitably choosing S_1 and S_2.

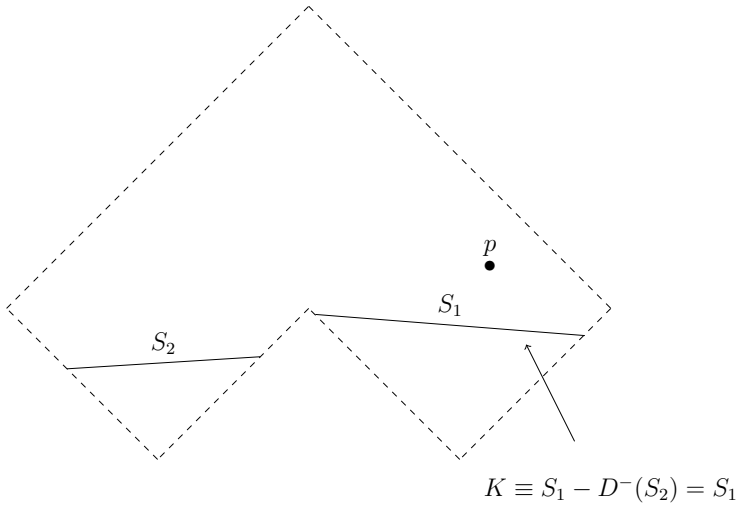


Figure II.3.2. This spacetime satisfies the KGW theorem but is globally hyperbolic. Other examples can be constructed which are less trivial than this one.

In view of such examples (and less trivial ones one can construct), we cannot rely on the KGW theorem alone as a basis for the suspicion that fully evaporating black holes lead to a failure in predictability. Here, I shall offer two theorems showing that spacetimes containing evaporating black holes either fail to be causally simple or causally continuous.[9]

[9]Recall that causal simplicity and continuity sit second and third from top in the causal hierarchy of spacetimes: Global Hyperbolicity ⇒ Causal Simplicity ⇒ Causal Continuity ⇒ Stably Causal ⇒ Strongly Causal ⇒ Distinguishing ⇒ Causal ⇒ Chronological ⇒ Non-totally Vicious.

The first theorem proves the failure of causal simplicity by extending the KGW theorem; in particular, conditions are added which rule out examples like the one above. The second is formulated entirely differently in that it assumes from the outset the existence of an event horizon. Its conclusion is stronger in that (M, g) is shown to be causally discontinuous. For simplicity, the event horizon is defined using Penrose's concept of aymptotic flatness at null infinity. As the proof reveals, the argument can be generalized to spacetimes with different asymptotics.

2. Geometric Preliminaries

We shall need the following standard results; see chapter 3 of [2] for the relevant details and background.

Definition. *A spacetime (M, g) is non-totally imprisoning when no future or past inextendible causal curve is contained in a compact set.*

Definition. *A spacetime (M, g) is said to be past reflecting if $I^+(p) \subset I^+(q) \Rightarrow I^-(q) \subset I^-(p)$. Future reflecting is defined dually. A spacetime is reflecting if it is both future and past reflecting.*

Definition ([2])**.** *The set-valued function I^+ is inner continuous at $p \in M$ if, for each compact set $K \subseteq I^+(p)$, there exists a neighborhood $U(p)$ of p such that $K \subseteq I^+(q)$ for each $q \in U(p)$.*

Definition ([2])**.** *The set-valued function I^+ is outer continuous at $p \in M$ if, for each compact set $K \subseteq M \backslash \overline{I^+(p)}$, there exists some neighborhood $U(p)$ of p such that $K \subseteq M \backslash \overline{I^+(q)}$.*

Lemma (Hawking-Sachs [9])**.** *For any spacetime (M, g) the outer continuity of I^+ and I^- is equivalent to the reflecting property.*

Definition ([2])**.** *A spacetime (M, g) is said to be causally continuous if and only if it is distinguishing and the set-valued functions I^+ and I^- are outer continuous.*

Definition. *A spacetime (M, g) is said to be causally simply if it is causal and $J^{+(-)}(p)$ is closed for all $p \in M$.*

Definition ([4])**.** *Let $S \subset M$ be achronal. Then $p \in S$ is an edge point of S provided every neighborhood $U(p)$ of p contains a timelike curve γ from $I^-(p, U)$ to $I^+(p, U)$ that does not meet S. We denote by edge(S) the set of edge points of S.*

Proposition ([4])**.** *Let S be closed. Then each $p \in \partial I^+(S) \backslash S$ lies on a null geodesic contained in $\partial I^+(S) \backslash S$, which either has a past endpoint on S, or else is past inextendible in M.*

Proposition ([4])**.** *Let S be an achronal subset of a spacetime M. Then $H^+(S) \backslash edge(S)$, if nonempty, is an achronal C^0 hypersurface of M ruled by null geodesics, each of which is either past inextendible in M or has a past endpoint on edge(S).*

The Lorentzian distance function $d(.,.)$ and the Lorentzian length $l_\gamma(p, q)$ is defined as in [23], from which we shall also use the following definition.

Definition (Definition 2.11 of [23])**.** *A continuous causal curve $\gamma : I \to M$ is maximizing if, for every $t_1, t_2 \subset I$, $t_1 < t_2$, $d(\gamma(t_1), \gamma(t_2)) = l(\gamma \mid_{[t_1,t_2]})$. A sequence of continuous causal curves $\gamma_n : I_n \to M$, is limit maximizing if defined*

$$\epsilon_n = \sup_{t_1,t_2 \in I_n, t_1 < t_2} [d(\gamma_n(t_1), \gamma_n(t_2)) - l(\gamma_n \mid_{[t_1,t_2]})] \geq 0$$

it is $\lim_{n\to\infty} \epsilon_n = 0$.

3. Failure of Causal Simplicity

The KGW theorem will now be extended so as to contradict causal simplicity. Recall that a spacetime is causally simple if it is causal and all the causal futures or pasts $J^+(p)$, $J^-(p)$ are closed.

In the KGW theorem, S_1, S_2 are achronal, edgeless (and thus closed by [15] p.

414, corollary 26) and S_1 is connected. Below, let us make the additional assumption that S_1 is non-compact and that S_2 is connected.

Theorem. *Let (M, g) be a non-totally imprisoning spacetime with connected, achronal, edgeless sets $S_1, S_2 \subset I^+(S_1)$ such that S_1 is closed and non-compact. Suppose the following:*

(i) $J^+(K) \cap S_2$ *has compact closure, where* $K = S_1 - D^-(S_2)$,

(ii) there is a point $p \in D^+(S_1) - J^+(S_2) \cup J^-(S_2)$,

(iii) $S_1 - V \subset J^-(S_2)$, *where* V *is a compact subset of* S_1,

(iv) $J^-(Q) \cap S_1$ *has compact closure where* $Q = S_2 - D^+(S_1)$.

Then (M, g) is not causally simple.

Remark. Condition (iii) blocks the kind of scenario described in the example of section 1, and it also expresses the idea that S_1 and S_2 are related as in the typical representation of evaporating black hole spacetimes, i.e., so that V represents the portion of S_1 that is inside the black hole. Condition (iv) expresses the idea that the non-compact ends of S_2 are contained in $D^+(S_1)$, so that in effect we are assuming a form of predictability 'far away' from the black hole.

Proof. We shall assume causal simplicity and obtain a contradiction.

First we apply the KGW theorem to deduce that $H^+(S_1) \neq \emptyset$. Since S_1 is closed and edgeless, we apply the proposition of section 2 to infer that $H^+(S_1)$ is ruled by past inextendible null geodesics not intersecting S_1.

Second, we observe that $H^+(S_1) \cap S_2 \neq \emptyset$. Since $S_2 \not\subset D^+(S_1)$ and S_2 is connected, $H^+(S_1) \cap S_2 \neq \emptyset$ will follow if we can show that $S_2 \cap D^+(S_1) \neq \emptyset$. Seeking a contradiction, we suppose that $D^+(S_1) \cap S_2 = \emptyset$. In that case, by defining Q to be equal to $S_2 - D^+(S_1)$, we have $Q = S_2$. Condition (iv) now becomes that $J^-(S_2) \cap S_1$ has compact closure. But this is in contradiction with (iii). In particular, since S_1 is closed and non-compact, the assumption that $S_1 - V \subset J^-(S_2)$ where V is a

compact subset of S_1 implies that $J^-(S_2) \cap S_1$ has non-compact closure.

These remarks now permit us to consider a point $x \in H^+(S_1) \cap S_2$. Since $J^-(x)$ is closed, every point $y \in \partial J^-(x)$ is also in $J^-(x)$. By a proposition in section 2, it then follows that for any such point y, there exists a past directed achronal null geodesic with endpoints y and x. These geodesics generate the boundary of the past lightcone of x.

By standard properties of Cauchy horizons, $H^+(S_1)$ is an achronal hypersurface that is generated by null geodesics. By restricting attention to a sufficiently small convex neighborhood $U(x)$ of x, it follows that there is a unique achronal null geodesic η which belongs to both $U(x) \cap H^+(S_1)$ and $U(x) \cap \partial J^-(x)$.

Since $U(x)$ is a convex neighborhood, we may take η to have two endpoints, x and z, where $z \in H^+(S_1) \cap U(x) \cap \partial J^-(x)$. Since S_1 is edgeless, if η is extended towards the past as an unbroken achronal null geodesic, then it will eventually become past inextendible, it will not intersect S_1, and it will remain entirely on $H^+(S_1)$.

The sought after contradiction will be to show that (iv) and the assumption of causal simplicity actually permit us to construct an achronal extension of η with past endpoint on S_1.

Consider the set $I^-(x) \cap S_1$. Clearly, this set is both non-empty and open. By assumption (iv) and by the standard identity $\partial I = \partial J$ - eg., see page 191, chapter 8 of [20] - it follows that $\overline{I^-(x) \cap S_1}$ is compact. Thus, any infinite sequence of points in $\overline{I^-(x) \cap S_1}$ has a limit point in $\overline{I^-(x) \cap S_1}$. We may then consider an infinite sequence of points $\{q_i\} \in I^-(x) \cap S_1$ with limit point $q \in \partial(I^-(x) \cap S_1)$. For each q_i, we consider a future directed timelike curve with past endpoint q_i and future endpoint $r_i \in I^-(x)$ where $r_i \to x$. We label each such curve by $\{\delta_i\}$, each of which are defined on $[0, t_i]$, so that $\delta_i(0) = q_i$ and $\delta_i(t_i) = r_i$ for some $0 < t_i$.

Since q is an accumulation point of a sequence of continuous timelike curves $\{\delta_i\}$, part (1) of theorem 3.1 in [23] permits us to extract a subsequence, $\{\delta_k\}$. Clearly, $q_k << r_k$ for all k and the curves δ_k have endpoints $r_k, q_k \in I^-(x)$. We

now show that causal simplicity implies that we can construct our sequence of curves, $\{\delta_i\}$, such that our subsequence $\{\delta_k\}$ converges to a smooth achronal limit curve δ with endpoints q, x. Since each distinct point on $\partial(I^-(x) \cap S_1)$ gives rise to a distinct limit curve, it follows that one of these coincides with η. Thus, the proof will complete for we will have produced an achronal extension of η with endpoint on $\partial(I^-(x) \cap S_1)$.

To show that δ can be suitably constructed, we consider the following result of Minguzzi.

Theorem. *A non-totally imprisoning spacetime $(M, \{g\})$ is causally simple if and only if, for every metric g in the conformal class $\{g\}$, the Lorentzian distance function $d_g(.,.)$ is continuous on the vanishing distance set.*

This theorem implies that we may take the distance function $d_g(.,.)$ to be continuous on the vanishing distance set. This in turn implies that, for any infinite sequence of points $\{p_n\}$, $\{s_n\}$ with $s_n \in I^+(p_n)$ and $p \in \partial I^+(s)$ where $p_n \to p$, $s_n \to s$, we have $d(p_n, s_n) < \epsilon$ as $n \to \infty$. In our case, given that $q \in \partial I^-(x)$, we have $d(q, x) = 0$. Continuity of $d(.,.)$ implies that $d(q_k, r_k) < \epsilon_k$, and thus that $d(q_k, r_k) \to 0$ as $k \to \infty$. By a definition in section 2, this implies that may take the sequence $\{\delta_k\}$ to be limit maximizing. By part (1) of theorem 3.1 in [23], the subsequence $\{\delta_k\}$ converges h-uniformly on compact subsets to a maximizing limit curve δ. Seeing as $q \in \partial I^-(x)$, δ must in fact be an achronal null geodesic with past, future endpoint q, x. □

It is likely that recourse to Minguzzi's theorem is unnecessary for the result of this section. Its inclusion here has the advantage of making the current argument particularly explicit.

4. Causal Discontinuity

Let us now consider an alternative approach whose virtue lies in proving a stronger conclusion. The price to be paid for this is the posit of an event horizon along with a black hole region obeying certain specific properties. In truth, these properties are entirely standard and they are in any case part of the arguments in favor of evaporation.

To formulate the theorem below, I will specify asymptotic boundary conditions in the form of the following definition, due to Penrose [17], of *asymptotic flatness at null infinity.*

Let (M, g) be a four dimensional, chronological, connected spacetime which can be conformally included into a spacetime-with-boundary (M', g') such that M is the interior of M', $\partial M = M' \backslash M$. Assume, with regards to the conformal factor, that there exists a smooth function Ω on M' such that:

(i) $\Omega > 0$ and $g' = \Omega^2 g$ on M,

(ii) $\Omega = 0$ and $d\Omega \neq 0$ along M'.

The boundary $\mathcal{J} \equiv \partial M'$ is assumed to consist of two components, $\mathcal{J}^+$ and $\mathcal{J}^-$, which are smooth null hypersurfaces representing, respectively, future and past null infinity. A spacetime (M, g) satisfying the above is said to be *asymptotically flat at null infinity.*

We now prove the following theorem.

Theorem. *Let (M, g) be a spacetime which is asymptotically flat at null infinity, such that its conformal boundary consists of two disconnected null components $\mathcal{J} \equiv \mathcal{J}^+ \cup \mathcal{J}^-$ each having topology $V \times \mathbb{R}$. Suppose that the following properties obtain:*

(i) there is a non-empty event horizon $\partial I^-(\mathcal{J}^+)$ and a non-empty black hole region defined by $B \equiv I^+(\partial I^-(\mathcal{J}^+))$ such that $\partial B = \partial I^-(\mathcal{J}^+)$ and $B \cap I^-(\mathcal{J}^+) = \emptyset$,

(ii) $\partial I^-(\mathcal{J}^+) \subset \overline{I^-(\Sigma)}$ *where* Σ *is a complete cross section of* $\mathcal{J}^+$*, i.e., a spacelike embedded submanifold of* $\mathcal{J}^+$ *with topology* V*.*

Then (M, g) *is causally discontinuous.*

That the event horizon is the boundary of a black hole region from which there is no escape to $\mathcal{J}^+$, and that $\mathcal{J}^+$ has topology $\mathbb{R} \times V$, are both conditions which are standard in the usual setting. Note that we need not assume that V has topology S^2.

The main new conditions capturing the notion of evaporation is (ii), which demands that the event horizon is entirely contained in the closure of the past of some cross section of $\mathcal{J}^+$. Here, we are imagining that the event horizon eventually ceases to exist, and thus, that causal curves γ which persist sufficiently far into the future have their pasts $I^-(\gamma)$ which almost make contact with the event horizon so that in particular their closures $\overline{I^-(\gamma)}$ could intersect portions of the event horizon. The diagram below provides a means of visualizing the kinds of spacetime for which the above theorem applies.

Proof. We assume both outer continuity and reflectivity and derive a contradiction.

Before going further, we note that by the above assumptions, there exists a non-trivial causal curve γ with future endpoint $y \in \mathcal{J}^+$ where y is the future endpoint to a non-trivial achronal null geodesic generator of $\mathcal{J}^+$ with past endpoint on Σ; consider the figure above.

First we make a preliminary observation. For any point $x \in \partial I^-(\mathcal{J}^+)$, we have by assumption that $x \in I^-(B) \cap \partial B$. Let $U(x)$ be a small open neighborhood of x. Consider the spacetime given by $(U(x), g')$ where g' is the spacetime metric of M restricted to $U(x)$. In this spacetime, $\partial I^-(\mathcal{J}^+) \cap U(x)$ is an achronal boundary without edge in $(U(x), g')$. Denote the achronal boundary $\partial I^-(\mathcal{J}^+) \cap U(x)$ in $(U(x), g')$ by V. By proposition 3.15 in [32] on achronal boundaries, V separates the spacetime $(U(x), g')$ such that $U(x) = I^+(V) \cup V \cup I^-(V)$. It then follows that any point in $I^-(V)$ is in $I^-(B)$, that any point in $I^+(V)$ is in B and that any point

on V is in $\partial I^-(\mathcal{J}^+)$.

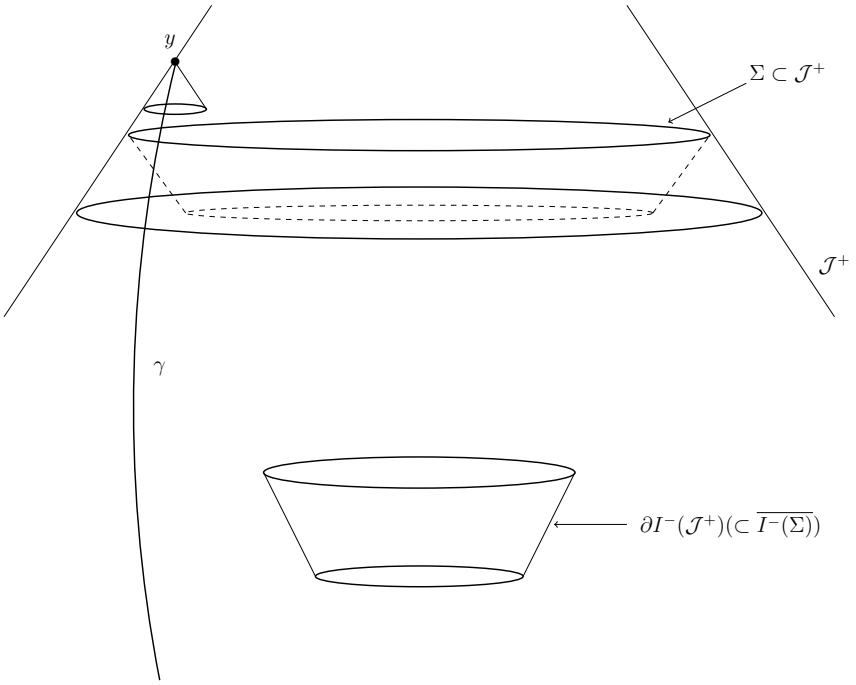


Figure II.3.3. The causal curve γ eventually reaches $\mathcal{J}^+$ and as a consequence of the event horizon $\partial I^-(\mathcal{J}^+)$ eventually disappearing, we have that $\partial I^-(\mathcal{J}^+) \cap \overline{I^-(\gamma)} \neq \emptyset$. We don't require that the whole event horizon be contained in the closure of the past of such a curve however. That instead is reserved to Σ, which is an entire cross section of $\mathcal{J}^+$.

By assumption, γ has endpoint $y \in \mathcal{J}^+$, and there is a point on the event horizon which is also in $\partial I^-(y) \cap M$. Denote such a point by x. Since $\mathcal{J}^+$ is open, we may consider an open neighborhood $U(y)$ of y such that $U(y) \subset \mathcal{J}^+$. In this neighborhood, we may consider a point $y' \in U(y)$ such that $I^-(y) \subsetneq I^-(y')$. Note that by condition (i), it follows that $I^-(y') \cap B = \emptyset$, and thus $I^+(x) \cap I^-(y') = \emptyset$.

Now consider a compact set $K \subset M$ with non-empty interior such that $K \subset$

$I^-(y') - [\overline{I^-(y)} \cup \overline{I^-(\Sigma)}]$. We know that such a set exists because (ii) implies that the region $I^-(y') - [\overline{I^-(y)} \cup \overline{I^-(\Sigma)}]$ is non-empty. Moreover, we know that such a set lies in an open region because I^- is open and $\overline{I^-}$ is closed.

By suitably choosing a compact set K and assuming reflectivity, we shall obtain a contradiction with outer continuity. To better follow the proof, note the diagram below.

First, we show that K can be chosen such that $K \subset M - \overline{I^+(x)}$. Since we have chosen K such that $K \subset I^-(\mathcal{J}^+)$, it follows that $K \not\subset I^+(x)$. This means that, if K intersects $\overline{I^+(x)}$, then it does so in a closed and achronal set, which we write as $Z \equiv K \cap \partial I^+(x)$. We note that we are here choosing K such that $K \not\subset Z$. This is clearly possible for K was originally chosen in the open region $I^-(y') - [\overline{I^-(y)} \cup \overline{I^-(\Sigma)}]$.

Let z be any point in Z. By the condition that $K \cap \overline{I^-(\Sigma)} = \emptyset$ and that $\overline{I^-(\Sigma)}$ contains $\partial I^-(\mathcal{J}^+)$, it is clear that $z \notin \partial I^-(\mathcal{J}^+)$. Thus, z lies neither on the event horizon nor inside the black hole region $B = I^+(\partial I^-(\mathcal{J}^+))$. It follows, then, that z is in $I^-(\mathcal{J}^+)$. Since this is true for all points in Z, we conclude that Z, which is an achronal set in $\partial I^+(x)$, is in $I^-(\mathcal{J}^+)$. Now, Z is either empty or not. If it is then of course $K \subset M - \overline{I^+(x)}$. If $Z \neq \emptyset$, we can show that we can amend our choice of K to obtain a compact set which is in $M - \overline{I^+(x)}$. In particular, we consider a new compact subset K' of K, such that $K' \in I^+(z)$ for some point $z \in Z$. By achronality of Z, $K' \cap Z = \emptyset$. Clearly, then, $K'(\subset K)$ is a compact set avoiding $\partial I^+(x)$ and thus $K' \subset M - \overline{I^+(x)}$. Thus, we may always take our compact set K to such that $K \subset M - \overline{I^+(x)}$ and $K \subset I^-(y') - [\overline{I^-(y)} \cup \overline{I^-(\Sigma)}]$.

If (M, g) is outer continuous at x, then there is an open neighborhood $O(x)$ of x such that for any $x' \in O(x)$, $\overline{I^+(x')} \cap K = \emptyset$. As remarked above, any open neighborhood of x contains a subneighborhood $U(x)$ such that $U(x) \cap I^-(\mathcal{J}^+) \neq \emptyset$. Let x' be any point in this intersection, i.e., $x' \in U(x) \cap I^-(\mathcal{J}^+)$. By conditions on γ, y, y', for a sufficiently small subneighborhood $U(x)$, any such x' lies in $I^-(y')$.

Now consider a point w in $I^-(k) - K$ for some point $k \in int(K)$. Choose w

such that $x' \in I^-(w)$. This choice is possible in virtue of the fact that $K \subset I^-(y')$ and $x' \in I^-(y')$. Now, if $x' \in I^-(k)$, then $I^-(x') \subset I^-(k)$. By reflectivity this implies that $I^+(k) \subset I^+(x')$. Since $K \cap I^+(k) \neq \emptyset$, we have $K \cap I^+(x') \neq \emptyset$. So $K \not\subset M - \overline{I^+(x')}$. Since $U(x)$ can be chosen to be arbitrarily small and x' is an arbitrary point in $U(x)$, (M, g) is not outer continuous at x. □

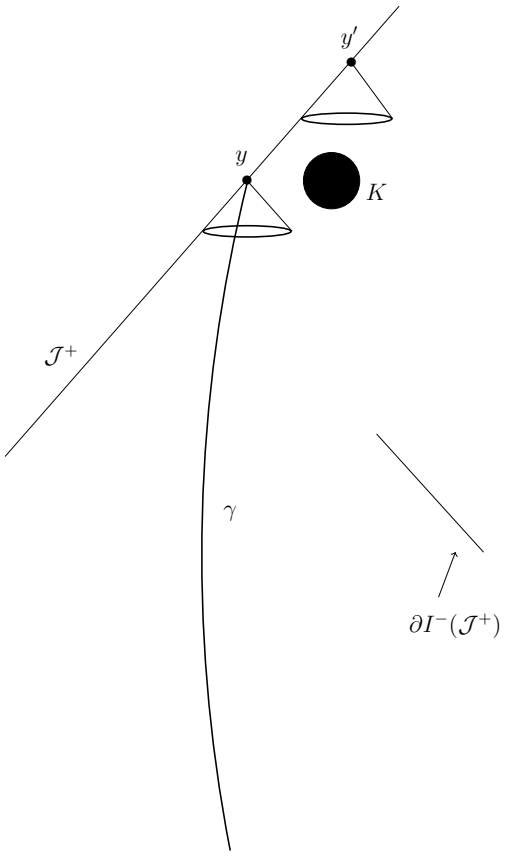


Figure II.3.4. This diagram behind the contradiction that is eventually set up with causal continuity.

5. Discussion

On causal discontinuity.

Hawking and Sachs [9] formulated causal continuity in search of a property weaker than global hyperbolicity but stronger than stable causality. They argue that global hyperbolicity may be too strong for a number of physically interesting models, and that stable causality includes spacetimes with curious causal properties. They consider the example of 2-dimensional Minkowski spacetime with a spatial segment removed. This example is stably causal, but it has the following pathological feature: the extent of the spacetime covered by $I^-(p)$ changes discontinuously upon arbitrarily small perturbations of the metric. Their proposal of causal continuity is at heart an attempt to prohibit such pathologies.

By the theorem of section 4, if spacetimes containing fully evaporated black holes are causally continuous, then one or more of the assumptions must be violated. The appeal of the theorem is that each condition is entirely standard and always assumed in semi-classical models of evaporating black holes. Thus, if Lorentzian manifolds provide a suitable background for semi-classical physics, then either they are causally discontinuous, or there are major semi-classical departures from the standard picture of evaporation.

On predictability.

Wald [20] remarks that non-global hyperbolicity need not imply the failure of predictability. He considers the example of a massless Klein Gordon scalar field ϕ in Minkowksi spacetime with a point removed, i.e., $(\mathbb{R}^4 \backslash p, \eta)$. If we take S_1 and S_2 to lie, respectively, below and above the removed point, then $S_2 \not\subset D^+(S_1)$, as in the conclusion of the KGW theorem. Yet Wald argues that, given suitable initial data, ϕ can still be determined globally and uniquely on $(\mathbb{R}^4 \backslash p, \eta)$.[10] On the basis of this example, Wald suggests that fully evaporating black hole spacetimes could be of

[10]Though his argument is heuristic, it looks as though it could be made rigorous and generalized to other matter fields.

this kind, i.e., non-globally hyperbolic but still permitting unique global evolution of matter fields residing in the spacetime.[11]

The important point here is that Wald's example *is* causally continuous. This raises the question of whether his argument goes through in causally discontinuous spacetimes. On the basis of some simple examples constructed by removing sets bigger than mere points in Minkowski spacetime, it seems not. As such, one can expect that a spacetime which fails to be reflective (and is thus causally discontinuous) prevents a system of hyperbolic PDEs from admitting globally unique solutions determined by initial data. Such a result would express, clearly and rigorously, once and for all, the suspicion that full evaporation conflicts with predictability already at the classical level.

On Maudlin, Manchak, Wallace, and Weatherall.

Maudlin [12] has proposed a view on the information paradox, where by 'paradox' he means the apparent inconsistency between the transition from a pure initial to a mixed final state and the quantum mechanical principle of unitarity. He has argued that there isn't really any paradox, and that it suffices to add a point at the evaporation event. This would allow one to draw a 'slice' through the typical Penrose diagram. In virtue of this new 'slice' being traversed by every inextendible timelike curve, Maudlin calls it a 'Cauchy surface'. His argument is that in the presence of this newly adorned 'Cauchy surface', predictability can be restored.

Very soon after Maudlin's paper appeared, Manchak and Weatherall pointed out [11] that his proposal suffers from various drawbacks. Not only does his model appear to be causally discontinuous (and thus not globally hyperbolic after all), they also note that it fails to be a manifold. The failure of being a manifold seems quite more damaging than the information loss that Maudlin was hoping to restore. A key point of Manchak and Weatherall's arguments is that, in virtue of the KGW theorem, evaporating black hole spacetimes aren't globally hyperbolic. The current

[11] Private communication: Wald still stands by the arguments in [20].

theorem can also be used in their argument, which in fact yields the stronger result that such spacetimes aren't even causally continuous.

Manchak and Weatherall go on to argue that the failure of global hyperbolicity conflicts with Penrose's 1979 conjecture of strong cosmic censorship, which posits that physically reasonable spacetimes are globally hyperbolic. Note that this interpretation is quite unlike the current view of mathematical relativists working on the SCC. For as described in III, the current mathematical content of the SCC is viewed as a hypothesis about *classical* general relativity in the context where there exists a maximal globally hyperbolic development. By contrast, evaporation relies on the presence of quantum fields and these badly violate all local classical energy conditions.

In my view, which is the right view regarding the SCC is mostly a matter of inclination, and as for the 'information paradox', if there is something paradoxical about evaporation, then it ought to be expressed in terms of the semi-classical physics that is thought of as responsible for the paradox. Whatever that may be, it is quite distant from the current mathematician's take on the SCC.

On this point, it seems worth mentioning the article of Wallace [22]. Wallace takes an entirely different approach and tackles evaporation head on as it is discussed in the mainstream physics literature. In that literature, there is a different paradox which has also featured prominently, the so-called Page time paradox. Wallace explains why that paradox is genuine, but his arguments go beyond the scope of this chapter.

AdS/CFT and Holography

In the last two decades, there have been a lot of interest toward the general terms of 'AdS/CFT', or more generally 'holography'. There is a widely discussed idea where evaporating black holes are modelled in spacetimes with timelike asymptotic boundaries, such as for instance the conformal boundaries of the Anti de Sitter or Schwarzchild Anti de Sitter spacetimes. In that setting, owing to a so-called du-

ality linking a quantum gravitational theory in the bulk with a purely quantum field theory on the boundary, it has been suggested that full black hole evaporation does not in fact lead to a failure of predictability.[12] Here, I merely wish to point out that the theorem of section 4 combined with recent results in [1] lead to an interesting perspective with regards to these ideas. To make the point, I shall need to make some preliminary observations.

First, as mentioned, the proof of the theorem in section 4 reveals that it is clearly possible to generalize it to the case of spacetime with timelike boundaries.

Second, the recent study [1] has offered a detailed investigation of the causal structure of spacetimes with timelike boundaries, formally written $\overline{M} = M \cup \partial M$. In virtue of having a timelike boundary, the spacetime $int(\overline{M})$ cannot be globally hyperbolic in the usual sense, eg., like the Anti de Sitter solution. That said, pacetimes with boundaries of the form of $\overline{M}$ can satisfy a stand-in condition for global hyperbolicity which is termed 'globally-hyperbolic-with-timelike-boundary'. This definition is defined analogously to the usual notion of global hyperbolicity, and it is the best causal property that spacetimes with timelike boundaries can satisfy. With this definition (which I omit recalling here), the authors of [1] prove that if $\overline{M}$ is globally-hyperbolic-with-timelike-boundary, then $int(\overline{M})$ is a causally continuous spacetime. Combining this with a timelike version of the theorem in section 4, we have that an evaporating black hole spacetime with timelike boundary cannot be globally-hyperbolic-with-timelike-boundary. So how might this be an interesting conclusion?

An important fact about spacetimes that are globally-hyperbolic-with-timelike-boundary is that they admit a smooth $3+1$ foliation (as in the globally hyperbolic case), and moreover that given suitable initial *and* boundary data, the initial - boundary value problem can be shown to be well posed under certain circumstances.

As such, we infer that evaporating black hole spacetimes, even if accompanied with a timelike conformal boundary, do not satisfy conditions that are sufficient for

[12]Note that a mathematically rigorous underpinning of this idea is lacking.

a well posed initial - boundary value problem. Although strictly speaking these are sufficient but perhaps necessary, it is my guess that no such initial - boundary value problem can be well posed if their interiors fail to be reflective. It would be great to have such a result in this direction, for it would provide a rigorous argument that evaporating black hole spacetimes, even if supplemented with boundary data at timelike conformal infinity, fail to be predictable in the classical sense.

If such a result can be proved, and if the aforementioned AdS/CFT argument is correct[13], then we would have a situation whereby there is a failure of classical predictability, but a restoration of predicatibility in the semi-classical (or fully quantum) context. This would be a wonderful thing to better understand.

[13]That is, if information is *not* in fact lost for evaporating black holes with timelike conformal boundaries.

References

[1] Aké, L., Flores J. L., Sanchéz, M., Structure of globally hyperbolic spacetimes with timelike boundary, arXiv:1808.04412 [gr-qc]

[2] Beem, J. K., Ehrlich, P. E., and Easley, K. L.: Global Lorentzian Geometry. New York: Marcel Dekker Inc. (1996)

[3] Bernal A. N., Sánchez, M., Globally hyperbolic spacetimes can be defined as "causal" instead of "strongly causal", Class.Quant.Grav. 24 (2007) 745-750

[4] Galloway, G. J., Notes on Lorentzian Causality, http://www.math.miami.edu/g̃alloway/vienna-course-notes.pdf

[5] Hawking, S. W., 1974. Black hole explosions. Nature 248 (5443), 30-31.

[6] Hawking, S. W., 1975. Particle creation by black holes. Communications in mathematical physics 43 (3), 199-220.

[7] Hawking, S. W., 1976. Breakdown in predictability in gravitational collapse. Physical Review D 14 (10), 2460.

[8] Hawking, S. W., Ellis, G. F. R., 1973. The Large Scale Structure of Space-time. Cambridge University Press, New York.

[9] Hawking, S. W., Sachs, R., 1974. Causally continuous spacetimes. Communications in Mathematical Physics 35 (4), 287-296.

[10] Kodama, H., 1979. Inevitability of a naked singularity associated with the black hole evaporation. Progress in Theoretical Physics 62, 1434-1435.

[11] Manchak, J., Weatherall, J., Found Phys, (2018). 48: 611.

[12] Maudlin, T., (Information) paradox lost (2017). arXiv:1705.03541 [physics.hist-ph]

[13] Minguzzi, E., 2008. Limit curve theorems in Lorentzian geometry. J.Math.Phys.49:092501,2008

[14] Minguzzi, E., Sanchéz, M., The causal hierarchy of spacetimes, in H. Baum and D. Alekseevsky (eds.), vol. Recent developments in pseudo-Riemannian geometry, ESI Lect. Math. Phys., (Eur. Math. Soc. Publ. House, Zurich, 2008), p. 299 – 358, ISBN=978-3-03719-051-7, arXiv:gr-qc/0609119

[15] O'Neill, B., Semi-Riemannian Geometry with Applications to Relativity, Academic Press (29 July 1983).

[16] Penrose, R., Techniques of differential topology in relativity, Society for Industrial and Applied Mathematics (1 Jun. 1972).

[17] Penrose, R., 1965. Zero rest-mass fields including gravitation: asymptotic behavior, Proc. Roy.Soc. Lond. A 284 159-203.

[18] Unruh, W., 1976, Notes on black hole evaporation. Physical Review D 14 (4), 870.

[19] Unruh, W., Wald, R., 2017. Information loss. Reports on Progress in Physics 80 (9), 092002.

[20] Wald, R., 1984, Black holes, singularities and predictability. In: Christensen, S. M. (Ed.), Quantum theory of gravity. Essays in honor of the 60th birthday of Bryce S. DeWitt. CRC Press, Boca Raton, FL, pp. 160-168.

[21] Wald, R., 1994, Quantum field theory in curved spacetime and black hole thermodynamics. University of Chicago Press, Chicago, IL.

[22] Wallace, D., (2017), Why black hole information loss is paradoxical, arXiv:1710.03783 [gr-qc].

Appendix. Proof of the KGW Theorem

Proof. First, note that by by p.414 corollary 26 and also p.413, proposition 25 of [15], that S_1 and S_2 are achronal and edgeless implies that these are also embedded topological hypersurfaces (see also the argument in proposition 6.3.1 of [14]). This will be used below.

Suppose, for contradiction, that $S_2 \subset D^+(S_1)$.

Then we can consider a smooth, non-vanishing timelike vector field t on M. Since S_1 and S_2 are achronal, no integral curve of t can intersect either of these sets more than once. However, by $S_2 \subset D^+(S_1)$, each integral curve which intersects S_2 must intersect S_1. Thus, by following these integral curves we obtain a one to one projection f from S_2 onto a subset A of S_1. Since t is smooth, and S_1 and S_2 are topological manifolds transverse to the $t-$flow, $f : S_2 \to A$ is a homeomorphism. This implies that $A = f(S_2)$ must be an open subset of S_1. However, we can prove that A is closed as follows.

Let $q \in \overline{A}$ and suppose $q \notin A$. Then, since $q \notin A$, we clearly have a future inextendible timelike curve from q which does not intersect S_2, and hence $q \notin D^-(S_2)$ - see proposition 6.5.1 of [14]. This implies that $q \in K$. On the other hand, since $q \in \overline{A}$, there exists a sequence $\{q_n\}$ in A which converges to q. Let $r_n = f^{-1}(q_n)$. Since infinitely many q_n enter $int(K)$, infinitely many r_n lie in $J^+(K) \cap S_2$. By compact closure, the sequence $\{r_n\}$ must have an limit point r in S_2. By continuity of f, we must have $f(r) = q$. This contradicts the assumption that $q \notin \overline{A}$ and thus proves that A is closed.

Since S_1 is connected, the fact that $f(S_2) = A$ is both open and closed means that $f(S_2) = A = S_1$. However, this is impossible since the integral curve of t which passes through the point $p \in D^+(S_1)$, $p \notin J^+(S_2) \cup J^-(S_2)$ must intersect S_1 at a point not lying in A. Thus $S_2 \not\subset D^+(S_1)$. □

III. Reasoning on the Basis of Past Lightcones

1. Introduction

Our focus will be on the following question: in general relativistic spacetimes, to what extent can observers make predictions and / or draw conclusions about the global structure of their spacetime from the contents of their past lightcones? To my knowledge the first studies of this question were published in a collection [10] of individual contributions by Glymour, Malament, and Geroch. The basic set-up for the question is as follows.

We consider an observer X in a spacetime (M, g) whose existence we represent by a timelike curve γ. For every point $p \in \gamma$, there is a past lightcone $I^-(p)$ representing the entire past of X at p. Assuming X to be all knowing about their past, we consider whether that would be sufficient for X to infer things about $M \backslash I^-(p)$. That is, we ask what observers can in principle infer about their spacetime from knowledge their own past, in the idealized setting where they are maximally knowledgeable about their past. If ideal observers do well, then perhaps more realistic ones can do too. But if it turns out that even ideal observers suffer from genuine epistemic constraints, then those facing realistic observers can only be as bad if not worse.

It pays to think about what one might say in the non-relativistic setting of say Newtonian (or its geometrized cousin Newton-Cartan) theory. Though I shall say more about this below, a major distinguishing feature of relativistic spacetime theories and in particular GR is that it is a theorem that, for every point $p \in (M, g)$, there exists a globally hyperbolic neighborhood $O(p)$ of p with its own Cauchy surface $S_{O(p)}$; see figure III.1. Global hyperbolicity, as I'll describe, is a causal property that facilitates predictions. Consequently, the theorem means that, at least quasi-locally, knowledge of $S_{O(p)}$ would put the observer in good stead to make predictions about $O(p)$. Note, however, that $S_{O(p)}$ cannot actually lie in $I^-(p)$ as drawn in Figure III.1, and as such that this theorem isn't to be thought of as bearing *directly* on the observer's epistemic constraint.

My reason for mentioning this theorem is that there is no analog in the non-

relativistic setting. Typically, in those settings, as described by Earman [9], one needs knowledge of an initial data set along with an appropriate set of boundary conditions; see figure III.2. The need for such boundary conditions stems from the infinite propagation speed which such non-relativistic theories permit, and the consequence is that, at least quasi-locally, predictability is easier in relativistic theories.

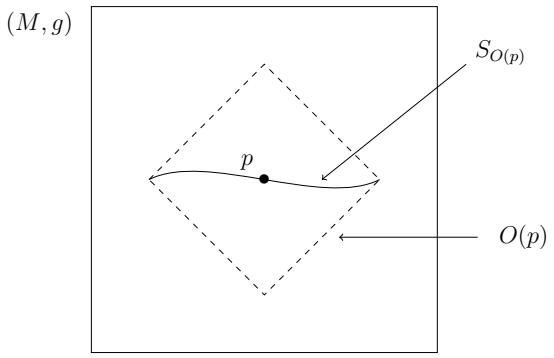


Figure III.1. Given any point p in any spacetime (M, g), there always exists a globally hyperbolic neighborhood $O(p)$, here with Cauchy surface $S_{O(p)}$. Knowledge of $S_{O(p)}$ is (almost) sufficient to know $O(p)$, and it is always guaranteed that every point admits such a neighborhood. That said, an observer at p cannot actually collect the data on $S_{O(p)}$ from p; and so this theorem, though useful, doesn't quite capture what we're after.

The key term here is 'quasi-locally'. For when we consider the global properties of a spacetime, things start to look less favourable for GR. The difference is that, in the non-relativistic setting, once the initial data set is known along with the relevant boundary conditions, then in a sense that's all there is to know. In GR however, knowledge of some initial data set is in general insufficient to recover the whole spacetime.

Coming back to the observer's point of view, the main results in GR were formulated by Malament in [18] and then later significantly generalized by Manchak

[19], [22]. Together, their work suggests that even in the aforementioned idealized setting, knowledge of $I^-(p)$ tends to be insufficient to make predictions or indeed to know things about $M \backslash I^-(p)$. Moreover, their sense of underdetermination remains even if we impose the laws of general relativity, which here I take to entail that spacetimes are Lorentzian manifolds on which Einstein's equations hold, possibly along with various energy conditions.

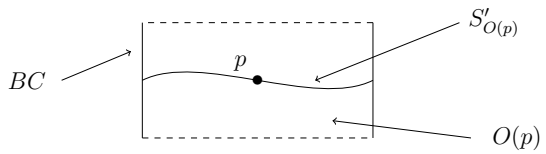


Figure III.2. Here, unlike the previous case, knowledge of $S'_{O(p)}$ doesn't get you as far, and additional boundary conditions are needed. As above, this isn't quite the relevant diagram for what the observer can know on the basis of their past, since data collection of the BC doesn't happen until the observer leaves $S'_{O(p)}$.

To better understand the meaning of their beautiful results, I shall formulate various new ones that provide sufficient conditions under which their underdetermination results disappear.

My description will also differ from theirs in that I've decided to refer and make analogies with more traditional philosophical concepts discussed in the context of the problem of induction, though my reason for doing so is mainly expository.

To summarize, we will now be steered by the following questions: are GR observers epistemically worse off than ones in non-relativistic spacetimes, and does knowing the laws of a spacetime help as much as it does in the non-relativistic setting?

1.1 The old problem of induction

The problem of induction concerns the warrant that we have (or lack) in asserting certain statements on the basis of observations, in particular those we think of as having well defined truth values.[14] The problematic part is that such assertions seem only justifiable by further inductive arguments, which in turn makes any business of justification seem circular. In view of this, the basic questions that arise are: what ought we make of this circularity, and whether it is truly problematic?

At the risk of stating the obvious, we often judge the soundness of an assertion on the basis of our belief in something that grounds our inference. In the context that will be relevant to us, our belief is grounded by what we usually call laws of Nature.

For instance, we believe various assertion around what might happen under certain circumstances because we believe in the existence of physical laws. We may come to better understand these law and find more refined accounts of them, but the point is that we believe there to be something regular and unfailing that accounts for such phenomena.

In that case, the existence of such laws seem to ground inductive inferrences, at least those which follow from whatever law is at hand. It seems however that this steers the discussion into questions about their existence, and on this point opinions vary as to whether the existence question is truly of interest.

Quine. "One part of the problem of induction, the part that asks why there should be regularities in nature at all, can, I think, be dismissed. That there are or have been regularities, for whatever reason, is an established fact of science; and we cannot ask better than that. Why there have been regularities is an obscure question, for it is hard to see what would count as an answer." [27] pg.48

[14] Inductive inferrences needn't be temporal. For example, both 'the sun will rise tomorrow' and 'all swans are white' count as inductive inferrences.

Quine's remark notwithstanding, one can say something rather elementary and deflammatory in response to the issues so far outlined. Note here that I don't mean to say that what follows *solves* the problem of induction, but rather that such responses are available and that they have been defended, for instance by Mellor [22] and van Cleve [3].

The first point is that inductive reasoning is simply an unescapable part of our coming to learn about the world that stems from there being agents with limited understanding, control, and knowledge of the natural phenomena surrounding them.

The second point is that inductive arguments tend to be reliable because the world is so constituted that they often, though not always, lead to correct conclusions.

The third is that we shouldn't seek to justify induction in any absolute sense, certainly not like the kind of justification that we may offer for deductive reasoning. Rather, all that we can really do is describe the various and relative strengths of warrant that we have in making certain inductive inferrences.

The fourth is that, whatever sense of warrant at hand, it is contextual, i.e., specific to particular domains of inquiry, and conditional on various facts being the case, and by 'facts' I mean a mixture of true mathematical and empirical statements.[15]

Finally, as for the laws, they are simply to be thought of as existing above and beyond our epistemic capabilities. This is sometimes referred to as the externalist solution to the problem of induction. The point being that we accept that the existence of laws isn't something that we can actually demonstrate by means of observation. Rather, we may do no more than collect evidence in favour of their existence. In that case, by appealing to the laws but by accepting that we cannot experimentally demonstrate their existence, we may dissolve the problem of induction, or at last one part of it.

[15]There are of course whole swathes of things that I'm sweeping under the rug here, eg., what makes a fact mathematical rather than empirical?

1.2 Spacetimes

I should first justify why the terms 'spacetime' and 'induction' can be put in the same sentence without there being some kind of category mistake. For indeed at face value these terms seem to refer to two very different things, i.e., problems about induction seem to concern philosophy of science more so than science itself. Moreover, by focusing on spacetimes and general relativity (GR), I am already guilty of the inductive leap that GR is a suitable framework to describe a universe that extends beyond our immediately observable region. Those troubled by induction will be troubled by such a leap, and indeed far more so than the issues that I'll be discussing, which ultimately concern how certain spacetimes make for better epistemic environments than others. So how does one justify putting induction and spacetime in the same breath?

At a broad level, my purpose here is to explore how far does the externalist's solution to induction take us in the context of spacetime. That is, we start by saying that the observer X at p in the spacetime (M, g) knows the laws of their spacetime (or of GR say), we assume that the observer is all knowing about $I^-(p)$, and then we ask how much warrant will such an observer have in inferring things about $M \backslash I^-(p)$?

Once we've taken the externalist's view and accepted the existence of laws and granted our knowledge of them, the term 'induction' is only there as a conceptual backdrop. Of course, that it is a good backdrop is something I ought to justify. But on this point, as I shall describe, it turns out that that the geometry of spacetime — points, curves, lightcones, slices, and so on — lends itself quite well to questions of (broadly) inductive character. Before showing this by means of an example below, let us recall some important and fundamental facts about GR.

The field equations of GR are a system of elliptic and hyperbolic equations which enjoy the property of having a well posed Cauchy problem in the sense of Hadamard. That is, given an initial data set satisfying the constraint equations, local spacetime solutions exist, are unique, and depend continuously on the initial

data. In addition to well posedness, a fundamental result of Choquet-Bruhat and Geroch [2] states that for any initial data set there exists a maximal globally hyperbolic development unique up to isometry that is entirely determined by the field equations and the initial data.

For our purposes, these results imply that if a spacetime has the property of being a globally hyperbolic Cauchy development, then it is safer to assert statements about its whole on the basis of partial knowledge of it, for instance about one of its Cauchy surface or perhaps some suitable collection of past lightcones. The warrant which we have in making such assertions are law-like since it is the laws of GR that ground our statements. Moreover, in view of the continuous dependence of solutions from initial data, our knowledge of the initial data needn't be exact; which in turn makes the sense of warrant all the more robust, since otherwise we'd need infinitely precise knowledge of the initial data set to make predictions. Finally, in virtue of global hyperbolicity being equivalent to the ability to admit a foliation by Cauchy surfaces, there is a sense in which such spacetimes lead to a certain regularity between events. In particular, every event is traceable into the future and the past, and every event can be seen as the playing out, according to the field equations, of some set of initial (or final) conditions specified on any one of the Cauchy surfaces that foliate the spacetime. That is, in sum, globally hyperbolic Cauchy developments lead to a great deal of warrant for assertions about (M, g) on the basis of partial knowledge.

More generally, in a general relativistic spacetime, the warrant one has in asserting things about the future or the past based on facts about some initial or final configurations depends on the spacetime's causal structure. And in view of the causal hierarchy of spacetimes, and the fact that global hyperbolicity is its strongest property, global hyperbolicity is actually a rather special property among spacetimes.[16] Note, moreover, that non-globally hyperbolic spacetimes aren't mere

[16]The causal hierarchy is a list of increasingly weak global causal conditions. See part I and or [29] for the full set of definitions.

mathematical curiosities. Take for instance the Kerr-Newman class of vacuum solutions to the field equations, whose exteriors are widely believed to be the asymptotic state of astrophysical black holes[17], and whose interiors contain closed timelike curves (thus failing to be globally hyperbolic). Quite generally, non-globally hyperbolic spacetimes will tend to have Cauchy horizons. These are boundaries defined with respect to some initial data set, beyond which the initial data set no longer fixes the spacetime into which it (isometrically) embeds. In such cases, one can no longer use the laws of GR to uniquely predict what lies beyond the Cauchy horizon.

Though apparently clear, the sense of warrant so far outlined is from a bird's eye view of the spacetime. Let us now turn to the observer's perspective, i.e., one that is actually in the spacetime.

1.3 New Riddles in Spacetime

Above, under the externalist's view of induction, I hinted that the geometry of spacetime lends itself well to questions of inductive character. And for the old problem of induction, this is rather straightforward, i.e., what can observers infer about (M, g) on the basis of $I^-(p)$ and knowledge of the laws in (M, g)? What is perhaps more surprising is that one can also do this for the new riddle of induction, whose origins let us now briefly recall.

In the classic text [15], Goodman devises an example whose purpose it is to point out a deficiency in the so-called 'confirmation theory' of induction, whose formal aspects had been developed cheifly by Hempel [16]. At the time of Goodman's writing, the theory seemed to offer a promising account of what counts as valid inductive inferrence.

Goodman starts with a typical inductive scenario.

Goodman. "Suppose that all emeralds examined before a certain time t are green.

[17] See Chapter 12 of [33] for a more detailed description of the Kerr-Newman class.

At time t, then, our observations support the hypothesis that all emeralds are green; and this is in accord with our definition of confirmation. Our evidence statements assert that emerald a is green, that emerald b is green, and so on; and each confirms the general hypothesis that all emeralds are green." pg. 73-4 [15].

He then defines the following more unusual predicate.

Goodman. "Now let me introduce another predicate less familiar than 'green'. It is the predicate 'grue' and it applies to all things examined before t just in case they are green but to other things just in case they are blue. Then at time t we have, for each evidence statement asserting that a given emerald is green, a parallel evidence statement asserting that that emerald is grue. And the statements that emerald a is grue, that emerald b is grue, and so on, will each confirm the general hypothesis that all emeralds are grue. Thus according to our definition, the prediction that all emeralds subsequently examined will be green and the prediction that all will be grue are alike confirmed by evidence statements describing the same observations. But if an emerald subsequently examined is grue, it is blue and hence not green. Thus although we are well aware which of the two incompatible predictions is genuinely confirmed, they are all equally well confirmed according to our present definition." pg. 74 of [15].

For Goodman, that the confirmation theory doesn't allow us to distinguish between 'green' and 'grue' is its irremediable shortcoming. More broadly, what remains unanswered for Goodman is the following question.

Goodman. "We have so far neither any answer nor any promising clue to an answer to the question what distinguishes law-like or confirmable hypotheses from accidental or non-confirmable ones; and what may at first have seemed a minor technical difficulty has taken on the stature of a major obstacle to the development

of a satisfactory theory of confirmation. It is this problem that I call the new riddle of induction." pg. 80-81 [15].

For Goodman, the resolution of the new riddle lies in understanding better what it is that makes a predicate projectible. By 'projectible predicates', one means predicates α and β whose shared instances all count, for whatever reason, toward confirmation of 'all α are β'. In the case of emeralds, what is it about 'green' and 'grue' that makes one inductively sound and not the other?

Without delving into Goodman's proposed solution and the relevant ensuing literature, I merely wish to point out that a similar game to Goodman's can be played in the context of spacetime, albeit with seemingly more troubling consequences.

Taking the observer's point of view and for the sake of argument, let us assume that our past lightcones, when considered as individual spacetimes, are globally hyperbolic.[18]

Based on this assumption, we could then inductively infer that all of past lightcones of our spacetime, as individual spacetimes, are globally hyperbolic. If true, it would then follow that the entire spacetime is globally hyperbolic.

Like Goodman, one could reason otherwise. First, let us say that for a time orientable spacetime (M, g), a past lightcone $I^-(p) \subset M$ is *grobally hyperbolic* with respect to some time T iff either p lies before T and $I^-(p)$ is globally hyperbolic or p lies beyond T and $I^-(p)$ isn't globally hyperbolic; see figure III.3. By T we may be referring to some affine parameter along some timelike curve γ where $p \in \gamma$, or, alternatively, we can assume stable causality so as to fix a global time function for (M, g).[19] We then say that a spacetime is grobally hyperbolic iff there is a timelike curve $\gamma \subset M$ and a time T (defined either with respect to a global time function or with respect to the affine parameter on γ) such that all past lightcones $I^-(q)$

[18]Perhaps we might seek to justify this by saying that the spacetime regions which we've experienced have seemed to be so.

[19]Stable causality is equivalent to the existence of a global time function which upon fixing makes unambiguous the notion of some time T.

with $q \in \gamma$ are grobally hyperbolic with respect to T.[20] Reasoning as before, our observations also confirm that all our past lightcones are grobally hyperbolic with respect to some time T. Thus, we could propose that in fact all of the past lightcones in our spacetime are grobally hyperbolic with respect to various and possibly distinct times. In that case however, our spacetime fails to be globally hyperbolic.[21]

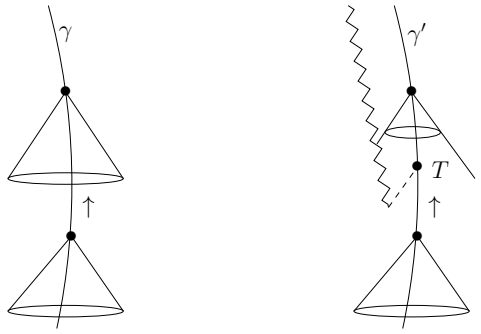


Figure III.3. The spacetime on the left is globally hyperbolic and as such all the past lightcones on γ are globally hyperbolic. On the right, the spacetime is grobally hyperbolic with respect to some T, and as such fails to be globally hyperbolic. The observer on the right inferred the grobal hyperbolicity of her spacetime merely by confirming the trend that all her past lightcones had been so.

So how is this analogous to the new riddle? Ultimately, it looks as though the analog to Goodman's question is whether the global hyperbolicity of our past lightcones is an accidental fact, or if it is part of a law-like generalization. This, of course, depends on the global structure of one's spacetime. This global dependence

[20]Note that Goodman's game can be played for various other spacetime properties of interest.

[21]Recall that it is key for Goodman that there are unexamined emeralds, for it is only then that some emeralds are blue if all emeralds are grue. Note that this is easily verified in our case. For if (M, g) had no unexamined past lightcones, then we would have $I^-(p) = M$ for all $p \in M$. In turn, that would imply that (M, g) is totally vicious, and thus contradict the assumption that at least some past lightcones are globally hyperbolic.

is an interesting phenomenon. For one thing, it indicates that the observer's epistemic status depends on the global structure of their spacetime.

Note also that Goodman purposefully designed 'grue' to be a rather ludicrous predicate, and that there lay the force of his example. The point being that we obviously know that 'green' is safer than 'grue' to reason inductively about, but that the confirmation theory doesn't tell us so. What is intriguing about 'grobal hyperbolicity' is that it is not so obviously tossed aside as a ludicrous. For all we know, our universe might very well be grobally hyperbolic with respect to some time T, and in particular it may be so for reasons that aren't as obviously unappealing as those which could make 'all emeralds grue' true. If anything, that makes the observer's epistemic situation worse. Not only is there is the issue of what one's spacetime is like, there is also the question of which spacetime predicates are good to reason inductively about, and that too depends on the global structure of one's spacetime.

2. Results

The first results on the observer's predictive abilities are in [1], where the authors define a deterministic universe as follows.

Definition (Budic, Sachs [1])**.** *A spacetime* (M, g) *is deterministic iff,* $\forall p \in M$*, each past inextendible causal curve which intersects* $I^+(p)$ *also interesects* $I^-(p)$*.*

The definition says that irrespective of where p lies in M, events occurring in the past of p are sufficient to track the entire causal future of p. In other words the entire future of any point is solely and entirely determined by its past.

Note that this is more a mathematician's definition, than a philosopher's one. In particular, ordinary Minkowski spacetime $(\mathbb{R}^{n\geq 2}, \eta)$ is not deterministic in the above sense, though it is when spatially compactified $(\mathbb{R} \times \mathbb{S}^1, \eta)$.

Given the strength of this definition, it is little surprise that these authors obtain a strong restriction on the spacetime's global causal structure.

Proposition (Budic, Sachs [1]). *A spacetime (M, g) is deterministic iff $\forall p \in M$, $I^-(p)$ contains a Cauchy surface.*

A more recent result by Manchak [19] paints a similar picture. He shows that in order to make verifiable predictions, the spacetime must admit a compact spacelike slice. Manchak's definition comes from Geroch, whereby $p \in M$ is predictable from q iff $p \in D^+(S)$ for some closed, achronal set $S \subset J^-(q)$, and p is verifiable iff $p \in D^+(S)$ and $p \in I^+(q) \backslash J^-(q)$, i.e., so that the observer at q can in principle access the prediction at p.[22] That p be in the domain of prediction of q is then written as $p \in P(q)$.

Also in [19], Manchak proves a result that's a little more subtle to interpret. He starts by considering the following definition.

Definition (Manchak [19]). *Let (M, g) be a spacetime and $q \in M$. We say that $p \in M$ is in the domain of genuine prediction of q, written $\mathcal{P}(q)$, iff $p \in P(q)$ and for all inextendible spacetimes (M', g'), if there is an isometric embedding $\phi : J^-(q) \to M'$, then there is an isometric embedding $\phi' : J^-(q) \cup J^-(p) \to M'$ such that $\phi = \phi'_{|J^-(q)}$.*

The idea here is that genuine predictions guarantee that observers with the same past make the same predictions. Manchak then obtains the following.

Proposition (Manchak [19]). *Let (M, g) be any spacetime and q a point in M. Then $\mathcal{P}(q) \subseteq \partial J^-(q)$.*

That is, the domain of genuine predictions from q, if non-empty, lies on the boundary of the causal past of q, and, thus, is on the verge of being a retrodiction. At first glance the result suggests a failure in one's ability to make predictions on the basis of one's past lightcones. Upon inspection of the proof however, a different attitude presents itself. The result follows by removing a set outside of

[22] When there are closed causal curves through q then $I^+(q)$ may cover some of $J^-(q)$ so that is strictly speaking why we require that $p \in I^+(q) \backslash J^-(q)$; otherwise $I^+(q) \backslash J^-(q)$ is just $I^+(q)$.

$J^-(q)$, considering a copy of the holed spacetime, and then gluing these copies by identifying the relevant holes so as to render the final spacetime inextendible. So the idea of the result is to allow for the spacetime to be ill-behaved outside of $J^-(p)$. One therefore wonders whether it is possible to get around it by assuming the spacetime to be well behaved in some sense. We shall now see how to prove various results in this direction.

For want of a better label, call a spacetime locally Cauchy iff it satisfies some condition sufficient for local existence and uniqueness of solutions to the general relativistic Cauchy problem whenever initial data is posed on a connected acausal edgeless set.

With this definition, we now observe the following.

Proposition. *Let (M, g) and (M', g') be two globally hyperbolic, locally Cauchy spacetimes. Let $S \subset M$ be an acausal Cauchy surface for M (i.e., so that $D(S) = M$) and let $q \in M$ be a point with $S \subset J^-(q)$. Suppose there is an isometric embedding $\phi : J^-(q) \to J^-(q')$. If (M', g') is inextendible, then there is an isometric embedding $\psi : M \to M'$ and $\psi(P(q)) \subset P(q')$.*

Proof. Denote $\phi(S) \equiv S'$. Since ϕ is a local diffeomorphism, S' is compact and edgless. Since S is acausal in (M, g), S' is at least locally acausal in (M', g'), i.e., there is a neighborhood $O(S') \subset M'$ such that S' is acausal as a subset of $O(S')$. For $O(S')$ small enough, $(O(S'), g'_{|O(S')})$ defines a globally hyperbolic spacetime with Cauchy surface S'. But then (M', g') is a globally hyperbolic extension of $(O(S'), g'_{|O(S')})$, S' is consequently a Cauchy surface for (M', g'), and we write $D(S') = M'$.

Consider now the maximal globally hyperbolic development $X(S')$ of $O(S')$. We know $X(S')$ is unique up to isometry. Since $D(S') = M'$ is a globally hyperbolic extension of S', we have $D(S') \subset X(S')$. Since (M', g') is inextendible, $X(S')$ is also inextendible and isometric to (M', g'). We now have an isometric embedding of (M, g) into $X(S')$ and thus into (M', g'), and the conclusion follows. □

The above proposition shows that if we restrict (M, g) and (M', g') to being very good arenas for predictability, then predictions do indeed carry over upon the isometric embedding of $J^-(p)$ into $J^-(p') \subset M'$.

It is worth considering why being locally Cauchy was required in addition to global hyperbolicity. As it turns out and as in the example below, without such a condition, one can have non-isometric globally hyperbolic spacetimes whilst having isometric Cauchy surfaces.

From now on, we say that any acausal edgeless set is a *partial Cauchy surface*.

Example 1. Let $(\mathbb{R}^4, \Omega^2\eta)$ be conformally related to Minkowski spacetime $(\mathbb{R}^4, \eta)$ by means of a bounded, smooth, non-zero, non-trivial function $\Omega : \mathbb{R}^4 \to \mathbb{R}$ which is equal to unity outside a compact set $K \subset \mathbb{R}^4$. Both $(\mathbb{R}^4, \eta)$ and $(\mathbb{R}^4, \Omega^2\eta)$ are inextendible and globally hyperbolic. Moreover, let S be a slice in $(\mathbb{R}^4, \Omega^2\eta)$ such that $S \cap K = \emptyset$. Clearly, for some neighborhood $O(S)$ of S, there is an isometric embedding $\phi : O(S) \to (\mathbb{R}^4, \eta)$. Yet in spite of this, an observer traversing the region where $\Omega^2 \neq 1$ will detect non-trivial curvature that isn't present in $(\mathbb{R}^4, \eta)$. Consequently, although $(\mathbb{R}^4, \eta)$ can be viewed as the maximal Cauchy development of $\phi(O(S))$, the same cannot be said of S and $(\mathbb{R}^4, \Omega^2\eta)$, and as such we can also see that $(\mathbb{R}^4, \Omega^2\eta)$ cannot satisfy the non-vacuum version of Geroch Choquet-Bruhat's theorem.

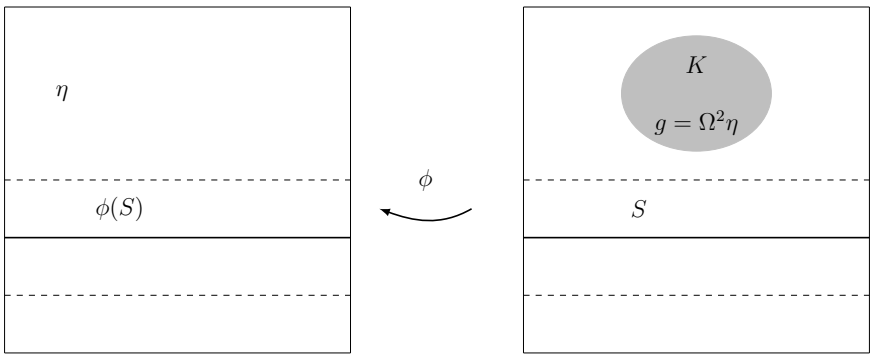


Figure III.4. The spacetime on the left is Minkowski spacetime, and on the right it is a conformally rescaled version of it. Although the slice S and a neighborhood of it isometrically embed into Minkowski spacetime, the respective domains of dependences differ.

One issue with the above proposition is that the strength of its assumptions trivialize the problem somewhat. A better result would be to assume less.

To this end, let us call a spacetime (M, g) standard iff it is locally Cauchy and there are no points $p \in M$ such that $J^-(p) \supset M$. The former condition comes from the above example, and the latter is satisfied in all non causally bizarre spacetimes.[23]

Below, we set up favorable conditions for prediction, and we find that verifiable and in a sense genuine predictions *do* carry over upon the isometric embedding of one's past.

Proposition. *Let (M, g) and (M', g') be two standard spacetimes. Let (M, g) admit a partial Cauchy surface S, and a point $q \in M$ with $S \subset J^-(q)$ such that $J^-(q) \subseteq D(S)$. Then, if there is an isometric embedding $\phi : J^-(q) \to J^-(q')$, there is another isometric embedding $\psi : A \to M'$ with $\psi_{|J^-(q)} = \phi$ such that A is in the domain of verifiable prediction of q and $\psi(A)$ is in the domain of verifiable prediction of q'.*

The proposition is basically captured by the following diagram.

[23] Not like for instance Minkowski spacetime compactified in the time direction $(S^1 \times \mathbb{R}^3, \eta)$.

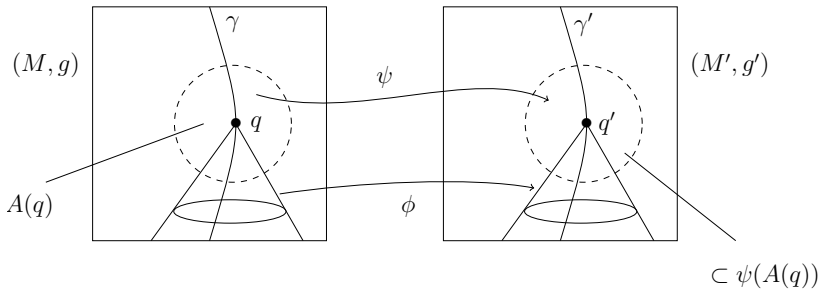


Figure III.5. The spacetime on the left is the one we start in, and on the right is the one into which one's past has been isometrically embedded. As a result of our assumptions regarding q and S, i.e., that $S \subset J^-(q)$, the domain of the isometric embedding ϕ extends to a greater region, now including part of $A(q)$, i.e., the domain of verifiable (and genuine) prediction. Though not drawn, S lies in $J^-(q)$.

Proof. First we make some basic observations. Since $J^-(q)$ lies in a globally hyperbolic set $D(S)$, it is a standard result that $J^-(q)$ is closed. It is also standard that, if S is acausal and edgeless, then in fact $D(S)$ must be open. Indeed, this follows from the property of $\partial D(S)$ being generated by achronal segments intersecting $edge(S)$. From these observations, we can show that the inclusion of $J^-(q) \subseteq D(S)$ is in fact strict. Suppose otherwise that $J^-(q) = D(S)$. In that case $D(S)$ is both open and closed in M. Since M is connected, that implies $D(S) = M$. But then $M = J^-(q)$, in contradiction with being standard. Thus $J^-(q)$ is a closed proper subset of the open set $D(S)$.

Now define $\phi(S) \equiv S'$. Since (M', g') is standard, S' is acausal, and since S' is compact and edgeless, S' is also a partial Cauchy surface in (M', g').

We now observe that any causal curve in $J^-(q')$ with future endpoint q', if sufficiently extended towards the past, intersects S'. Since ϕ is an isometry, there is a neighborhood of $0 \in TM_q$ such that the exponential map of the past non-spacelike vectors in that neighborhood is contained in $D^+(S)$. By the isometry, the same is true for q' with respect to $\phi(D^+(S))$, in particular there is a neighborhood U'_q of q' such that $U'_q \cap J^-(q') \subset \phi(D^+(S))$. Seeking a contradiction, we suppose there

is a past inextendible causal curve γ' with future endpoint q' that does not intersect S'. By the property aforementioned, there is at least a segment of γ' which is contained in $\phi(D^+(S))$. The curve defined by $\gamma \equiv \phi^{-1}(\gamma')$ is causal and ends at q, and is thus entirely contained in $D^+(S)$. But then γ does not intersect S, which is a contradiction since $J^-(q) \subset D(S)$.

By the same argument as before, the inclusion $J^-(q') \subseteq D(S')$ is in fact strict. That is, since S' is acausal and edgeless, $D(S')$ is open, and since we have just seen that $J^-(q')$ is closed (in virtue of lying in a globally hyperbolic subset), if $J^-(q') = D(S')$, then $M' = J^-(q')$, which again contradicts being standard.

We now have two closed sets $J^-(q)$ and $J^-(q)$ each strictly contained in the open sets $D(S')$ and $D(S)$. We now seek to show the existence of an isometry ψ as defined. Although there may be no isometric embedding $\psi : D(S) \to M'$ extending the domain of ϕ, we need only show that the domain of ϕ can be extended to cover at least a portion of $D(S)$.

Consider now the maximal globally hyperbolic development $X(S)$ of S. Since S and S' are isometric, $X(S)$ is both a maximal globally hyperbolic development of S and S'. In virtue of being standard spacetimes, both $D(S)$ and $D(S')$ isometrically embed into $X(S)$. Denote these isometries by $\rho : D(S) \to X(S)$ and $\rho' : D(S') \to X(S)$. Clearly, we have $\rho(D(S)) \cap \rho'(D(S')) \neq \emptyset$. Note that since ρ and ρ' is a local diffeomorphism, both $\rho(D(S))$ and $\rho(D(S'))$ are open. Also note that since $J^-(q)$ and $J^-(q')$ are isometric, we may set ρ and ρ' to be such that $\rho(J^-(q)) = \rho'(J^-(q'))$. Note that $\rho(J^-(q))$ is closed in $X(S)$. We now define the following set

$$[\rho(D(S)) - \rho(J^-(q))] \cap [\rho'(D(S')) - \rho(J^-(q'))] \equiv I$$

Observe that if $I \neq \emptyset$, we can define a non-empty set $A \equiv \rho^{-1}[\rho'(J^-(q')) \cup I]$. The isometric embedding of A into (M', g') may then be performed via $X(S)$ using ρ' and ρ^{-1}. That is, $\rho'^{-1} : \rho(A) \to D(S')$ where $A \subset D(S)$ and $D(S') \subset M'$.

Thus we are done if $I \neq \emptyset$. But in fact that follows from our earlier observations. In particular, since ρ and ρ' are isometries, $\rho(D(S))$ and $\rho(D(S'))$ are both open. Both strictly contain $\rho(J^-(q))$ and $\rho(J^-(q'))$, which are both closed, and the conclusion follows. □

We note that the assumption that $J^-(q) \subseteq D(S)$ is key and guarantees that $D(S)$ is in a sense large enough. To see its meaning, consider the following example.

Example 2. Take two-dimensional Minkowski spacetime, and remove a spacelike slit of finite extent at the same time co-ordinate, say $t = t_1$. There is now a disconnected acausal edgless set S_{t_1} given by the level set of that same time co-ordinate function. Consider some connected acausal edgless set $S_{t_0<t_1}$ given by the level set of some time function to the past of the slit. Now consider an observer's worldline passing near the edge of the slit. Consider the past lightcone of that observer. In moving along the worldline, the intersection of the observer's past lightcone with $S_{t_0<t_1}$ grows continuously up until some time $t_2 > t_1$ is reached beyond the slit. The amount of information about $S_{t_0<t_1}$ ceases to grow smoothly as the observer moves up their worldline as the observer stops having access to some past configuation. The assumption that $J^-(q) \subseteq D(S)$ is basically a way of ruling out such circumstances, at least for some sufficiently small neighborhood of q.

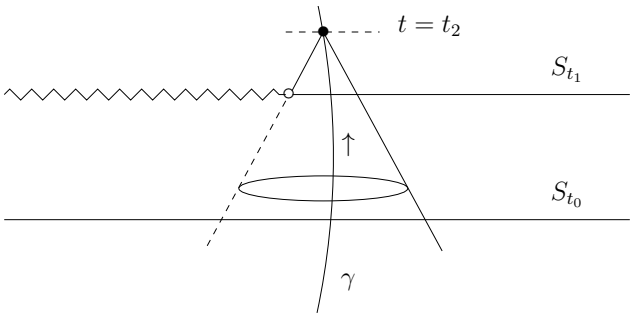


Figure III.6. This spacetime is described in Example 2. The point is that the observer's past knowledge of S_{t_0} doesn't grow smoothly with time. At $t = t_2$ and

beyond, the time spent on γ ceases to bring new information about S_{t_0}. Requiring $J^-(q) \subset D(S)$ in the above proposition blocks this kind scenario at q. It may happen further in the future, but at least not for some sufficiently small time beyond t_2.

Whether the assumptions of the above proposition are justified is of course up for debate. What's worth considering is whether the justification that can be given mirrors what can be said in the usual context of inductive reasoning. Note that what fails in example 2 is the equivalence between knowing your past at some time along with a little more than what your past affords, and knowing your new and bigger past after some elapsed time. If this equivalence holds, then in a sense there was nothing particularly abrupt that occurred in that interval. Speaking roughly, we'd say that one's knowledge of the past progresses as it has been doing up to that point.

Consider for a moment Russell's chickens, as described in chapter 6 of [30]. We'd expect two chickens that have had the same life up to t to think about the immediate future in much the same way. If at some later time $t+\delta t$ we're told that one of them has experienced nothing different from the usual, then it might seem reasonable to expect the same of the other chicken. This is essentially what is captured in the proposition above. In going from t to $t+\delta t$ with δt sufficiently small, observers make the same predictions, and that this happens is basically engineered by the assumption that $J^-(q) \subset D(S)$ at q.

The moral here is that various local and global properties can restore the observer predictive abilities, which seemed to have been jeapordized by Manchak's results. It is therefore natural to consider whether those properties are things that an observer could in principle know of their own spacetime. The epistemically ambitious observer would certainly hope so.

Unfortunately for such an observer, a number of results by Manchak and Malament show that there are obstructions to inferring such things merely on the basis of past lightcones. We begin with the definition of Malament [18].

Definition (Malament [18])**.** *A spacetime (M, g) is observationally indistinguishable from (M', g') just in case there is an isometric embedding $\phi : I^-(p) \to M'$ for every point $p \in M$; and, likewise, with M and M' interchanged.*

Note that the relation is transitive and symmetric, and moreover that since we're considering every possible point in (M, g), we're really thinking of an epistemic community rather than a single observer.

The idea of the definition is that if (M, g) is observationally indistinguishable from (M', g'), then an observer in (M, g) at $p \in M$ cannot in principle know whether they are at $p \in M$ or whether they are at $p' \in M'$, and vice versa. Malament explains the suitability of this definition over others but we avoid repeating those arguments. Malament also considers a weaker, asymmetric version of this definition which drops 'with M and M' interchanged'.

With the above definition, one can ask whether various global properties of interest are preserved under observationally indistinguishability, i.e., does (M', g') inherit property P if (M, g) is observationally indistinguishable from (M', g') and (M, g) has P? One result in this direction is that if an observer inhabits a universe with a Cauchy surface, then that spacetime may be observationally indistinguishable from one that isn't globally hyperbolic. Thus, Malament says that an observer living in a globally hyperbolic spacetime doesn't have the means of checking so.

One of Malament's most important results is the construction of two spacetimes, each built by gluing together a series of de Sitter past lightcones, which have the property of being (symmetrically) observationally indistinguishable yet non-isometric. Since his result relies on the specific causal geometry of the lightcones in the de Sitter spacetime, one wonders whether the result holds in virtue of some specific property of de Sitter.

Remarkably, Manchak [22] showed that this didn't seem to be the case for the asymmetric version of the definition. Manchak proved that the asymmetric sense of epistemic underdetermination actually occurs for all non-causally bizarre space-

times.[24]

Proposition (Manchak [22]). *Let (M, g) be any non-causally bizarre spacetime. Then there is a spacetime (M', g') such that (M, g) is observationally indistinguishable (in the asymmetric sense) from (M', g') but with (M, g) non-isometric to (M', g'). Moreover, the result holds if (M, g) and (M', g') are assumed to satisfy any number of local conditions, such as the Einstein equations, and also under the global assumptions of stable causality and inextendibility.*

To see the meaning of the result, we consider its proof. The construction is best envisioned with the diagrams offered in [22].

First, we imagine a countably infinite collection of points $p_i \in M$ such that $\cup_i I^-(p_i)$ covers M with $i \in \mathbb{N}$. Such a covering collection is always possible in any non-totally vicious spacetime. Now imagine two countably infinite collections of locally isometric copies of (M, g). Index these by $(M_{(i,\alpha)}, g)$ and $(M_{(i,\beta)}, g)$.

Each manifold in the collection $\{M_{(i,\alpha)}\}$ is constructed as follows. Given any open set $I^-(p_i)$, there is always a slightly bigger open set $V \supset I^-(p_i)$ extending beyond $I^-(p_i)$. In $V \backslash \overline{I^-(p_i)}$, one can consider a non-empty open set O_{p_i}, where the subscript refers to the fact that we're considering an extension beyond $I^-(p_i)$. This open set necessarily contains two non-intersecting closed spacelike hypersurfaces (i.e., with edge), which we label K_i^+ and K_i^-. For $i \neq 1$, the manifold $M_{(i,\alpha)}$ is $M - K_i^+ \cup K_i^-$, and the spacetime based on $M_{(i,\alpha)}$ just inherits g. For $i = 1$ we construct $M_{(1,\alpha)}$ just by removing one such hypersurface, i.e., $M - K_1^+ = M_{(1,\alpha)}$.

As for $M_{(i,\beta)}$, the manifold is $M - K_i^+ \cup K_{i+1}^-$. Note that by construction, we may pick K_i^+ and K_{i+1}^- to be non-intersecting. The spacetime based on $M_{(i,\beta)}$), as before, inherits g.

We now have two countably infinite families of spacetime based on manifolds $M_{(i,\alpha)}$ and $M_{(i,\beta)}$. We now glue them in the following order $M_{(1,\alpha)}$ - $M_{(1,\beta)}$ - $M_{(2,\alpha)}$ - $M_{(2,\beta)}$ - $M_{(3,\alpha)}$ - ... The gluing is done in a standard way: by identifying the top

[24] (M, g) is causally bizarre iff there is a point $p \in M$ such that $M \subset I^-(p)$.

side of K_i^+ in $M_{(i,\alpha)}$ with the bottom side of K_i^+ in $M_{(i,\beta)}$ and vice versa, and by identifying the bottom side of K_i^- in $M_{(i,\beta)}$ with the top side of K_i^- in $M_{(i+1,\alpha)}$ and vice versa. This identification forcibly removes the edge points from $K_i^{+(-)}$.

There are two important facts about this argument. The first is that Manchak's construction of (M', g') involves disturbing the global properties of (M, g), and perhaps more importantly, that the asymmetric definition of observational indistinguishability and Manchak's instantiation of it could actually enable observers in (M', g') to measure that they are not in (M, g) (though not the other way around of course).

To see for instance that (M', g') will in general be holed, let us recall Minguzzi's definition of hole-freeness.

Definition (Minguzzi [24]). *A spacetime (M', g') is hole-free iff it isn't holed, and it is holed if there is a future or past hole. A future hole exists if there is a partial Cauchy hypersurface $S' \subset M'$ and an isometric embedding $\phi : \tilde{D}(S') \hookrightarrow M$ such that $\phi(S')$ is acausal in (M, g) where (M, g) is any spacetime, and $\phi(H^+(S')) \cap D^+(\phi(S)) \neq \emptyset$. A past hole is defined dually.*

Note here that $\tilde{D}$ is like D but with the restriction that the curves be timelike, rather than causal. With this definition, Minguzzi shows the following.

Proposition (Minguzzi [24]). *A causally simple inextendible spacetime is hole free.*

With these, we now have the following.[25]

Observation. *Let (M, g) be a globally hyperbolic spacetime. Then (M', g'), constructed as above, is holed and causally discontinuous.*

To see this is straightforward (which is why I've merely called this an 'observation', as opposed to a 'proposition'), but we shall still go through it. Let us label

[25]Note that we assume global hyperbolicity of (M, g) to speed up the argument, but that this is likely to be replaceable by something weaker.

(M, g) and (M', g') as they appear in [22]. Let S be an acausal Cauchy surface for (M, g). Let $S_{1,\alpha}$ be the corresponding surface in $M_{(1,\alpha)} \subset M'$ and moreover choose S in M such that $K_1^+ \subset J^+(S)$. (Such a choice is always possible in virtue of the global hyperbolicity of (M, g).) As such, because of the missing points in $M_{(1,\alpha)}$ (which are also lacking in M'), we have $H^+(S_{1,\alpha}; M_{(1,\alpha)}) \neq \emptyset$. Moreover, since (M', g') is locally isometric to (M, g), and since $\tilde{D}(S_{1,\alpha}; M_{(1,\alpha)}) \subsetneq M_{(1,\alpha)} \subset M$, there exists an isometric embedding $\psi : \tilde{D}(S_{1,\alpha}; M_{(1,\alpha)}) \hookrightarrow M$. This isometric embedding gives $\psi(S_{1,\alpha}) = S$, which is acausal. Since (M, g) is globally hyperbolic, $M \subset J^-(\psi(S_{1,\alpha})) \cup D^+(\psi(S_{1,\alpha}); M)$. So far, then, we have observed that $M_{(1,\alpha)}$ is holed. Wanting to show that (M', g') is holed, we consider $H^+(S_{1,\alpha}; M')$ and the possibility of a larger isometric embedding $\phi : \tilde{D}(S_{1,\alpha}; M') \hookrightarrow M$. By construction, $\tilde{D}(S_{1,\alpha}; M')$ is larger than $\tilde{D}(S_{1,\alpha}; M_{(1,\alpha)})$, since $\tilde{D}(S_{1,\alpha}; M')$ extends into $M_{(1,\beta)}$, and, depending on how K_2^- is positioned with respect to K_1^+ in $M_{(1,\beta)}$, $\tilde{D}(S_{1,\alpha}; M')$ could also extend into $M_{(2,\alpha)}$. (Indeed this could extend into other spacetimes further along in the chain.) Since portions of $\tilde{D}(S_{1,\alpha}; M')$ which extend into $M_{(1,\beta)}$ can only be a subset of $\tilde{D}(K_1^+; M) \backslash edge(K_1^+)$ (because the boundary points of K_1^+ in both $M_{(1,\alpha)}$ and $M_{(1,\beta)}$ are absent). So $\tilde{D}(S_{1,\alpha}; M')$ doesn't cover the whole of M and thus (M', g') is holed, which by Minguzzi's result means that it isn't causally simple. A similar line of arguing easily leads to (M', g') failing to be causally continuous. There will be a failure of reflectivity along points lying on achronal null segments above and below the missing points.

The point of this observation is to say that if one starts in a well behaved spacetime, then Manchak's construction forces a major change in the global structure of the original spacetime. If however we start with a spacetime that is pathological in various respects, eg., causally discontinuous, holed, acausal, etc. - then Manchak's construction, in spite of introducing more pathologies, might not actually change the global structure of (M, g). This is interesting because it tells us that epistemic underdetermination is more problematic for pathological spacetimes. For such spacetimes however, perhaps we might have expected that to be the case, i.e.,

that past lightcones don't get you all that far in fixing the global structure of M. For well behaved spacetimes however, Manchak's (M', g') is significantly pathological in comparison.

The key point here is that observations in (M, g) and (M', g') actually differ from one another. For simplicity take (M, g) to be globally hyperbolic and inextendible. In that case, observers in (M', g') will after all be able to distinguish between (M, g) and (M', g'). Thus, observational indistinguishability only works in one direction under Manchak's construction; that is, there are experiments that a collection of causally interacting observers in (M', g') could make that differ depending on whether they inhabit (M, g) or (M', g').

Say for instance that Bob and Alice start in $M_{(i,\alpha)} \subset M'$ and that they are communicating via a reflecting lightpulse. Suppose moreover that Bob is heading towards the lower portion of K_i^+, but that Alice's trajectory remains in $M_{(i,\alpha)}$. After some finite time, Bob will find himself in $M_{(i,\beta)}$, and for no apparently clear reason Alice will suddenly fail to receive a lightpulse. Note that such a thing wouldn't happen for the same trajectories in (M, g). Therefore, the sum total of observations by Alice and Bob differ depending upon whether they live in (M, g) or (M', g').

Now although Manchak's construction introduces holes, that doesn't imply that non-isometric observationally indistinguishable spacetimes are necessarily pathological. Indeed, Malament's original de Sitter type examples are hole-free. The difference here is that Manchak's construction applies to *any* (non-totally vicious) spacetime, whereas Malament's example is based on the global causal geometry of the de Sitter spacetime. Manchak's remarkable generalization is achieved by cutting holes and introducing causal discontinuity which observers can distinguish from the case in which no such pathologies occur.

In view of this, we might look to look for a set of sufficient conditions '...' such that "observationally indistinguishable + ... $\Rightarrow$ isometric". We shall now see how a result of this kind can be proved. Part of ... will be the requirement that there be

no null lines, i.e., that there are no inextendible achronal causal curves.[26] To prove the result we start with two lemmata.

Lemma. *Let (M, g) be a spacetime without null lines. Then given any compact region K, there exists another compact $K' \supset K$ such that if $p, q \notin K'$ and $q \in J^+(p) - I^+(p)$, then any causal curve γ connecting p to q cannot intersect K.*

Proof. Suppose otherwise for some K_0 and $K_1 \supset K_0$. Then consider a sequence of ever bigger compact sets $K_{i+1} \subset K_i'$. By assumption, each K_i will have horizmos[27] related outer points $p_i, q_i \notin K_i$, $q_i \in J^-(p_i) - I^-(p_i)$ that are connected by a causal curve, necessarily achronal, that intersects K_0. In considering larger compact sets, we make these causal curves longer in the sense of an auxiliary Riemannian metric. All of these curves intersect K_0. Taking the limit, by compactness of K_0, there is a limit point for the sequence of points lying in K_0 for each causal curve linking p_i, q_i, which moreover is in K_0. By the limit curve lemma, there passes an inextendible limit curve through this limit point. That the limit curve is achronal also follows by the limit curve lemma. Thus we have a null line. □

Lemma. *Let (M, g) be a spacetime with compact Cauchy surface that does not admit any null lines. Then (M, g) admits a point $p \in M$ such that $S \subset I^-(p)$ for some Cauchy surface S.*

Proof. Since (M, g) is globally hyperbolic, there exists a continuous global time function $T : M \to \mathbb{R}$ such that each level set is a Cauchy surface. Let $K = S_1$ be a Cauchy surface and let $K' \supset K$ be as in the lemma above. Denote $T_1, 2$ the maximum and minimum respectively of T on K'. Let S_1 be any Cauchy surface with $T < T_1$ and let S_2 denote a Cauchy surface $T = T_2$. Let $q \in \partial I^+(S_2)$, $p \in S_1$ and suppose that $p \in I^-(q)$. Since (M, g) is globally hyperbolic, $J^-(q)$ is closed and so there is a causal curve connecting p to q. It now follows from the lemma that

[26] A curve is achronal if no points on it can be related by a timelike curve. A null geodesic needn't be achronal. If it sits entirely on the boundary of a past or future lightcone then it will be, but otherwise it won't be.

[27] The future horizmos of p is defined as $J^+ \backslash I^+(p)$.

such a curve doesn't intersect S_1, which contradicts the fact that S_1 is a Cauchy surface. Consequently, there cannot exist a $p \in S_1$ such that $p \in \partial I^-(q)$. However, $I^-(q)$ is open and since $\partial I^-(q) \cap S_1 = \emptyset$, the complement of $I^-(q)$ in S_1 is also open. Since $I^-(q) \cap S_1 \neq \emptyset$, and S_1 is connected, we have $S_1 \subset I^-(q)$. □

Using these, we obtain the following.

Corollary. *Let (M, g) be observationally indistinguishable from (M', g'). Suppose that both are locally Cauchy, globally hyperbolic, and inextendible. If (M, g) has no null lines, then (M, g) is isometric to (M', g') and vice versa.*

Proof. By the second lemma, (M, g) admits a point with a Cauchy surface in its chronological past. By observational indistinguishability, that surface isometrically embeds into (M', g'). The conclusion then follows in virtue of being Cauchy, globally hyperbolic and inextendible. □

One can then consider various conditions on spacetimes that lead to the absence of null lines. If it turns out that we have good reason to suspect that physically reasonable spacetimes don't have null lines, or indeed that our own spacetime doesn't have any, then of course the above result would be strengthened. On this point, we note the following rigidity theorem associated with null lines.

Theorem (Galloway-Solis [11])**.** *Let (M, g) be a spacetime satisfying Einstein's field equations with no matter fields and with a positive cosmological constant. Suppose that (M, g) is asymptotically de-Sitter and that there is a null line γ from $\mathcal{J}^+$ to $\mathcal{J}^-$. Then (M, g) isometrically embeds into de-Sitter spacetime.*

Note that this theorem is likely to admit generalizations so as to include matter models satisfying various kinds of energy conditions.

As a consequence, we have the following.

Corollary. *Let (M, g) and (M', g') be two globally hyperbolic, inextendible spacetimes satisfying Einstein's field equations with no matter fields and with a positive*

cosmological constant. Suppose that (M, g) *is asymptotically de-Sitter but not isometric to de-Sitter. If* (M, g) *is observationally indistinguishable from* (M', g')*, then* (M, g) *is isometric to* (M', g') *and vice versa.*

So far as we believe, our universe *looks* like it will be future asymptotically de Sitter, and it is obviously not isometric to de Sitter. Unfortunately for the epistemically ambitious observer, our spacetime isn't believed to be asymptotically de Sitter towards the past. Rather it is believed that there is a big bang and an expansion phase with broadly speaking similar structure as in the FLRW spacetime. A more desirable statement might therefore be one that is based on an asymptotic past that looks like an early stage FLRW spacetime, and one that is future asymptotically de Sitter.

It may also be worth noting the following additional no-null line, ray conditions; note the time reverse version of this result where 'future' replaces 'past'.[28]

Proposition. *Let* (M, g) *be a chronological spacetime. Suppose that* (M, g) *is past null geodesically complete and that one or both of the following conditions hold:*

(i) T_{ab} *is sufficiently slowly decaying towards the past*[29]*,*

(ii) every past inextendible null geodesic encounters a past trapped surface.

Then (M, g) *admits a compact Cauchy surface, no null rays (and thus no null lines).*

We might expect (i) to be satisfied in certain cosmologies, and indeed we might think that (ii) holds on account of an early period of great expansion towards the future. On the other hand, null geodesic completeness is rarely thought to be a property of our own universe; though note Minguzzi's recent argument that null geodesic completeness is in fact an interesting possibility [23].

So what do we make of these results? In my view, they serve to illustrate, rather than to quell the doubts of the ambitious observer. For in truth, I don't know of a

[28] A null ray is a null half-line.

[29] This is made precise in [23].

truly good argument that null lines aren't to be expected, let alone global hyperbolicity, inextendibility, and local Cauchy determinism.

It is also important to note that even if more attractive results were proven, all that such these could show is that underdetermination is escapable *if* spacetimes are a certain way. Crucially however, such restrictions cannot be fully justifiable by observation. Rather, they are at best motivated by various observational and mathematical considerations. This is particularly explicit for the above 'no null line' condition, for in a globally hyperbolic spacetime, a null line cannot sit in a past lightcone.

The resulting picture is that the epistemically ambitious observer (or a community thereof) can do no more than hope that their spacetime is a certain way. It may be that certain spacetimes properties can be partially justified, but such a justification can be neither complete nor justified by means of observation alone. The best one can do is appeal to a mixture of observational and mathematical arguments, and what these might be is a point of discussion.

3. Strong Cosmic Censorship

For illustrative purposes later on, let us start by recalling some ideas of Quine [27]. Quine conceives of the problem induction as the task of understanding why induction about some predicates and not others is successful. In short, his proposal is that predicates that grip onto natural kinds are good predicates to reason inductively about. In particular, it is the concept of similarity that permits grouping things into natural kinds, and, upon further conditions, it's that which accounts for performance under induction.

Quine also argues that in better understanding some physical phenomenon, we usually become able to bypass the initial sense of similarity. He illustrates this with the example of soluble things, where here solubility guarantees membership in the natural kind of such things.

Quine. "One can redefine water-solubility by simply describing the structural conditions of that mechanism. [...] once we can legitimize a disposition term by defining the relevant similarity standard, we are apt to know the mechanism of the disposition, and so by-pass the similarity." [27]

The idea here is that in better understanding solubility — say by describing the phenomenon in terms of molecular goings-on — we become able to bypass the similarity and re-define somewhat the relevant natural kind. Quine explores this in various ways, but here we need only keep in mind his idea that similarity leads to the notion of a natural kind, and that progress will often lead to a bypassing of the original similarity standard.

Changing gears somewhat, consider the original statement of the SCC by Penrose [31].

Conjecture. *Physically reasonable spacetimes are globally hyperbolic.*

This version of the SCC was essentially borne out of the singularity theorems that were developed mainly by Geroch, Hawking, and Penrose from 1965 to 1970. These theorems are existence results for geodesic incompleteness for spacetimes satisfying conditions which, it is argued, we may reasonably expect of our own. Important as they are, these theorems don't inform us as to what kind of singularity is expected, or indeed what results of the spacetime's global structure. The SCC was envisioned partly as the idea that the singularities forming in realistic situations result in spacetimes that are globally hyperbolic, like for instance the Schwarzschild black hole solution and the FLRW class of cosmological solutions.

Phrasing Penrose in terms of Quine, the SCC seeks to establish a fact about a certain natural kind. The natural kind is the class of physically reasonable spacetimes, and the property shared among this class is their physical reasonableness. The SCC is then a Cauchy surface existence claim for each member of the class, or, more long-windedly, the claim that membership in the natural kind of physically

reasonable spacetimes implies membership in the class of globally hyperbolic spacetimes.

Over the years however, the SCC has become a claim about the inextendibility of Cauchy developments to generic initial data.

Conjecture. *Let (V, h, K) be a 'generic' initial data set with V a connected Riemannian 3-manifold. Then the maximal Cauchy development of (V, h, K) is inextendible as a 'solution' of the field equations.*

The terms 'generic' and 'solution' are highlighted because it is still somewhat open as to what they should be. As stated, although the latter conjecture lacks precise definitions, it can be made provably true or false by instantiating these terms with precise definitions. Moreover, we can expect the conjecture to be true under some definitions, but false under others. The real question, then, is whether the ones under which it is true capture some notion of physical reasonableness. For that to happen, there needs to be a coincidence between mathematical and physical senses of relevance, which can indeed happen in GR.

Take for example geodesic completeness. It is a mathematical property of natural importance, but it's also physically relevant because we think of free-falling observers and or light rays as being tracked (perhaps approximately so) by geodesics.[30] This coincidence is celebrated and singularity theorems have long been a cornerstone of GR, both in physical and mathematical circles.

In many cases however, mathematical and physical senses of relevance come apart. Take for instance questions concerning the regularity of g. Mathematically it is absolutely crucial to keep track of regularity, and indeed recently the community has become increasingly interested in low regularity spacetimes, with some authors even dealing with C^0 metrics. But despite various results and an increasing interest, it isn't at all clear how questions of regularity lead to physically relevant distinctions. Could we for instance think of an experiment which distinguishes

[30] A recent and precise expression of this idea has been offered in [14].

whether g is $C^{1,1}$ and not C^3? Not to my knowledge.

Despite the physical opacity of regularity, the concept is centrally important for the Cauchy problem and indeed for the modern SCC. For indeed, C^0-inextendibility is quite clearly much stronger than C^∞-inextendibility.[31] To see this, we note that it was only in the last couple of years that a difficult argument by Sbierski [31] was able to prove the C^0-inextendibility of the Schwarschild solution, whereas its C^∞-inextendibility has long been known.

The other key idea of the modern SCC is genericity. Before considering its meaning, a word or two ought to be said as to why it's become so important. Here, the analogy with Quine is revealing. What seems to have happened is that in obtaining a better understanding of the mathematical forces at work, we have come to bypass Penrose's loose and intuitive grip on the notion of similarity and the associated natural kind of physically reasonable spacetimes. Both have been replaced by concepts that are believed capable of bringing eventual adjudication. The mechanism is the Cauchy problem and consequently the earlier standard of similarity has morphed into a notion of genericity in the sense of initial data. What remains to be seen is whether this transformation can in any way preserve the original intuition of Penrose.

3.1 Man is Torn by the Norm

As Quine says, the notion of things being similar to one another is key in initially guiding our thinking. Moreover, it is often the case that with better understanding of some phenomenon, one's initial sense of similarity becomes grounded in some more fundamental mechanism. Quine also writes that the initial sense of similarity and the more fundamental mechanism which eventually replaces it can conflict with one another.

[31] 'C^k-inextendible' is the impossibility of extensions resulting in spacetimes with C^k metrics.

Quine. "Living as he does by both bread and basic science, man is torn. Things about his innate similarity standards that are helpful in the one sphere can be a hindrance in the other." [27] pg. 49.

Quine's bread refers to everyday worldly properties like colours, length, and so on. His science refers to things that we consider more fundamental. With that in mind, one wants to know whether the modern SCC can ground Penrose's version. Since the latter is essentially a statement about the Cauchy problem, the question becomes whether the Cauchy problem can be made physically relevant or transparent. As hinted at in Part I however, the language of the Cauchy problem is based on quite technical definitions that are difficult to find a clear physical interpretation for.

Now in truth, the modern SCC lies well beyond currently available methods. The closest statements obtained involve highly specific restrictions on the class of initial data sets that specify the topology of V, the type of initial data permitted, and the kind of possible matter models involved. These choices affect the relevant functional spaces in which the initial data is defined, which in turn the relevant norms and the associated concept of genericity. The result is that there are actually a variety of subconjectures, more or less in the spirit of the modern SCC, but each carrying its own meaning and physical interpretation.

Consider for instance the simplified case of spatially homogeneous spacetimes.[32] With such a restriction, the set of initial data can be given the structure of a finite dimensional manifold, and as a result there are various notions of genericity that become available:

(i) the complement is of measure zero with respect to the measure induced by a Riemannian metric on the manifold of initial data,

[32]Recall that spatially homogenous means that there exists a one-parameter family of spacelike hypersurfaces V_t foliating the spacetime such that for each t and for any points $p, q \in V_t$ there exists an isometry of the spacetime metric g which takes p into q.

(ii) the complement is a countable union of submanifolds of positive codimension,

(iii) the set is open and dense with respect to some natural topology,

(iv) it is a countable intersection of open and dense sets with respect to some complete topological metric.

Though in general inequivalent, all of these notions share the crucial property that if a subset is generic, then the complement of that subset isn't.[33] The notion which is eventually decided upon will, naturally, depend on what the techniques permit to prove.

It ought to be said that that in the infinite dimensional case, measure theoretic notions of genericity as above become more subtle and it can be still harder to find a physical interpretation for them.

In essence, then, the modern SCC is a wide PDE problem concerning the long-time behavior of solutions to the field equations. So what kind of long-time results has the community been able to produce?

One set of studies of the long-time behavior of solutions are so-called stability results. Such results show that under sufficiently *small* deviations of an initial data from one that would isometrically embed into some given reference spacetime, the maximal development of that initial data reproduces the large-scale properties of the reference spacetime.

Christodolou-Klainerman's [7] monumental achievement in proving their result on the global non-linear stability of Minkowski spacetime is of this kind. They consider an open set of initial data that differs in a 'small' way from one that would isometrically embed into Minkowski spacetime, and they show that the full evolution of such data sets are asymptotically flat and geodesically complete, thus preserving the important global properties of Minkowski spacetime. I've highlighted the word 'small' because their definition of it, given for instance on page 16 of [7], is highly non-trivial. In any case, their result is widely seen as a turning point in

[33]Note for instance that being 'dense' isn't enough to count as generic, eg., under the usual topology, both the rational and irrational real numbers are dense.

general relativity, and it has since been re-proved and generalized in various ways, each time yielding more insight into the problem.

Although there's is little doubt that their achievement (and those following) count as remarkable steps forward in our understanding of hyperbolic PDEs, the physical meaning of such results remains to be more fully understood. To avoid what would be a sizeable digression in re-stating their assumptions, one need only consult [7] to feel the opacity of their definition of suitably small deviations from a Minkowski slice.

A simple observation one can make is as follows. As is easily seen, there are black hole solutions with arbitrarily small ADM mass. Given that Minkowski spacetime is the only spacetime with zero ADM mass, such arbitrarily small mass black hole solutions are in one sense arbitrarily close to Minkowski, despite of course being radically different from Minkowski spacetime in the large.[34] As such, their assumption that the data deviates from a Minkowski slice must somehow grip on the possibility of forming black holes. As such, perhaps it is to be understood as a bound on the concentration of curvature that is sufficient to avoid dynamically forming trapped surfaces. To my knowledge however, I am not aware of any results that clarify this point, or more generally the physical meaning of their small data assumption.

More pragmatically, one can simply ask whether it is in principle possible to check whether some spacetime region does indeed involve a 'small' deviation, in their sense of the term, from a slice that embeds into Minkowski spacetime. Results in this direction would be of interest, but at present they seem difficult to formulate let alone obtain.

[34]The ADM mass is a global quantity measuring the total energy associated with an asymptotically flat spacetime. If the dominant energy condition is satisfied, it is zero if and only if the spacetime is Minkowski.

3.2 Reason for Hope?

The remarks above suggest a certain degree of pessimism as to the physical meaning of the modern SCC, even if some mathematically profound version of it turns out to be true. I would now like to offer some thoughts of optimism.

In the last couple of years, there has been remarakble rise in interest of the stability of black hole solutions. Since a review of what has been discovered in the last decade would involve far too long a digression, I shall merely mention that a number of results have been proved for specific perturbations of Schwarzschild black holes, for certain kinds of wave equations and matter-models on fixed black hole backgrounds, and in some more generality for the reference black hole spacetime of Kerr de Sitter. As has long been known, the stability of certain black hole solutions is actually non-trivially related to questions of SCC. Remarkably, in recent years, this link has been significantly more precise, and a number of striking results have been obtained.

Black hole solutions (other than the Schwarzschild solution) have Cauchy horizons in their interiors, beyond which the spacetime can be extended, which results in a non-globally hyperbolic spacetime. As such, if it can be shown that such Cauchy horizons are stable under perturbations of black holes solutions, then the modern SCC would seem to be put in serious doubt. Conversely, if it turns out that Cauchy horizons are unstable and that they generically disappear, then one has evidence in favor of the SCC.

Recently, Dafermos and Luk [8] have found precise relations between the stability of the spacetime exterior to the Kerr black hole and the stability of the Cauchy horizon which lies in its interior. In particular, they show that if the Kerr exterior is strongly stable in a certain precise sense (which is widely believed to be the case), then in fact its interior Cauchy horizon will be C^0 stable; that is, C^0 extensions beyond the horizon will exist. As a consequence of their work, if the Kerr solution is strongly stable, then the C^0 version of SCC is in fact false, though of course a weaker form of the SCC version could still be true.

Also very recently, some numerical studies (including but not limited to [5], [6]) have found evidence of violations to the SCC in certain settings of highly charged Kerr de Sitter black holes. The mechanism in these studies is similar to that above in that it involves linking the decay of perturbations to the black hole's exterior to the stability of its interior Cauchy horizon. It remains to be seen whether the result is true of black holes that are not highly charged Kerr de Sitter black holes. With regards to the astrophysical relevance of such results, whether such black holes actually exist is of course a subject of debate.

What is exciting here is that even if a master theorem like what was suggested above is currently far beyond reach, the link between black hole stability and Cauchy horizon stability has produced a new, concrete, and even analytically accessible mechanism by which to test the SCC.

3.3 In Sum

Penrose's version of the SCC was intended to be conceptually clear and of physical relevance. If somehow it could be shown to be true, then we could envisage appealing to it as a means of solving a great deal of problems which otherwise occur in spacetimes that are not globally hyperbolic. Unfortunately, without any definition of 'physically reasonable', the original SCC as stated was void of mathematical content.

Mathematical relativists inspired by Penrose sought to propose a more precise and potentially provably true or false version of the SCC. The result of that is now a collection of global existence and uniqueness claims in the theory of hyperbolic PDEs on various kinds of Lorentzian manifolds.

Under some definitions that are suited to the analysis of certain highly specific cases, some of these subconjectures have been shown to be true; and under other definitions that are stronger than the usual version, certain statements in the spirit of the SCC have been shown likely to be false, both by numerical and purely ana-

lytical arguments.

For each subconjecture in the spirit of the SCC, in making the statement precise enough to be provably true or false, one considers a certain notion of genericity defined with respect to a certain space and norm for the space of initial data, and a certain regularity for the possible extensions beyond Cauchy horizons. Both of these components — genericity and regularity — are difficult to interpret physically. This is because the Cauchy problem involves technical definitions that are more difficult to interpret than the more elementary concepts of Lorentzian geometry — lightcones, geodesics, slices, etc. As such, whatever the truth value of these various subconjectures, it isn't at all clear what their physical relevance would be. This leads to at least two questions.

1. Is it possible to come up with a mathematically precise definition of physical reasonability that doesn't trivialize the SCC?

2. Can we make more physically transparent the definitions and results described in today's studies of the Cauchy problem?

If the answer to the first is 'no', then Penrose's version of the SCC will be dissolved. If there is no progress on the second, then the modern SCC will remain more of a mathematical pursuit than a physically interesting one.

At least until discussions about the SCC became dominated by studies of the Cauchy problem, GR had been rather exemplary in its ability to yield clear and physically satisfying descriptions. It may be that the SCC is pushing GR in a realm where it is unable to yield a clear physical picture. Perhaps in that case we ought to adjust our standards somewhat. For now at least, it is probably wisest to think of the SCC as a difficult and sprawling program in a certain branch of hyperbolic PDE theory.

References

[1] Budic, R., Sachs, R.K., Deterministic space-times. Gen Relat Gravit (1976) 7: 21.

[2] Choquet-Bruhat, Y., Geroch, R., Commun. Math. Phys. (1969) 14: 329.

[3] Cleve, J., Reliability, Justification, and the Problem of Induction, Midwest Studies In Philosophy: 555–567, (1984).

[4] Earman, J., Glymour, C., Stachel, J. (eds.) Foundations of Space-Time Theories. Minnesota Studies in the Philosophy of Science, vol. 8. University of Minnesota Press, Minneapolis (1977).

[5] Cardoso, V., Costa, J. L., Destounis, K., Hintz, P., Jansen, A., Quasinormal modes and Strong Cosmic Censorship, Phys. Rev. Lett. 120, 031103 (2018).

[6] Cardoso, V., Costa, J. L., Destounis, K., Hintz, P., Jansen, A., Strong cosmic censorship in charged black-hole spacetimes: still subtle, Phys. Rev. D 98, 104007 (2018).

[7] Christodoulou, D., Klainerman, S., The Global Nonlinear Stability of the Minkowski Space. Princeton Mathematical Series, 41. Princeton University Press, 1993.

[8] Dafermos, M., Luk, J., (2017) The interior of dynamical vacuum black holes I: The C^0-stability of the Kerr Cauchy horizon, arXiv:1710.01722 [gr-qc].

[9] Earman, J., A primer on determinism, The Western Ontario Series in Philosophy of Science, Vol 32, Springer, (1986).

[10] Earman, J., Glymour, C., Stachel, J. (eds.) Foundations of Space-Time Theories. Minnesota Studies in the Philosophy of Science, vol. 8. University of Minnesota Press, Minneapolis (1977).

[11] Galloway, G. J. G., Solis, D., Uniqueness of de Sitter space, Class.Quant.Grav.24:3125-3138, 2007.

[12] Geroch, R., (1970) Singularities. In Relativity, ed. M. Carmeli, S. I. Fickler, and L. Witten, 259-91. New York: Plenum.

[13] Geroch, R., Prediction in General Relativity. In: Earman, J., Glymour, C., Stachel, J. (eds.) Foundations of Space-Time Theories. Minnesota Studies in the Philosophy of Science, vol. 8. University of Minnesota Press, Minneapolis (1977).

[14] Geroch, R., Weatherall, J.O., Commun. Math. Phys. (2018) 364: 607.

[15] Goodman, N., Fact, Fiction, and Forecast. Harvard University Press 1983.

[16] Hempel, C., A Purely Syntactical Definition of Confirmation, Journal of Symbolic Logic, Volume 8, Issue 4 (1943), 122-143.

[17] Jackson, F., Grue. Journal of Philosophy 72 (5):113-131 (1975).

[18] Malament, D., (1977) Observationally Indistinguishable Space-Times. In Foundations of SpaceTime Theories, ed. J. Earman, C. Glymour, and J. Stachel, 61-80. Minnesota Studies in the Philosophy of Science 7. Minneapolis: University of Minnesota Press.

[19] Manchak, J., (2008). Is prediction possible in general relativity? Foundations of Physics, 38(4): 317-21.

[20] Manchak, J. (2009). Can We Know the Global Structure of Spacetime? Studies in History and Philosophy of Modern Physics 40:53–56.

[21] Manchak, J., On the Inextendibility of Space-Time. Philosophy of Science, 84 (December 2017) pp. 1215–1225.

[22] Mellor, D. H., The Warrant of Induction, Chapter 15, Matters of Metaphysics, Cambridge: Cambridge University Press (1988).

[23] Minguzzi, E., Chronological spacetimes without lightlike lines are stably causal, Commun.Math.Phys.288:801-819,2009.

[24] Minguzzi, E., Causally simple inextendible spacetimes are hole free, Journal of Mathematical Physics 53 (2012) 062501.

[25] Minguzzi, E., Sanchez, M., The Causal Hierarchy of Spacetimes, in H. Baum and D. Alekseevsky (eds.), vol. Recent developments in pseudo-Riemannian geometry, ESI Lect. Math. Phys., (Eur. Math. Soc. Publ. House, Zurich, 2008), p. 299.

[26] Penrose, R., (1979) Singularities and Time-Asymmery. In General Relativity: An Einstein Centenary Survey, ed. S. Hawking and W. Israel, 581–638. Cambridge: Cambridge University Press.

[27] Quine, W.v.O., (1970). Natural Kinds. In Nicholas Rescher; et al. Essays in Honor of Carl G. Hempel. Dordrecht: D. Reidel. pp. 41–56. Reprinted in: Quine (1969), Ontological Relativity and Other Essays, Ch. 5, Columbia Univ. Press.

[28] Ringstromg, H., The Cauchy Problem in General Relativity. European Mathematical Society, Zurich 2009.

[29] Ringstrom, H., On the Topology and Future Stability of the Universe, Oxford University Press 2013.

[30] Russell, B., The problems of philosophy, 1912.

[31] Sbierski, J., The C^0 inextendibility of the Schwarzschild spacetimes and the spacelike diameter in Lorentzian geometry, J.Diff.Geom. 108 (2018) no.2, 319-378.

[32] Wald, R., General Relativity, University of Chicago Press, 1984.

Appendix A. Jackson's Argument

Note below that 'examined' means 'examined prior some time T'.

For Jackson, what we're really after is the following modified version of the *straight rule* (SR) — the practice by which one extrapolates that all Fs are G on the basis of observing some Fs to be G.

Jackson. "When we apply the SR in practice, we commonly argue on the modified pattern: certain Fs which are H being G supports Fs which are not H being G ('H' often being 'examined', 'sampled', etc.)." [17].

For Jackson, discussions surrounding Goodman's example havn't appreciated the importance of the counterfactual 'the Fs would not have been G if they had not been H'. He grants this counterfactual its own name.

Jackson. "The condition: that certain Fs which are H being G does not support other Fs which are not H being G if it is known that the Fs in the evidence class would not have been G if they had not been H, will be referred to as the counterfactual condition." [17].

For Goodman's example, upon substituting the relevant terms, Jackson is saying that the examined emeralds being grue does not support unexamined emeralds being grue, if it is known that the examined emeralds would not have been grue if they had not been examined. The point being that we know, in our world, that examined emeralds would not have been grue if they had not been examined (for otherwise they would have been blue). Thus we are not warranted in going from 'all examined emeralds are grue' to 'all emeralds, examined or unexamined, are grue'. Jackson considers the following further example.

Jackson. "Every lobster I have observed has been red. This supports that the next lobster I observe will be red, and no doubt it will be. But every lobster I have observed has been cooked, and I know that it is the cooking that makes them red - that is, that the lobsters I have observed would not have been red if they had not been cooked. Hence I do not regard myself as having good evidence that the next uncooked lobster I observe will be red." [17].

For Jackon, it is by not appreciating such counterfactuals that we go wrong in going from 'all examined emeralds are grue' to 'all emeralds, examined or unexamined, are grue'. The conclusion of Jackon is that there is nothing truly new or problematic about the new riddle so long as we have a good understanding of the relevant counterfactuals at work.

Jackson "When we argue on this modified pattern, we take it that it is not the case that the Fs which are H would not have been G if they had not been H. This guarantees that we can never be led from the same evidence to opposite predictions concerning whether the non-H Fs are G." [17].

For Jackson, we know in our world that the following is false: if the emerald had not been examined then it would not have been green.

Admittedly, deferring to counterfactuals looks like opening a philosophical can of worms. Still, Jackson's argument does have some appeal to it and thus one wonders how it fares in the spacetime context. In that context, we seek to understand whether the following counterfactual is true: if an examined past lightcone had not been examined, then it would not have been globally hyperbolic. Unfortunately, we don't know the truth value of such a counterfactual. Indeed, its truth value depends on the global properties of (M, g), which, as I've described, which we don't readily have access to on the basis of $I^-(p)$. Thus, Jackson's way out of the new riddle by means of counterfactuals doesn't seem to work so straightforwardly in spacetime.

Of course that's probably to be expected: we understand emeralds better than the global causal of our spacetime. We shall consider how to assess the truth value of such counterfactuals in IV, and there we shall see how the strong cosmic censorship conjecture could potentially play a role.

IV. Counterfactuals and Laws for Spacetimes

1 Introduction

Lewis [16] sets out truth conditions for counterfactual statements like 'given a world i, if it were that ϕ, then it would be that ψ'. In short, these conditions say that a counterfactual is true at world i iff there is a collection of nearby possible worlds such that in each world if ϕ is true, then ψ is also.

Standing in the way of concrete adjudications are vagueness and context dependence. If there is vagueness in ϕ, ψ, there will also be in the counterfactual. As for context dependence, there may be ambiguities as to what constitutes the collection of nearby possible worlds. In view of these obstructions, we shouldn't generally expect to have the ability of adjudicating beyond doubt whether a counterfactual is true. That doesn't matter however. The point of Lewis' account is to describe what it means for a counterfactual to be true, or indeed of how it may earn its truth value.

My main aim will be to explore whether and how Lewis' account can be used to assign truth values to counterfactuals that we can formulate about spacetimes. And by this I mean statements of the form 'given a spacetime (M, g), if it were that ϕ, then it would be that ψ', where ϕ and ψ both precisely defined using the differential geometric framework of spacetime theories. For example, given a reference spacetime (M, g), I shall consider statements of the following form.

- 'If there had been a trapped surface, then the spacetime would have been geodesically incomplete.'
- 'If g had been $\tilde{g}$ with $\tilde{g} - g$ everywhere sufficiently small, then $(M, \tilde{g})$ would be without closed causal curves.'
- 'If within some compact set $K \subset M$, g had been $\tilde{g}$, then flat regions of $(M, \tilde{g})$ wouldn't be compact.'

Since it isn't a priori clear how Lewis' account could be used for such counterfactuals, my purpose will be to propose a way of doing so.

The interest in studying spacetime counterfactuals comes from the important

role they occupy in our standard talk and thought about spacetimes. As for my choice to consider Lewis' account, it is mainly due to the influence that it has had on philosophers and the appealing challenge of interpreting some of his ideas in the context of a successful physical theory.

In virtue of considering Lewis' account, I shall naturally be drawn to his views on laws of Nature, which in the current context will be the laws of spacetime theories and general relativity (GR) in particular. To see part of the connection between counterfacfuals and laws, I shall briefly recall a recent exchange between Carroll [6] and Beebee [2], which neatly illustrates the link.

To simplify and widen the scope of my analysis, I shall only focus on those broader aspects of Lewis' views. As such, I will not explore the various brands of his Humean Supervenience view, and I will also avoid his modal realism and the ontological status of possible worlds; a summary of the Humean Supervenience view along with various references can be found in [20] [21].[35]

Of the broader aspects of Lewis' view, I will pay particular attention to the maxims he proposes for assessing world similarity [18]. Lewis proposed these in response to an example raised by Kit Fine [9] which, despite being intuitively true, seems to be ruled false under Lewis' original account. Lewis introduces these maxims to rectify this glitch and to make better overall sense of his original account.

Before going any further, given that I shall be considering Lewisian ideas in the context of a physical theory, I should like to make a remark as to my overall purpose. When Lewisian views are considered in the context of a physical theory, it is often with the intention of assessing their physical co-tenability. In particular, a number of authors have taken issue with his Humean Supervenience view on the

[35]Lewis describes the view as follows. "It says that in a world like ours, the fundamental relations are exactly the spatiotemporal relations: distance relations, both spacelike and timelike, and perhaps also occupancy relations between point-sized things and spacetime points. And it says that in a world like ours, the fundamental properties are local qualities: perfectly natural intrinsic properties of points, or of point-sized occupants of points. Therefore it says that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past and present and future." [19] pg. 474.

basis of the apparent contradiction with quantum mechanics. The apparent tension between his view and quantum mechanics stems from the fact that Lewis seems to put great emphasis on the fundamentality and separability of local, point sized properties, whereas quantum entanglement seems to imply a form of non-locality and non-separability going in exactly the opposite direction. Lewis was indeed aware of the apparent tension.

Lewis. "Really, what I uphold is not so much the truth of Humean Supervenience as the tenability of it. If physics itself were to teach me that it is false, I wouldn't grieve. That might happen: maybe the lesson of Bell's Theorem is exactly that ... But I am not ready to take lessons in ontology from quantum physics as it now is. ... If, after [quantum theory has been cleaned up], it still teaches non-locality, I shall submit willingly to the best of authority. What I want to fight are philosophical arguments against Humean Supervenience." [18] .

Note here that Lewis suggests the possibility that physics could rule negatively on his ontology. My position here will not be of this kind. I won't attempt to argue either positively or negatively in this direction. My purpose is merely to see how far one can go in providing an interpretation of his views in the context of a successful physical theory. Consequently, I won't regard physics as having the final word on certain aspects of his views, and I certainly won't seek to mould spacetime theories or GR so as to better accommodate them. That I'm able to do this is owed to the fact that my focus is on the broader aspects of his views, and in particular the ones which don't involve specific ontological claims. To say it in Lewis' own terms, if he says that he isn't "concerned with physics proper, but with the preliminary metaphysics done by physicists" p. 91 [16], then I too am concerned with that metaphysics, but with the difference being that I shall focus first and foremost on what the physicists actually do.

1.1 Law Governance and Counterfactuals

There are two particular popular accounts of physical laws, and at a broad level these can be grouped under the labels 'Humean' and 'anti-Humean'. Speaking roughly, the former sees laws as supervening over particular non-modal matters of fact, whereas the latter sees them as governing over them; and by 'particular non-modal matters of fact', I mean facts about objects and their natural, non-nomic properties and relations, such as for example the fact that my worldline has some specific geometrical property.

In Lewis' version of the Humean view, the idea is that laws are those statements that appear as axioms and or theorems of the most informative yet streamlined axiomatization of the total set of such facts. Lewis gives us the following definition.

Lewis. "A contingent generalization is a *law of nature* iff it appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength." [16] p. 73.

The picture in mind is that any given world can be broken down into its own set of facts, and moreover that this set is somehow amenable to axiomatization. An extreme form of such an axiomatization could be the list of all the relevant set of facts that are true of that world. Such a system is very strong in the sense of being highly informative, but it is also highly non-simple and of course problematically unwieldy. On the other end of the spectrum, one could have a very simple axiomatization, comprised of only a couple of axioms, but which is consequently very weak in that it doesn't permit recuperating anything close to the total set of facts. In the Humean view, one imagines axiomatizations that strike a balance between these two extremes. These axiomitizations achieve a certain degree of strength and simplicity. They capture a good deal of the total set of facts, but without being anything like a piecewise specification of them. Given such a system, its axioms and theorems are then what we call the laws of that world.

From now on, I shall use the term 'Humean' to refer to the above.

Since I won't consider the anti-Humean view in any real detail, my description of the view will be kept brief; see [2] for a gentle introduction, [1], [6] for more detailed accounts, and [27] for a physics-based rendition of the view.

The basic idea of the anti-Humean view (of which there are various versions) is that physical laws *govern* over particular matters of fact. Namely, laws of Nature exist over and above the total set of facts, and moreover the facts are what they are in virtue of what the laws fix them be. The sense of this governance is understood differently among anti-Humeans. For instance, Maudlin [27] views it as primitive and irreducible, whereas Armstrong [1] describes it in terms of various components that are interwoven together into a rich metaphysical structure.

Naturally, there are debates as to which view is superior. One kind of argument against the Humean view involves the analysis of specific examples that are intended to show that physical laws don't in fact supervene on particular matters of fact. The mirror example by Carroll is a prime example [6]. Carroll's analysis can be cast as a consequence of the following modal principles, which he proposes as intuitively convincing.

Carroll's Modal Principles. Let $\diamondsuit$ denote "physically possible" and $\Box$ denote "physically necessary", and consider 2 and 3 as following from 1.

1. If $\diamondsuit P$ and $\Box(P \supset Q)$, then if P were the case, Q would (still) the case.
2. If $\diamondsuit P$ and Q is a law, then if P were the case, Q would (still) be a law.
3. If $\diamondsuit P$ and Q is not a law, then if P were the case, Q would (still) not be a law.

A few years after [6], Beebee responded to Carroll by pointing out that his modal principles are at odds with the Humean view, and, as such, that it is no surprise that his analysis of the mirror example leads him to the conclusion that laws don't supervene. More precisely, she shows that Carroll's modal principles lead to a view

on counterfactuals which is totally unlike the typical Humean view, and by 'typical' she refers to Lewis' account as described in [16]. She shows this by means of the following example.

Note below that Beebee understands the term 'physically possible' in the sense of Lewis: a world j is physically possible with respect to world i if and only if the laws of i are all true in j.

Beebee's Example Let i be a deterministic world. As a consequence of determinism, worlds that are physically possible relative to i will in general be quite different with respect to their particular matters of fact.

Now let $P \equiv$ 'Natalie drops her pen at t', and suppose that in i, Natalie actually hangs on to her pen at t, i.e., so that P is false at i. Naturally, we expect there to be physically possible worlds where P is true. Let j be the closest such world. Of course, since j is physically possible (and thus deterministic), it must differ from i in some respects for all times before t, and so as aforementioned, j is likely to be pretty distant from i with respect to its particular matters of facts.

Now consider the closest but not necessarily physically possible world where P is true. Let this world be k. On Lewis' analysis, k is identical to i with respect to its particular matters of fact until just before t, and then, owing to a small miracle in k occurring just before t, Natalie ends up dropping her pen. Let us denote by Q the law of i that's broken at k.

Relative to i, P is physically possible and Q is a law, and therefore the antecedent of 2 is true. However, the consequent is now false for it's false that if P had the been the case, then Q would still be a law. As such, 2 is false because the closest world where P is true isn't necessarily a physically possible one.

We see here that Carroll's modal principles are inconsistent with Lewis' account of counterfactuals and laws as described in [16]. Note moreover that what is key for Beebee here is that, given some fixed world, the collection of physically

possible worlds needn't coincide with that of the closest possible worlds. It is that which ultimately leads to a contradiction with Carroll's modal principles.

1.2 Spacetime and Humeans

In considering jointly the terms 'Humean' and 'spacetime', we ought to consider whether and how other authors have done so. To my knowledge, the contributions that are most relevant to what follows are due, respectively, to Earman and Callender.

In his treatise [8] on general relativity, Earman makes use of the Humean view to study some of the conceptual puzzles surrounding the possibility of closed timelike curves (CTCs). Earman studies certain modern versions of the grandfather paradox and considers the status of those consistency constraints required for CTCs to be logically permitted. He investigates whether these constraints should be thought of as laws, and, if so, whether they may be explained by physical arguments themselves expressible in GR. The relevance of the current study is that Earman explicitly endorses a Humean account of laws, and moreover that he explicitly states and rejects Carroll's modal principles.

More recently, aiming to understand what it is about Time that distinguishes it from Space, Callender [5] proposes that part of the answer lies in the "best systems" account of laws. For Callender, Time is special because it is the direction in which Physics tells its best, simplest, and most informative stories. Among his various illustrations of this point, one that is strongly inspired by GR is his suggestion to consider the Cauchy problem "sideways" — that is, in the sense of developing a spacetime from a timelike hypersurface, rather than from the usual spacelike hypersurface. On the basis of the PDEs describing the dynamics of physical fields, he points out that the sideways problem is always harder to extract mathematical and physical content from. Thus, he argues, the PDEs single out Time as the best story-telling direction. Like Earman, the idea here is that the Humean view permits for a greater understanding of the various questions that we can ask about

spacetimes.

In view of these contributions, it becomes of interest to look for a consistent Humean account of counterfactuals and laws in the context of spacetimes.

2. Lewis' counterfactual semantics

Let us first recall the main aim of Lewis as stated in [16].

Lewis' Aim. "Counterfactuals are, for the most part, contingent. We have been trying to give their truth conditions in general: to say how the truth value of a counterfactual at *any* world depends on the truth values at various worlds of its antecedent and consequent." [16] p.61

Lewis' account is designed to deal with counterfactuals that are essentially unrestricted in scope. Since here our focus is on spacetimes, it is only a small portion of his overall account that we shall need to consider.

Lewis starts by formulating the relevant logical operators of necessity and possibility. Necessity counterfactuals are of the form described above, whereas possibility counterfactuals are of the kind 'given a world i, if it were that X, then it *might* be that Y'. To simplify the current discussion, I won't consider such might-be counterfactuals. A necessity operator, in general, is an operator that acts like a restricted universal quantifier over possible worlds. Necessity of a certain sort is truth at all possible worlds that satisfy a certain restriction. With a view to talk about spacetimes, we're interested in *physical necessity*, and by that, given a reference spacetime, we mean truth at all worlds where those laws hold true.

Lewis then describes his centrally important notion of spheres of possible worlds. Lewis' spheres are nested collections of possible worlds that capture a qualitative and comparative notion of similarity of between worlds. The smaller the sphere, the more similar its members relative to i. Lewis also demands his system of spheres

to satisfy the following.

Spheres of Worlds. Let $\mathcal{S}$ be an assignment to each possible world i of a set $\mathcal{S}_i$ of sets of possible worlds. Then $\mathcal{S}$ is called a centred system of spheres, and the members of each $\mathcal{S}_i$ are called *spheres* around i, iff, for each world i, the following conditions hold.

(i) $\mathcal{S}_i$ is centred on i; that is, the set $\{i\}$ having i as its only member belongs to $\mathcal{S}_i$.

(ii) $\mathcal{S}_i$ is nested; that is whenever S and T belong to $\mathcal{S}_i$, either T is included in S or T is included in S.

(iii) $\mathcal{S}_i$ is closed under unions; that is, whenever C is a subset of $\mathcal{S}_i$ and $\cup C$ is the set of all worlds j such that j belongs to some member of C, $\cup C$ belongs to $\mathcal{S}_i$.

(iv) $\mathcal{S}_i$ is closed under non-empty intersections; that is, whenever C is a nonempty subset of $\mathcal{S}_i$ and $\cap C$ is the set of all worlds j such that j belongs to every member of C, $\cap C$ belongs to $\mathcal{S}_i$.

Lewis notes the analogy between his spheres and topological neighborhoods, though he points out that his spheres offer a more comparative sense of similarity.[36] With these systems of spheres, Lewis then proposes the following truth conditions.

Spherical Truth Conditions. A counterfactual 'given a world i, if it were that X, then it would be that Y' is true according to a system of spheres $\mathcal{S}_i$ iff there is sphere centred on i every member of which has the property that, if X is true, then Y is also true.[37]

[36]Indeed, topological neighborhoods do not in and of themselves provide a comparative measure of distance with respect to some fixed element, and moreover they needn't be nested around some element.

[37]There is also the trivial empty case where the counterfactual is deemed true when there is no sphere centred on i containing any world where X is true. This kind of vacuuous truth won't be of much interest to us however.

To say it another way, a counterfactual is true at i iff, if there is an accessible world j in which the antecedent is true, then the consequent is true, and moreover it is also true in every antecedent world k that is at least as close to i as j.

Lewis' truth conditions crucially rely on the concept of world *similarity*. On this notion, Lewis advises against trying to be too precise or quantitative.

Measurability of Similarity. "I have sometimes spoken informally of *degrees* of similarity, as if similarity of worlds could be measured numerically; but I have not assumed that it could be. I have not used any quantitative concept of similarity, but only a comparative concept. One world is more similar than another to a third; but we need never say how much more, and the question how much more need not make sense." p. 50 [16]

Later in [16], Lewis further stresses the point that we ought not to try to make too exact the concept of similarity between worlds.

Inexactness of Similarity. "It is tempting to try to define some exact measure of the similarity 'distance' among worlds [...] We must resist temptation. The exact measures thus defined cannot be expected to correspond well to our own opinions about comparative similarity. [...] Not only would we go wrong by giving a precise analysis of an imprecise concept; our precise concept would not fall within — or even near — the permissible range of variation of the ordinary concept." [16] p. 94-5.

There is something a little unnerving about Lewis' position. His spheres, whose purpose is to capture world similarity, are defined with mathematical precision, and yet in spite of this precision, when we come to think more about the concept simi-

larity, Lewis tells us that it is unavoidably inexact and beyond measure. So in spite of having a relatively clear idea of what makes a counterfactual true, it seems that we shouldn't expect to have the ability of saying whether some counterfactual really is true. As aforementioned, this indeterminateness seems partly due to a form of context dependence, i.e., a counterfactual may be true in one system of spheres yet false under another, but also due to a lack of discernability in the sense that even if we fix some system of spheres, there may be no way of quantifying similarity so as to permit precise adjudication of whether some two worlds are more similar than another.

For worlds as intricate as our own, Lewis' advice against a precise notion of similarity seems wise. Our case is different however. We are interested in spacetimes and in virtue of these being precise mathematical entities, we should at least endeavour to offer a more precise concept of similarity.

3. Spheres of Spacetimes?

Since we're looking to implement Lewis' ideas in the language of spacetimes, the first and obvious thing to do is to think of worlds as spacetimes. More precisely, in spacetime theories such as GR we define a spacetime (M, g) to be a Lorentzian manifold, connected, without boundary, time orientable, Hausdorff, and with countable basis.

The second point is to have some idea as to what we should use for the notion of accessible worlds. Following Lewis, for whom the accessibility relation is highly permissive, let us say that the worlds that are accessible from (M, g) consists in those and only those that are spacetimes of the same dimension.

A natural thing to consider at this point is the notion of laws. Although we can expect there to be various inequivalent choices, at minimum we should require that the laws entail that physically possible worlds are a subset of the set of accessible worlds. Speaking more specifically of GR, the laws would also force us to satisfy the

Einstein field equations, which in the presence of matter fields $(\phi, ...)$ are as follows

$$T[\phi, ...] = G(g) \equiv \mathrm{Ric}(g) - \frac{1}{2}\,\mathrm{Scal}(g)$$

where $T[\phi, ...]$ denotes the stress energy tensor associated with $(\phi, ...)$, and $G(g)$ is the Einstein tensor of g, defined in terms of the Ricci and scalar curvature of g.

Having mentioned the field equations, there immediately arises the issue of whether or not we should the energy conditions. For as is well known, without these the field equations are largely robbed of content. Lewis' concept of physical possibility is thus weaker or stronger depending on whether or not we should impose the energy conditions.

Another thing we might do is to think of some way of capturing Lewis' idea of spheres of worlds around some world (M, g). To get things off the ground, let us fix the topological manifold M so that all the accessible worlds from (M, g) are of the form $(M, \tilde{g})$. With this restriction, we can think of the space of accessible worlds as the space of Lorentzian metrics that M admits, a space usually denoted by $L(M)$. In the case of GR, we could reduce $L(M)$ by getting rid of all those metrics that don't satisfy the above field equations. If however we impose no restrictions on the collection of matter fields $(\phi, ...)$, then the reduced space may turn out no smaller than $L(M)$. This is because if g satisfies the field equations on M, then to make some metric $\tilde{g} \in L(M)$ satisfy the field equations, we simply evaluate its Einstein tensor and fix some collection of matter fields $(\phi, ...)$ on M so that their total stress-energy tensor is equal to the Einstein tensor of $\tilde{g}$. If however we impose restrictions on $T[\phi, ...]$ such as energy conditions, then $L(M)$ will of course be reduced to a much smaller space $E(M) \subsetneq L(M)$. We could go even further and choose to investigate the purely gravitational regime, in which case we only consider metrics $\tilde{g}$ that satisfy the vacuum field equations, $\mathrm{Ric}(\tilde{g}) = 0$. The space of such metrics, say $V(M)$, is of course much smaller still, $V(M) \subsetneq E(M) \subsetneq L(M)$.

Putting these subtleties aside, let us say for the moment that our space of acces-

sible worlds is $L(M)$. By choosing a topology τ on $L(M)$, we obtain a topological space $(L(M), \tau)$. This space has a system of neighborhoods, and we can consider these as candidates for Lewis' concept of spheres.[38] For these neighborhoods to faithfully embody Lewis' spheres, the topology τ on $L(M)$ ought to yield a comparative notion of similarity between spacetimes.

As well as $L(M)$, we can also consider the quotient of $L(M)$ by equivalence classes of metrics that are related by a conformal transformation $g \to \Omega^2 g$. We denote this smaller space by $C(M)$ and it too can be topologized. Naturally, topologies on $L(M)$ will induce a quotient topology on $C(M)$. The main important fact about $C(M)$ is that, because properties concerning the causal structure of g are invariant under conformal transformations, any genericity or stability statement about causal structural properties expressed in terms of $C(M)$ can be equivalently expressed in terms of $L(M)$ (under the corresponding topology).

On $L(M)$, there are basically three sorts of basic topologies commonly in use; see the recent article by Fletcher [10] for a detailed study of these. In order of ascending fineness, these are the so-called C^k compact-open topologies, the C^k open topologies, and the C^k fine topologies. The C^0 open topology on $L(M)$ induces a quotient topology on $C(M)$ known as the interval topology, and its open sets are defined by being composed of metrics g lying in some interval $g_1 < g < g_2$ where $g < g_2$ is shorthand for the lightcones of g being everywhere narrower than those of g_2.

At this point, one might hope that these topologies give us good comparative notion of similarity between spacetimes, and in particular of distance between worlds.[39] One thing to look for are spaces and topologies on these which make

[38] We say that a topology is finer (coarser) if there are more (less) open sets partitioning the space.

[39] In general, topological spaces lack the structure to make sense of the relative closeness of its members. By contrast, spaces with more structure like say metric spaces do yield a precise notion of distance. For instance if (N, d) is a metric space with metric $d : N \times N \to [0, \infty)$, then for three elements $x, y, z \in N$ we say that y is at least as close to x than z is to x iff $d(x, y) \leq d(x, z)$.

these metrizable. Some spaces and topologies won't do. In particular, $C(M)$ with the interval topology is non-metrizable. To see this, note that metrizable spaces are necessarily first countable, but that by a result of Lerner [15], if M is non-compact (which we usually take it to be), then $C(M)$ with the interval topology isn't first countable.[40,41]

The other aforementioned topologies on $L(M)$ are in a sense more well behaved than the interval topology. The C^k open topology for example is defined as follows (note the analogy with the Euclidean topology for $\mathbb{R}^n$).

First, let h^{ab} be the inverse of a smooth Riemannian metric on M. We then define the 'distance' function between the k^{th} partial derivatives of two Lorentz metrics g_{ab},g'_{ab} relative to h^{ab} as

$$d(g, g'; h, 0) = [h^{ru}h^{sv}(g_{rs} - g'_{rs})(g_{uv} - g'_{uv})]^{1/2}$$

$$d(g, g'; h, k) = [h^{a_1b_1} \cdots h^{a_kb_k}h^{ru}h^{sv} \otimes \nabla_{a_1} \cdots \nabla_{a_k}(g_{rs} - g'_{rs})\nabla_{b_1} ... \nabla_{b_k}(g_{uv} - g'_{uv})]^{1/2}$$

With this distance function, we express the basis for the C^k open topology by sets of the form

$$B_k(g, \epsilon; h) = \{g' : \sup_M d(g, g'; h, 0) < \epsilon, \ldots, \sup_M d(g, g; h, k) < \epsilon\}$$

where g ranges over all Lorentz metrics, ϵ ranges over all positive constants, and h ranges over all Riemmanian (inverse) metrics.

Note that $B_k(g, \epsilon; h) \subset B_k(g, \epsilon'; h)$ iff $\epsilon < \epsilon'$. Consequently, the similarity of g' and g'' with respect to g becomes reducable to the question of whether

[40] A space is first countable if, for each point x in a topological space X there exists a sequence $N_1, N_2, \ldots$ of neighbourhoods of x such that for any neighbourhood N of x there exists an integer i with N_i contained in N.

[41] On this point, it may be worth noting that Minguzzi [30] proves the stronger statement that the interval topology isn't even T-sequential. To see the force of the result, recall that first countable $\Rightarrow$ Fréchet-Urysohn $\Rightarrow$ T-sequential.[42] In view of Minguzzi's result, it may be of interest to look for variations on the interval topology that are T-sequential.

there exist a pair of real numbers ϵ' and ϵ'' such that $g' \in B_k(g, \epsilon'; h)$ and $g'' \in B_k(g, \epsilon''; h) \backslash B_k(g, \epsilon'; h)$, in which case we'd say that g'' is more similar than g', as in Figure IV.1.

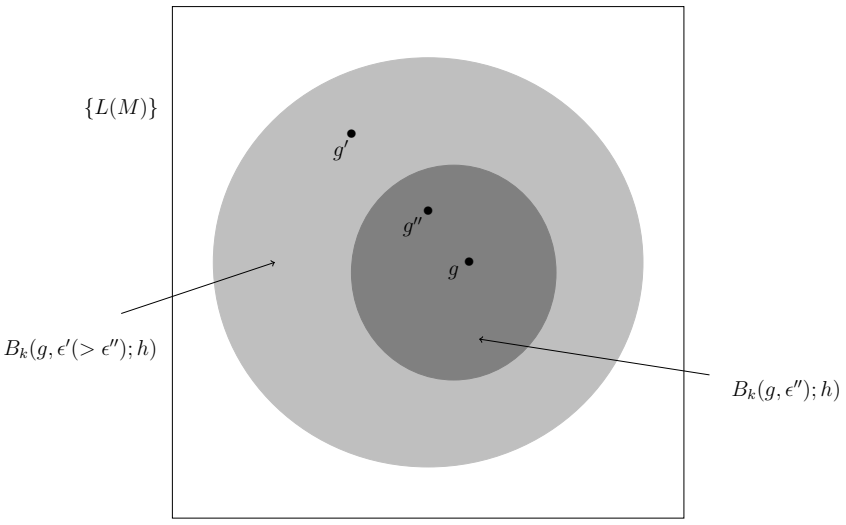


Figure IV.1. Here, the set $B_k(g, \epsilon'; h)$ denotes a collection of possible worlds around g. Membership in such set means some degree of closeness of to g, and here with g'' closer than g'.

Although it may be that such a notion doesn't faithful embody the similarity of spacetimes, it is worth noting that it leads to the following formal translation of Lewis' truth conditions.

Topological Truth Conditions. The counterfactual 'if it were that ϕ, then it would be that ψ' is true in the C^k open sense at (M, g) iff, for all Riemannian metrics h on M, there is a set $B_k(g, \epsilon'; h)$ such that ϕ, ψ are both true for some $g' \in B_k(g, \epsilon'; h)$ and moreover that if ϕ is true for some $g'' \in B_k(g, \epsilon''; h)$ with $\epsilon' > \epsilon''$, then ψ is also.

Remark. Note of course here that the quantification over all metrics h makes these truth conditions rather difficult to assess in a given concrete scenario. Moreover, fixing one metric h isn't any good either because the truth will then depend on one's choice, which is to some extent arbitrary.

Here, the set $B_k(g, \epsilon'; h)$ denotes a collection of possible worlds around g. Membership in such a set means some degree of closeness of to g, and here with g'' closer than g'. So the counterfactual 'if it were that ϕ, then it would be that ψ' is true iff there is some nearby spacetime g' in which ϕ and ψ are true, and moreover that for any ϕ-world g'' which is at least as similar as g', ψ is also true.

In spite of the formal analogy, it isn't clear that the C^k open sense of similarity faithfully embodies Lewis' spheres. A worrying sign is that neither the field equations nor the global structure of the spacetime are visible in the above C^k open sense of similarity, whereas for instance Lewis speaks explicitly about laws (from the outset) as contributing weightily to the overall similarity of worlds.

On this point, we defer to the article of Fletcher [10], which explains in detail what stands in the way of endowing the aformentioned topologies with clear physical significance. Fletcher's main conclusion is that each topology comes with its own sense of similarity, and moreover that it is unlikely that there be a unique and once for all topology that satisfyingly and wholly captures our notion of similarity.

4. Lewis' Similarity Maxims

Lewis' view on laws in [16] is well illustrated by his discussion of the following roulette example. Lewis starts with a deterministic world i in which a roulette lands on black at some time t. He then considers antecedent worlds where the roulette lands on red at time t. Lewis considers two such antecedent worlds, deterministic ones in which facts will have been different from the reference black world for all time prior to t, and ones that are just as the reference world up to some time just

before t but at which point a small miracle occurs and results in the roulette landing on red. He then asks which of these antecedent worlds is closest to i. In prioritizing similarity with respect to particular matters of fact, he suggests that the miracle worlds could be closer.

One Miracle Worlds. "Laws are very important, but great masses of particular fact count for something too; and a localized violation is not the most serious sort of difference of law. The violated deterministic law has presumably not been replaced by a contrary law. Indeed, a version of the violated law, complicated and weakened by a clause to permit the one exception, may still be simple and strong enough to survive as a law. Therefore some of the antecedent-worlds where the law is violated may be closer to i than any of the ones where the particular facts are different at all times before t." p.75 [16]

Lewis then considers what would be said if, in the one-miracle red worlds, there occurs a second miracle erasing the effects of the first. Merely on the basis of facts, those two-miracle worlds would seem to be more similar (to i) than both the deterministic red worlds and the one-miracle red worlds. But here Lewis prefers the one-miracle worlds.

Two Miracle Worlds. "There are others where a second minor miracle occurs just after t, erasing all traces of the first miracle, so that we have two violations of law instead of one but the particular facts from that time on are just as they are in i. If I have decided that a small miracle *before* t makes less of a difference from i than a big difference of particular fact at all times *before* t, then why do I not also think that a small miracle *after* t makes less of a difference from i than a big difference of particular fact at all times *after* t? That is not what I do think: the worlds with no second miracle and divergence must be regarded as closer, since I certainly think it true (at i) that if the wheel had stopped on red at t, all sorts

of particular fact afterward would have been otherwise than they are at i. The stopping on red would have plenty of traces and consequences from that time on." p. 76 [16]

To question this kind of reasoning, Kit Fine [9] proposed the following counterfactual, which we would intuitively like to think of as true.

- 'If Richard Nixon had pressed the button, then there would have been nuclear war.'

Fine points out that Lewis' account (as described) rules counterintuively on this counterfactual. To see this, consider a small miracle occurring just after Nixon's pressing of the button that renders ineffective the deployment of the warhead. In such worlds, things have been as in our world up to Nixon's pressing, and, after a short time (which can made be arbitrarily small), things will proceed more or less as in our world. With respect to the facts, therefore, such worlds are quite like our own. If however no miracle occurs, then the warhead will be deployed, there will be nuclear war, and we obtain worlds that differ wildly from our own, more so it seems than those no-war one miracle worlds. Consequently, the counterfactual seems to be ruled false.

In a later article [18] partly motivated to deal with examples like Fine's, Lewis offers the following maxims to specify what he means by the similarity of worlds. In that same article he also reformulates his truth conditions as follows.

Truth of Counterfactuals. A counterfactual 'if it were that ϕ, then it would be that ψ' is (non-vacuously) true if and only if some (accessible) world where both ϕ and ψ are true is more similar to our actual world, overall, than is any world where ϕ is true but ψ is false.

And now the term 'similarity' is described in terms of the following maxims.

Similarity Maxims. To look for a world that is most similar to another, we are guided by the following maxims, listed in order of decreasing importance.

1. Avoid big, widespread, diverse violations of law.
2. Maximize spatiotemporal regions throughout which there is perfect agreement with respect to particular matters of local fact.
3. Avoid small, localized, simple violations of law.
4. Pay little or no attention to approximate similarity in some particular fact, even if this fact is of great importance to us.

Under these maxims, Lewis argues that worlds in which the nuclear launch mechanism fails are actually less similar (to the actual world) than those in which war occurs. The miracle is now thought of as a big violation of law and thus contributes weightily to the world's overall difference. The violation of the first maxim outweighs the satisfaction of the second, and consequently the nearest worlds become those in which there is war. As desired, Fine's counterfactual may now be ruled as true.

Note that Lewis' refinement makes the concept of laws more important. Not only must we know what the laws are, we must also know what counts as a big or small violations of them. For our sake, the task is now set: what do Lewis' maxims look like in the context of spacetimes?

5. Spacetime Similarity

Lewis' most important maxim is to avoid big, widespread violations of law. We can hazard a guess as to what that should mean in the spacetime context, but let us return to this point with some examples in hand.

His second maxim is that there hold, to the largest degree, perfect agreement between particular matters of local fact. Since g is that which is endowed with chrono-geometric significance, a natural translation of 'local matters of fact' is 'local

properties of the geometry'. Thus, a spacetime version of the maxim could be to maximize the extent of regions that can be isometrically embedded into the reference spacetime.[43] To express this more precisely, the following comes to mind.

Definition. *Let (M, g) be a fixed reference spacetime and μ a Borel measure on M so that $\mu(M)$ is finite.*[44] *Let (M', g') a spacetime containing an open connected subset $U \subsetneq M'$, and let $\phi : (U, g') \hookrightarrow (M, g)$ be an isometric embedding into M. Let (M'', g'') be another spacetime with a corresponding open connected subset $V \subsetneq M''$ and an isometric embedding $\psi : (V, g'') \hookrightarrow (M, g)$. Let $\tilde{U}$ be the maximal open connected domain of ϕ and $\tilde{V}$ the corresponding one for ψ. By paracompactness, the collection of such domains is necessarily countable, and thus we obtain a countable collection $\{\tilde{U}_j\}$ and $\{\tilde{V}_j\}$ with $j \in \mathbb{N}$. We say that (M', g') is at least as 'locally similar' to (M, g) than (M'', g'') iff $\sum_i \mu(\tilde{U}_i) \geq \sum_j \mu(\tilde{V}_j)$, and more so iff $\sum_i \mu(\tilde{U}_i) > \sum_j \mu(\tilde{V}_j)$.*

Remark. Note here of course the issue of which Borel measure is most adequate. Some spacetimes will be deemed more similar under one measure and not so under another. Somehow the measure one uses ought to represent and capture some intuitive notion of how we assign importance to a given region of spacetime. Ought we weight all regions equally, and should we consider the entirety of spacetime in our assessment of observer-based counterfactuals ? More concretely, there is also the issue that such measures may not respect or even tie with the natural notion of a size of a spacetime region as indicated by the spacetime metric g.

This definition provides a necessary and sufficient condition for spacetime local similarity on the basis of agreement of local facts; Figure IV.2 shows the definition at work.

[43] It's worth noting that isometries can sometimes behave a little counterintuitively. To illustrate the point, see the example of Geroch [11] where he has two spacetimes (M, g) and (M', g') which are inequivalent, despite being isometric to one another, i.e., $\phi : M \hookrightarrow M'$ and $\psi : M' \hookrightarrow M$ such that $\phi(M) \subseteq M'$ and $\psi(M) \subseteq M'$.

[44] This notion of measure is described in chapter 3 of [3].

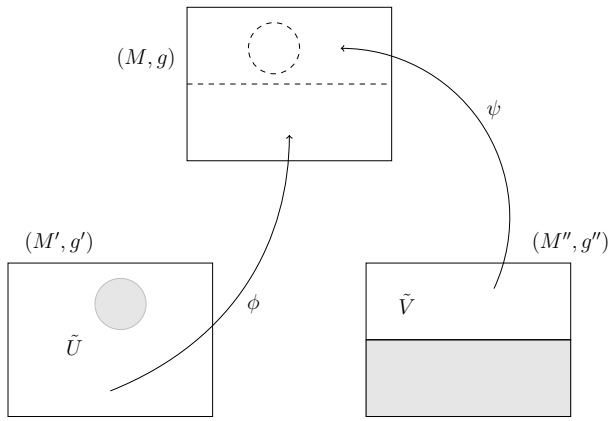


Figure IV.2. The reference world is (M, g) and we're looking to maximize the extent of regions in (M', g') and (M'', g'') that can be isometrically embedded into (M, g). For (M', g'), $\tilde{U}$ is drawn, and likewise for (M'', g''). Once embedded into (M, g), we may compare the size of these regions with a measure coming from (M, g). Here, (M', g') is clearly more locally similar.

Lewis' third maxim is that we should minimize small violations of the laws. Depending on what we take these to be, this could mean that we only permit for small violations of either the field equations, or perhaps the relevant energy conditions. I won't seek to provide a precise definition, but such a thing is presumably possible.

As for Lewis' fourth maxim, we shall simply take it to mean that the approximate agreement of particular matters of local facts is unimportant.

Like Fine, we are now positioned to start considering actual spacetime counterfactuals. We can anticipate there to be at least two sources of inconclusivity to our assessment of such counterfactuals.

The first kind is trivial in that it is simply due to a lack of precision. For example, if a term like 'almost flat' is stated without there being any intention toward a precise definition, then the counterfactual will be inconclusive, but for uninteresting

reasons.

A more important source of inconclusivity is due to a lack of contextual information. Consider for instance the following counterfactual made with respect to a spacetime (N, σ) which satisfies the null energy condition.

- 'If there had been a trapped surface, then there would be an incomplete null geodesic.'

Though none of the terms are vague, there is clear source of inconclusivity arising from a lack of information. For one thing, we can trivially satisfy the consequent by removing sets. That said, we'd like to avoid such cases on the grounds that they miss the point of the counterfactual (which involves a trapped surface and thus concerns the dynamics of the gravitational field).

To aid our assessment of this counterfactual, we have the well known theorem of Penrose [31].

Theorem (Penrose [31])**.** *Let (M, g) be a spacetime admitting a non-compact Cauchy surface, satisfying the null energy condition, and containing a trapped surface $S \subset M$. Then there is an incomplete null geodesic.*

It is important to note that although there are null geodesically complete spacetimes satisfying the null energy condition with trapped surfaces, these invariably involve either CTCs or compact achronal slices (or both). As such, if we had been told that (N, σ) admitted a compact slice or CTCs, then we'd be less inclined to rule the counterfactual as true, than if we'd been told that (N, σ) admits a non-compact Cauchy surface, in which case we'd want to rule it as true. The truth of this counterfactual therefore seems to depend on the global structure of (M, g).[45]

To see the effects of global structure in more detail, let us now consider the

[45] A global property is a property that is not local, and a local property is one that is preserved in all locally isometric spacetimes. Two spacetimes (M, g) and (M', g') are locally isometric iff for every $p \in M$ and $p' \in M'$, one can always find an open neighborhood $O(p)$ and $O(p')$ which isometrically embeds, respectively, into (M, g) and (M', g').

following counterfactual.

Example. Take Minkowski spacetime $(\mathbb{R}^4, \eta)$ with Cauchy surface S as the reference spacetime and consider the following counterfactual.

- 'If the data on S differed in a suitably small way from flat data only within a compact proper subset of a new initial data set V, then the spacetime into which V isometrically embeds would be flat outside a set with compact boundary.'

Let (M, g) be the maximal globally hyperbolic development of V, as in Figure IV.3.

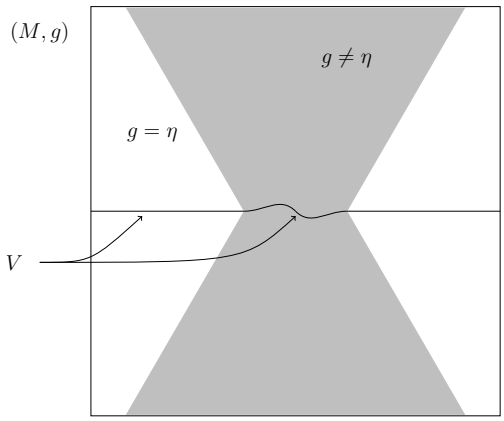


Figure IV.3. The white region is flat, and the shaded region is curved, due to the curvature introduced in V, which then spreads indefinitely.

Here, the antecedent is true but the consequent is false, and moreover the spacetime, like Minkowski spacetime, is globally hyperbolic and inextendible (and thus hole-free).

Now consider a different kind of spacetime (M', g') as in Figure IV.4.

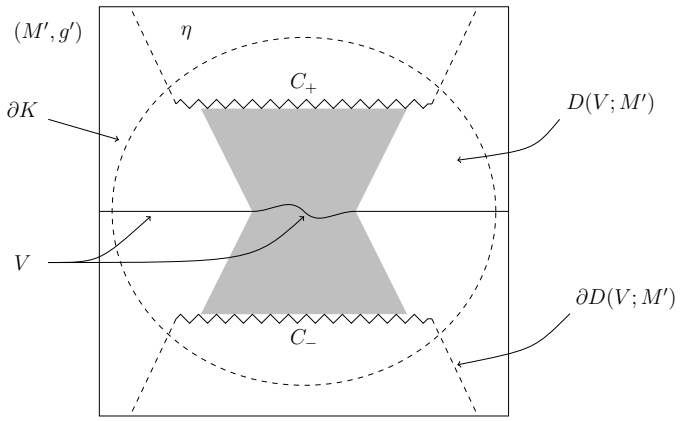


Figure IV.4. With the sets C_+ and C_- removed, we can consider a new spacetime created by patching in flat regions outside of $D(V; M')$. The resulting spacetime (M', g') is non-flat in a region that can be made arbitrarily small, and the set K has a compact boundary ∂V.

These are constructed as follows. Cut out of (M, g) two compact spacelike hypersurfaces C_- and C_+ that sit above and below V so that both C_- and C_+ extend into flat portions of (M, g). In the resulting spacetime $(M \backslash C_- \cup C_+, g)$, the set $D(V; M \backslash C_- \cup C_+)$ no longer covers the whole of $M \backslash C_- \cup C_+$. Since the boundary of $D(V; M \backslash C_- \cup C_+)$ is flat by continuity, we can remove $M \backslash C_- \cup C_+ - D(V)$ from $M \backslash C_- \cup C_+$ and replace these with flat regions. Call the resulting spacetime (M', g'). Here, both the antecedent and consequent are true.

With respect to the reference spacetime $(\mathbb{R}^4, \eta)$, we now have two classes of antecedent spacetimes — hole-free and holed — with only the latter satisfying the consequent. So which is most similar to $(\mathbb{R}^4, \eta)$? In the sense of local similarity, the answer is clearly (M', g'). However if we prioritize global structure, then it is (M, g).

The point of this example is that it is basically analogous to Lewis' roulette. The small miracle which leads to a 'small' lapse in determinism is like the hole

created by removing C_-. For Lewis the miracle permitted to land on red; and for us the hole allowed us to introduce some curvature where previously there hadn't been any. Likewise, the second hole caused by C_+ is like Lewis' second miracle that erases the effects of the first, which in our case erases the non-flatness that had been introduced.

Note that the two-miracle spacetimes are more locally similar to $(\mathbb{R}^4, \eta)$ than the one-miracle spacetimes. Note also that there is an important difference with Lewis' judgement in that he thinks of the two-miracles worlds as more different than the one-miracle worlds. This difference is owed to Lewis' distinction between future and past directions, which the above definition of similarity doesn't capture (which is perhaps more natural in a time symmetric theory like GR). Putting two-miracle spacetimes aside, we note that the local similarity definition would rule as in Lewis' roulette. Namely, that one-miracle worlds may be closer to the black deterministic world than the red deterministic world. Here, by prioritizing local similarity over global structure, we come to the same verdict.

Recalling Fine's example, if we were to turn around (as Lewis does) and all things considered treat the one-miracle spacetime as more different, then we'd favor a singly holed spacetime over (M', g'). Clearly, then, what Lewis calls 'big' violation of law is what we think of as differences in global structure. The task is therefore to understand how the global structure of spacetimes may differ from one another and how to judge relative global similarity. The task initially sounds daunting because the term 'global structure' in GR actually includes an infinite number of possible definitions; see for instance chapter 3 of [3]. Let us thus start by considering those which seem most fundamental.

Perhaps the most basic one resides in the properties of the underlying spacetime manifold M and its topology. As such, it seems natural to prioritize spacetimes whose manifolds are homeomorphic. Note that sameness of topology is rather weak on its own: $(\mathbb{R}^4, \eta)$ and the Godel spacetime $(\mathbb{R}^4, g_G)$ have the same topology but the former is flat and globally hyperbolic, whereas the latter is curved and totally

vicious.[46]

The next most important global property concerns the spacetime's causal structure. As such, we might seek to prioritize spacetimes with the same causal structure.

Having now prioritized topological and causal structure, we might speculate that when spacetimes are topologically and causally equivalent, then we defer to our previous notion of local similarity to complete our assessment of similarity. As we shall now see however, sameness in topological and causal structure is still insufficient to fix a unique class of most similar spacetimes. To see this, consider first the following definition.

Let (M, g) and (M', g') be any two spacetimes. We say that (M', g') is an X-extension of (M, g) iff, for any open connected subset O' of M' with compact closure in M', if there is an isometric embedding $\phi : O' \to M$, then $\phi(O')$ has compact closure in M. An X-extension is called trivial if it is simply the same of the original spacetime, i.e., the same manifolds and metrics related by an isometry. To see an example, we note for instance that Minkowski spacetime $(\mathbb{R}^4, \eta)$ isn't an X-extension of the bottom half of it $(\mathbb{R}^4_-, \eta)$.

Example. Let the reference spacetime be $(\mathbb{R}^4_-, \eta)$, and consider the following counterfactual.

- If $(\mathbb{R}^4, \eta)$ were a non-trivial X-extension, then there would be non-flat regions but only with compact boundary.

Let $\{A\}$ be the class of spacetimes of the form $(\mathbb{R}^4_-, \sigma)$ where σ is flat outside a infinite slab $\subset \mathbb{R}^4$, as in Figure IV.5. Note here that the antecedent is true but the consequent is false.

[46] A spacetime is totally vicious iff there passes a CTC through every point. See [14] for more details about Godel's spacetime.

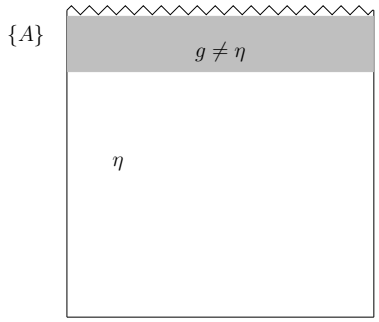
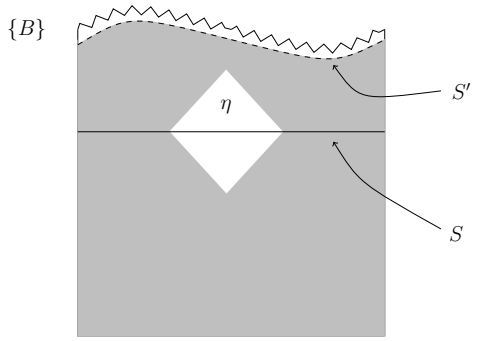


Figure IV.5. Start with Minkowski spacetime, cut it along the spatial direction, change the flat metric η within an infinite horizontal slab, and then smoothly merge it with another half of Minkowski spacetime. The flat portion of the spacetime certainly doesn't have a compact boundary.

Now let $\{B\}$ be the class of vacuum spacetimes as in Figure IV.6., which are constructed as follows. Consider the maximal Cauchy development of an asymptotically flat Cauchy surface S with topology $\mathbb{R}^3$ that contains some compact flat subset but that doesn't isometrically embed into $(\mathbb{R}^4, \eta)$.[47] Finally, remove from this maximal Cauchy development everything $\subset J^+(S')$ for some Cauchy surface $S' \supset J^+(S)$ extending beyond the flat portions. Here, both the antecedent and consequent are true.



[47] Consider for instance sufficiently small deviations to flatness.

Figure IV.6. Start with an asymptotically slice S which contains some flat portion, whose domain of dependence will also be flat. Let the slice develop according to the field equations under some energy conditions, and then get rid of half of it by removing what lies on or above S'. The resulting spacetime is flat in a set with compact boundary.

So now we ask: which of $\{A\}$ or $\{B\}$ is most similar to $(\mathbb{R}^4_-, \eta)$? Judging by local similarity, it is $\{A\}$, but if we include the global property of Cauchy determinism, then it is $\{B\}$. Here, we say that a globally hyperbolic spacetime is Cauchy deterministic iff for any one of its Cauchy surfaces, the spacetime is the unique Cauchy development that respects the field equations along with a suitable energy condition.[48]

So which is it? One natural choice is to prioritize Cauchy determinism. The reason for this is that the property has global consequences when global hyperbolicity holds, and, as such, it fits into our Humean description of law, more so it seems than the above sense of local similarity.

Combining the intuitions gained from these examples, we now obtain the following similarity maxims.

Spacetime Similarity Maxims. To look for spacetimes most similar to (M, g), we are guided by the following maxims.

1. Prioritize above all spacetimes with manifolds that are homeomorphic to M.
2. Prioritize spacetimes with the same causal structure as that of (M, g), i.e. those which lie on the same level of the causal hierarchy.
3. Seek to preserve local Cauchy determinism if true in (M, g).
4. Maximize local agreement with respect to (M, g) in the sense of the above

[48]For example, $(\mathbb{R}^4, g)$ with g only flat outside a compact set isn't Cauchy deterministic despite being globally hyperbolic and admitting a Cauchy surface that isometrically embeds into $(\mathbb{R}^4, \eta)$.

definition.

5. Minimize quasi-local violations of the field equations, or energy conditions (if imposed).

Note that 1-4 have been defined precisely, and that 5 can also be. Note also that with these maxims, the examples above may be adjudicated upon conclusively (i.e., the hole-free class for the former, and $\{B\}$ for the latter). Another interesting thing about these sets of maxims is that Lewis' original emphasis on local similarity seems to have been relegated somewhat.[49]

6. The View from our World

In the current picture, counterfactuals made from our world depend on the global and local properties of our spacetime. Unfortunately, since we don't know these[50], the best we can do is to explore what the properties might be and what that would imply.

6.1. What are the Properties of our Spacetime?

On this question, the first point to note is that properties of a spacetime can be of various types. Other than the distinction between global and local properties encountered above, another one is between categorical and modal properties.

Here, a property of (M, g) is said to be categorical if it refers explicitly to M and g (and $(\phi, ...)$ if there are matter fields). For example, global hyperbolicity is the requirement that (M, g) be causal and that $J^+(p) \cap J^-(q)$ be compact for all $p, q \in M$, and is thus categorical.

By contrast, a property is said to be modal if it refers to some background class

[49]Whether and how that might impact the relevance of Humean Supervenience views in the context of spacetime theories like GR is a question that goes beyond the scope of this article.

[50]See III for discussion of this point.

of spacetimes. For example, the extendibility of (M, g) is modal since it refers to another spacetime into which (M, g) could isometrically embed.

Naturally, modal properties can imply categorical ones and vice versa. A classic example is stable causality. One of its modal definitions is that a spacetime (M, g) is stably causal iff there is a metric $\tilde{g} \in C(M)$ such that the lightcones of $\tilde{g}$ are strictly everywhere wider than g and $(M, \tilde{g})$ is causal. By a well known theorem (whose rigorous proof only came a few years ago), that turns out to be equivalent to the categorical property of admitting a temporal function, i.e., a smooth function $f : M \to \mathbb{R}$ with past directed timelike gradient ∇t.

Theorem (Sanchez [29])**.** *A spacetime is stably causal iff it admits a temporal function.*

Less well known is the interesting and recent result that global hyperbolicity is stable in the C^0 open topology, or equivalently the interval topology. That is, the categorical property of being globally hyperbolic implies the modal one of being surrounded by an open set of spacetimes also sharing this property.[51]

Theorem (Benavides Navarro - Minguzzi [4])**.** *Global hyperbolicity is stable in the interval topology, and thus in the C^0 open topology.*

In addition to these distinctions, there is also the question of whether certain properties are stable and or generic under some topology.[52] Note of course that statements about stability and genericity are sensitive to the topology one chooses. In particular, a stable property will be do so in any finer topology though not necessarily in any coarser topology, and likewise a property that is generic in one topology needn't be so in a finer one. Stability and genericity are often thought of as necessary

[51]Recall that a property of $g \in L(M)$ is said to be stable under some topology τ just in case there is a set $O(g) \subset L(M)$ that is open with respect to τ all of whose members also have the property.

[52]Recall that a property is said to hold generically on a collection $S \subset L(M)$ iff it holds on an open subset of S that is dense in S, and that a subset A of S is dense when all elements of S are either in A or are limit points of A.

and or sufficient conditions for physical significance. Consider Hawking for instance.

Hawking "[T]he only properties of space-time that are physically significant are those that are stable in some appropriate topology" [...] "For physical purposes it is sufficient to prove that a theorem holds generically." p. 395 [13].

The intuition here is that our world's properties shouldn't be what they are in virtue of some atypical configuration among the space of possibilities. That is, it should be the case that if the state of the world were different from what it actually is in a sufficiently small way (the meaning of which is topology dependent), then the spacetime's overall properties should be the same as they are in our world, or at least approximately so.

Although the idea that physical significance is tracked by genericity and or stability is practical, there is the issue of how to best choose the topology. Consider for instance the following recent result of Fletcher.

Proposition (Fletcher [10])**.** *Chronology violating spacetimes are generic in $L(M)$ for every compact open topology.*

As Fletcher describes, this result says more about the compact-open topology than it does about chronology violation. Choosing the right topology for the problem can be difficult, one needs to strike a good balance between utility and the physical intuition in question.

Bar issues about which topology to use, Hawking's suggestion initially seems to provide a clear guideline as to which properties we might reasonably expect of our spacetime. However, what isn't made clear by Hawking is that stability and genericity properties actually depend on the underlying global structure of the spacetime; and on the following point I would like to express my gratitude to Professor Manchak for englightening discussions, and for his articles [24], [25], the latter of which makes a related point.

Take for instance Manchak's recent result [23] that the property of 'effective

completeness' is unstable under the C^k open topology.[53] Under Hawking's intuition, Manchak's result would cause one to wonder whether effective completeness is a physical significant property. Upon inspection of the proof however, one sees that it involves acausal spacetimes: a causality violating compactification of Minkowski spacetime and Misner spacetime. As such, it may very well be that effective completeness is stable in the C^k open topology if we assume for instance that the spacetime is globally hyperbolic.

To see that such a thing may happen, consider the property of being geodesically complete. Williams [34] has produced examples showing that both geodesic completeness and geodesic incompleteness may fail to be stable in the $C^{r\geq 0}$ open topology. In view of these examples, one cannot deduce that geodesic completess or incompleteness is a physically significant property. However, if restrictions are made on the spacetimes in consideration, then it *is* possible to show both the stability of geodesic completeness and incompleteness. Consider for instance the following results by Beem, Ehlrich, and Easeley.

Theorem (Beem, Ehrlich, Easley [3] theorem 7.23)**.** *Let (M, g) be a null geodesically incomplete FLRW spacetime. Then there is a C^1 neighborhood $U(g)$ of g in $L(M)$ of globally hyperbolic metrics such that all null geodesics of $(M, \tilde{g})$ for $\tilde{g} \in U(g)$ are incomplete.*

Here, we start with a Friedman-Lemaitre-Robertson-Walker (FLRW) cosmological solution, and it is proved that it is stably incomplete.

Another result shows that geodesic completeness is under certain circumstances stable.

Theorem (Beem, Ehrlich, Easley [3] corollary 7.27)**.** *Let (M, g) be a globally hyperbolic spacetime. If (M, g) is non-spacelike geodesically complete, then there is a C^1 neighborhood of g in $L(M)$ such that each $\tilde{g} \in U(g)$ is nonspacelike complete.*

[53]The term is defined in a previous paper [22] and lies in the following sequence of implications: geodesic completeness $\Rightarrow$ effective completeness $\Rightarrow$ hole-freeness.

The resulting viewpoint here is that physical significance, when cashed out in terms of stability and or genericity, is not only topology sensitive, it is also world relative. That is, a property may be physically significant in one world but not in another. Consequently, if we want to know which properties are physically significant, then not must we only investigate their stability and genericity with respect to various topologies, we must also keep track of the global structural properties under which our statements are proven. On a purely practical note, this means proving many more theorems and coming up with many more examples.

6.2. Is Strong Cosmic Censorship Relevant?

I should first note here that what follows is best read in conjunction with III, where the conjecture of strong cosmic censorship (SCC) is discussed in more detail.

Being interested in the properties of our spacetime, it seems that we ought to consider the widely discussed yet undefined notion that a spacetime be 'physically reasonable'. Whatever the notion is, it seems definitional that our spacetime should qualify as such.[54] This notion isn't simply discussed by philosophers, it also lies at the centre of Penrose's original formulation of the SCC [32], which states that physically reasonable spacetimes are globally hyperbolic.

The potential relevance of the SCC is that if it were true in some suitable way, then we could hope to use it to simplify our analysis of counterfactuals. In particular, with our own reference spacetime now globally hyperbolic, it would follow from the previous similarity maxims that we prioritize spacetimes that have a Cauchy surface topologically equivalent to our own. The result of this is a dramatic simplification of counterfactuals, at least those whose antecedent worlds are consistent with global hyperbolicity. For in that case, assessing similarity within this reduced comparison class of nearby spacetimes becomes a question of similarity between two sets of Cauchy surfaces. Since these are complete Riemannian 3-manifolds with

[54]What would be the point of such a notion if that were not so?

specific fields specified on them, the technical task of comparison is made much more tractable.

Notice of course that the reverse is also a possibility. Namely, that if any one (or more) of the following occurs, then the similarity of spacetimes from our own will be harder to assess.

(i) the SCC is false,

(ii) whatever statement that is true doesn't actually allow us to infer something about our own spacetime,

(iii) we somehow find a good reason to believe that our spacetime isn't best described by a globally hyperbolic spacetime.

Moreover, since the SCC may be false to varying degrees, the more it fails, the harder it gets to judge similarity.

7. Conclusion

Under broadly Lewisian guidelines, we've now seen how we may come to think of certain spacetime counterfactuals as true. Rather than any particular stand-alone result, the analysis has led to the following set of perspectives.

1. It is possible to formally translate Lewis' counterfactual semantics in the spacetime context, but there are various problems in doing so.

2. The glitch in Lewis' original account pointed out by Fine has an analog in the spacetime context, and that its resolution is brought by prioritizing the global structure of spacetimes. The result is a set of spacetime similarity maxims, analogous to those proposed by Lewis.

3. Spacetime counterfactuals made from our some reference spacetime (such as our world) depend on a number of mathematical questions whose answers are currently unknown. This includes a refinement of stability and or genericity

as world relative concepts, and a connection with the SCC in the sense that its truth would greatly simplify our assessments of similarity.

References

[1] Armstrong, D.M., (1983), 'What Is A Law Of Nature?', Cambridge University Press.

[2] Beebee, H., (2000), 'The Non-Governing Conception of Laws of Nature', Philosophy and Phenomenological Research, Vol. 61, No. 3, pp. 571-594.

[3] Beem, J. K., Ehrlich, P. E., and Easley, K. L.: Global Lorentzian Geometry. New York: Marcel Dekker Inc. (1996)

[4] Benavide Navarro, J. J., Minguzzi, E., (2011), 'Global hyperbolicity is stable in the interval topology', Journal of Mathematical Physics 52, 112504.

[5] Callender, C., (2017), 'What makes Time special?', Oxford University Press.

[6] Carroll, J.W., (1994), 'Laws of Nature', Cambridge University Press.

[7] Cohen, J., Callender, C., (2009), 'A better best system account of lawhood', Philosophical Studies, 145(1):1-34.

[8] Earman, J., (1995), 'Bangs, Crunches, Whimpers, and Shrieks, Singularities and Acausalities in Relativistic Spacetimes', Oxford University Press.

[9] Fine, K., (1975), 'Critical Notice of *Counterfactuals*', Mind, 84:451–458.

[10] Fletcher, S., (2016), 'Similarity, Topology, and Physical Significance in Relativity Theory', British Journal for the Philosophy of Science 67.2: 365–389.

[11] Geroch, R., (1971), 'General Relativity in the Large', General Relativity and Gravitation 2.1:61–74.

[12] Geroch, R., (1970), 'Singularities', Relativity. Ed. Moshe Carmeli, Stuart I. Fickler, and Louis Witten. New York: Plenum.

[13] Hawking, S., (1971), 'Stable and Generic Properties in General Relativity', General Relativity and Gravitation 1.4: 393–400.

[14] Hawking, S., Ellis, G.F.R., (1973), 'The large scale structure of space-time', Cambridge University Press.

[15] Lerner, D., (1973), 'The Space of Lorentz Metrics', Communications in Mathematical Physics 32: 19–38.

[16] Lewis, D., (1973), 'Counterfactuals', Oxford: Blackwell

[17] Lewis, D., (1979), 'Counterfactual Dependence and Time's Arrow', Nous, Vol. 13, No. 4, Special Issue on Counterfactuals and Laws (Nov., 1979), pp. 455-476.

[18] Lewis, D., (1986), Philosophical Papers, Volume II, Oxford University Press.

[19] Lewis, D., (1994), 'Humean Supervenience Debugged', Mind, 103: 473–490

[20] Loewer, B., (1996), 'Humean Supervenience', Philosophical Topics 24, 101-127

[21] Loewer, B., (2004), 'Humean supervenience', In Carroll, J., editor, Readings on Laws of Nature, pages 176-206. University of Pittsburgh Press.

[22] Manchak, J., (2014), On Space-Time Singularities, Holes, and Extensions, Philosophy of Science 81(5):1066-1076

[23] Manchak, J., Some "No Hole" Spacetime Properties are Unstable, Foundations of Physics 48 (11):1539-1545 (2018)

[24] Manchak, J., Epistemic "Holes" in Space-Time, Philosophy of Science, 83 (April 2016) pp. 265–276.

[25] Manchak, J., General Relativity as a Collection of Collections of Models, preprint via philsci-archive.pitt.edu/14882/

[26] Marqués, F., Deforming three-manifolds with positive scalar curvature, Annals of Mathematics Second Series, Vol. 176, No. 2 (Sept. 2012), pp. 815-863

[27] Maudlin, T., (2007), 'The Metaphysics Within Physics', Oxford University Press.

[28] Minguzzi, E., (2012), 'Causally simple inextendible spacetimes are hole-free', Journal of Mathematical Physics 53-062501.

[29] Minguzzi, E., Sanchèz, M., (2008), 'The causal hierarchy of spacetimes', in H. Baum and D. Alekseevsky (eds.), vol. Recent developments in pseudo-Riemannian geometry, ESI Lect. Math. Phys., p. 299 – 358.

[30] Minguzzi, E., Rinaldelli, M., (2009), Widening the lightcones on subsets of spacetime: some variations to stable causality, Class. Quant. Grav. 26:175004

[31] Penrose, R., (1965), 'Gravitational collapse and Space-Time Singularities', Phys. Rev. Lett., 14, 57-59.

[32] Penrose, R., (1979), Singularities and Time-Asymmery. In General Relativity: An Einstein Centenary Survey, ed. S. Hawking and W. Israel, 581–638. Cambridge: Cambridge University Press.

[33] Wald, R., (1984), 'General Relativity', University of Chicago Press.

[34] Williams, P. M., (1985), Completeness and its stability on manifolds with connection, Ph. D. Thesis, Dept. Math., Univ. of Lancaster.

Appendix

The point of this appendix is to motivate some deep and difficult mathematical questions on the basis of the Lewisian ideas so far described in the setting of general relativistic spacetimes.

It is clear from sections 5 and 6 that globally hyperbolic spacetimes are of great importance and moreover that the simplifications they bring render them ideal case studies. In particular, given a globally hyperbolic reference spacetime (M, g) with Cauchy surface V satisfying some suitable energy condition, we can inquire about the space of nearby spacetimes. Under the similarity maxims (1)-(5), this is a collection of spacetimes $(M, \{\tilde{g}\})$ with Cauchy surface having the topology of V and $\{\tilde{g}\}$ denoting the class of Lorentzian metrics satisfying the field equations with the relevant energy condition. By standard theorems aforementioned, this collection may be thought of as the space of initial data sets that satisfy the constraint equations with the desired energy condition. This moduli space for the constraint equations is a topological space with various characteristics (eg., connectedness, compactness, homotopy, etc.) which capture the mathematical landscape associated with the collection of the nearest possible spacetimes.

The interest in studying this moduli space is to understand whether proximity of spacetimes within the moduli space faithfully captures an intuitive notion of similarity. Suppose for instance that for a specific reference spacetime we find that the moduli space of a particular class of solutions to the equations is disconnected. A natural question is then whether we should take spacetimes belonging to different components as further apart than members within the same component. Such questions are presumably quite difficult to answer, especially given that our current understanding of the moduli space is virtually null, and that our understanding of how global spacetime geometry given knowledge merely of a Cauchy surface is only in its infancy.

At present, the only result I know of in this direction is the theorem of Marqués [26] that the space of asymptotically flat and maximal (i.e., trace $k = 0$) solutions to

the vacuum constraint equations on $\mathbb{R}^3$ is path connected in the relevant topology. This difficult theorem is of profound mathematical interest, and there will be many more like it.

V. Conclusions and Future Outlook

The main original elements of this thesis have been its mathematical results and the explanations which have come with them. These are designed to improve or build upon previous ones, some more famous than others, due to Geroch, Hawking, Kodama, Malament, Manchak, Penrose, Tipler, and Wald.

On II.1 and II.2

For II.1 and II.2, the meaning of the results are somewhat self-evident. In particular, for II.1 we have seen that what suffices for the area non-decrease of black holes are positive energy contributions within a vicinity of the event horizon, and for II.2 that the Hawking Penrose singularity theorem can be superseded somewhat by a theorem which is geared specifically towards black holes and which allows for chronology violation within their interior. The questions left open for II.1 concern quantum field theory in curved spacetime, and for II.2 the formation of closed timelike curves upon collapse. The prospect of seeing concrete answers in the next few of years is unlikely, but these issues may come to be more firmly understood.

On II.3

The results in II.3 are more difficult to interpret. The reason is that they are tied to the semi-classical physical description of black holes, which is not yet fully understood. As if often the case when rigour is lacking and a subject is poorly understood, the literature diffuses into isolated bubbles which basically evolve independently of one another. Among this unchartered sea of contradictory ideas, there do seem to be some important ideas that are worth investigating from a foundational perspective.

One of these is whether it is possible to define and do quantum field theory in curved spacetimes that are causally discontinuous. Finding good and concrete reasons explaining why it this is or isn't possible would bring us closer to a rigorous understanding.

Another is to understand better the recent soft hair arguments of Hawking,

Perry, and Strominger [1]. Their ideas have led to a flurry of activity on the subject which some mathematical relativists have started to pay attention to. Their idea has yet to be treated from a foundational angle, though it would seem like a good thing to do in view of the growing interest.

Finally and most ambitiously, the foundational and physics community is in need of a better grip of the broad umbrella term of 'holography' and 'AdS/CFT'. The enormous growth of this literature symbolises something of a new paradigm in high energy physics, and it would be good to see this area more fully understood from the foundational side. Wallace's article is likely to be a good entry point here.

On III

Here I have sought to consider the epistemic status of observers in classical general relativistic spacetimes. The number of results in the literature had essentially been contained in the aforementioned papers of Glymour, Earman, Geroch, Malament and Manchak. Of these authors, the most interesting and general ones are undoubtedly those of Manchak, which are based on those of Malament.

As explained, it was found that Malament's example relies directly on the existence of observer horizons (as in de Sitter spacetime), and that Manchak's realization of Malament's much more ambitious proposal of underdetermination is achieved by cutting holes and destroying the causal structure of the spacetime. That isn't to say that there is a less ambitious underdetemination result that doesn't so radically affect the global causal structure of the original spacetime.

Among the key observations is that the definition of observational indistinguishability as it appears in Manchak's construction of (M', g') isn't really a faithful embodiment of the concept that there are no experiments that a collection of observers could perform to distinguish between (M, g) and (M', g'). Moreover, part of the subtlety here is that if two spacetimes are genuinely observationally indistinguishable when there are no experiments which can in principle be carried out by a collection of observers which would come out as different, then Malament's example

is but that Manchak's (M', g') isn't.

In spite Malament's observer horizons and Manchak's introduction of holes, the overall message of their arguments remains. That is, observers in general relativistic spacetimes do in fact face epistemic constraints. The results proved here only say that 'if the world is a certain way, then there may not be such a strong sense of epistemic undetermination'. Crucially however, observers won't be capable of checking whether their own spacetime is indeed of that type.

The second task in III was to consider the conjecture of strong cosmic censorship. Owing to the absence of a definition for the concept of 'physically reasonable' spacetimes, the mathematical emptiness of Penrose's original hypothesis was outlined. The mathematicians' take on the conjecture and its translation in terms of the Cauchy problem was explained. Various difficulties with regards to questions of interpretation were described. The essential problem being that the language of the Cauchy problem is rarely physically transparent. As described however, recent progress suggests that there we will soon obtain a better understanding of what mathematical and physical mechanisms could bear on the question of censorship, even if the idea of a master theorem is currently way out of reach.

On IV

Here I have not sought to prove any one particular theorem. Rather, I have tried to translate certain broad and basic aspects of Lewis' counterfactual semantics for the purpose of talking precisely about counterfactuals involving spacetime and their properties. The idea, in short, has been to explain, perhaps only in part, what Lewis might have said about such counterfactuals if he had known about general relativity and the subtleties associated with global spacetime structure.

The outcome has been broadly positive. His ideas can be translated somewhat, and they motivate a number of further mathematical questions with definite answers that we don't yet know. It is now a matter of time and energy whether these are worked out.

References

[1] Hawking, S. W., Perry, M., Strominger, A., Soft Hair on Black Holes, Phys. Rev. Lett. 116, 231301 (2016)