

Time Travel and Time Machines

Chris Smeenk and Christian Wüthrich

Version 2.17

1 Introduction

The general theory of relativity allows an abundant variety of possible universes, including ones in which time has truly bizarre properties. For example, in some universes the possible trajectory of an observer can loop back upon itself in time, to form what is called a closed timelike curve (CTC). In these universes time travel is possible, in the sense that an observer traversing such a curve would return to exactly the same point in spacetime at the “end” of all her exploring. Such curves and other similarly exotic possible structures illustrate the remarkable flexibility of general relativity (GR) with regard to the global properties of spacetime. Earlier theories such as Newtonian mechanics and special relativity postulate a fixed geometrical and topological spacetime structure. In contrast, GR tolerates a wide variety of geometries and topologies, and these are dynamical rather than fixed *ab initio*. This toleration does have bounds: the theory imposes weak global constraints on possible universes, such as requiring four-dimensionality and continuity, along with local constraints imposed by the basic dynamical laws (Einstein’s field equations) and the requirement that locally the spacetime geometry approaches that of the special theory of relativity. But within these bounds flourish an embarrassingly rich collection of possible topologies and geometries that depart quite dramatically from the tame structures apparently compatible with our experience.

What does the existence of solutions with such exotic structures imply regarding the nature of space and time? Below we will assess different answers to this question before offering our own. But the question itself has to be disambiguated before we can sketch answers to it. First, what do we mean by the “existence of solutions” with exotic structures? At a minimum we require that these are solutions of the field equations of GR, and in that sense “physically possible” models according to the theory. But one recurring theme of the discussion below is that this is a very weak requirement, due to the possibility of constructing “designer spacetimes” that satisfy the field equations only, in effect, by stipulating the right kind of matter and energy distribution to produce the desired spacetime geometry. Second, much of the discussion below will focus on GR, our best current spacetime theory. But part of the reason for interest in these solutions is the light their study may shed on the as yet unformulated theory of quantum gravity, so we will also discuss hybrid theories that include quantum effects within GR. Third, our focus below will be on one kind of “exotic” structure – namely, CTCs. We expect that arguments roughly parallel to those below could be run for other kinds of exotic structure. Finally, we hope to isolate the *novel* consequences of the existence of exotic spacetimes, distinct from other lessons of special and general relativity.¹

One answer to our question simply denies the relevance of exotic spacetimes entirely. Many physicists and philosophers have argued that solutions with CTCs are logically or metaphysically

¹Two other essays in this volume address these other implications of relativity theory for understanding the nature of time: Savitt focuses on the implications of special relativity, and Kiefer discusses the problem of time in quantum gravity.

impossible, based on paradoxes of time travel such as the grandfather paradox. Suppose Kurt travels along a CTC and has Grandpa in his rifle sight, finger at the trigger. Either outcome of this situation seems wrong: either Kurt succeeds in killing Grandpa, preventing the birth of his father Rudolf and his own conception, or something such as a well-placed banana peel mysteriously prevents Kurt from fulfilling his murderous intentions. In §2 below we will argue that this paradox (and others) do not show that spacetimes with CTCs are logically incoherent or improbable; what they show instead is that in spacetimes with CTCs questions of physical possibility — e.g., whether it is possible for Kurt to kill Grandpa — depend upon global features of spacetime. We further argue against the idea that one should rule out exotic structures *a priori* by imposing stronger restrictions on “causal structure” than those imposed by GR. §3 reviews the ideas of causal structure in GR that give content to this debate.

A second answer delimits the opposite extreme. On this view, the existence of models with CTCs is taken to imply directly that there is no “objective lapse of time” according to GR. Gödel famously offered an argument for this conclusion based on his discovery of a spacetime with the property that a CTC passes through every point. We will discuss Gödel’s solution and his arguments based on it in detail in §4. The problem we will focus on in assessing Gödel can be stated more generally: if we do not take the exotic spacetimes to be physically viable models for describing the entire universe or particular systems within it, what does their existence reveal regarding the nature of space and time?

Our own answer falls between the two extremes. The question proves to be a fruitful and difficult one due to our lack of understanding of the large scale dynamics of GR and the space of solutions to the field equations. Suppose we allow that the existence of alternative cosmological models such as Gödel’s that are not viable descriptions of the observed universe does not have direct implications for the nature of time in our universe. We can still ask the following question: what is the nature of time in a class of solutions that are directly accessible from our universe, in the sense that an arbitrarily advanced civilization could reach them by locally manipulating matter and energy? If it were possible to *create* CTCs via local manipulations — in effect, to operate a *time machine* — then we could argue much more directly than Gödel against the existence of an objective lapse of time. In §5 below we offer a definition of a “time machine” along these lines. However, even if a time machine so defined proves to be impossible it is tremendously important to understand *why* it is impossible. Our general approach to this issue is familiar from Earman (1986)’s treatment of determinism: pushing a theory to its limits often reveals a great deal about its content, and may lead to refinement of core principles or even provide stepping stones to further theory. In this case, a proof that TMs are impossible could take the form of delimiting some range of solutions of the field equations as “physically reasonable” and showing that these solutions do not give rise to CTCs or other exotic structures. This way of formulating the problem brings out the parallels with Penrose’s “cosmic censorship conjecture,” as we will discuss in §6. Finally, the vast physics literature on time travel and time machines has been inspired by intriguing connections with quantum field theory and quantum gravity (the topics of §7 and §8).

2 The paradoxes of time travel

So what is time travel? The standard answer among philosophers, given by David Lewis (1976, 68), is that time travel occurs in case the temporal separation between departure and arrival does not equal the duration of the journey. However, this is not a necessary condition for time travel. Presumably, Lewis and everyone else should want to include a case when the time lapse between departure and arrival equals the duration of the journey but the arrival occurs *before* the departure.

More significantly, we also claim that Lewis’s definition does not state a sufficient condition for an interesting sense of time travel within the context of modern physics. Readers familiar with special relativity may have already asked themselves what Lewis might mean by temporal separation between arrival and departure. Due to the relativity of simultaneity, observers in relative motion will generally disagree about the temporal separation between events. We could try to skirt this difficulty by defining the temporal separation as the maximal value measured by any observer (corresponding to the proper time elapsed along a geodesic connecting the two events) or by taking advantage of symmetries in a particular model in general relativity. For example, we could exploit the symmetries of the models usually taken to be the best approximation to the large-scale structure of spacetime, the *Friedmann-Lemaître-Robertson-Walker (FLRW) spacetimes*, in order to define an objectively preferred frame of simultaneity, a privileged way of foliating the four-dimensional spacetime into space and time. The objective time elapsed between departure and arrival would be the time lapse according to this cosmologically privileged frame. Either of these proposals would allow us to assign an objective meaning to Lewis’s temporal separation between arrival and departure. But the resulting definition of time travel is far too promiscuous. On the first proposal, everyone who departs from geodesic motion – due to the slightest nudge from a non-gravitational force – counts as a time traveler, and on the second proposal everyone who moves with respect to the cosmologically privileged frame earns the distinction. Just imagine: even if the earth didn’t move with respect to the privileged frame, you would be time-traveling each time you go to the fridge. Admittedly, Lewis’s definition does seem to capture an intuitive sense of “time travel” that is useful for some purposes. But it is too broad to capture a useful distinction within relativity, given that nearly every observer would qualify as a time-traveler.

Thankfully, an alternative conception of time travel that avoids these problems is close at hand in GR. There is a sense in which GR permits time travel into the past: it allows spacetimes containing closed timelike curves (CTCs), i.e. spacetimes with unusual causal structures. Loosely speaking, a CTC is a path in space and time that can be carved out by a material object and is *closed*, i.e. returns to its starting point not just in space, but also in time. A curve is *everywhere timelike*, or simply *timelike*, if the tangent vectors to the curve are timelike at each point of the curve. A timelike curve represents a possible spatio-temporal path carved out by material objects, a so-called *worldline*. Of course, we also presuppose that the curves representing observers are continuous. GR declares as compatible with its basic dynamical equations spacetimes that contain closed worldlines that could be instantiated by material objects. It is evident that the presence of worldlines that intersect themselves is a sufficient condition for time travel to take place. For the rest of this essay, we shall also assume that it is a necessary condition.²

Both the popular and the philosophical time travel literature contain vivid debates regarding whether time travel in this sense is logically impossible, conceptually or metaphysically incoherent, or at least improbable. Let us address these three issues in turn.

Although less prevalent than a decade or two ago, the belief that various paradoxes establish the logical or metaphysical impossibility of time travel is still widespread in philosophy. The grandfather paradox introduced above is no doubt the most prominent of these paradoxes. It allegedly illustrates either how time travel implies an inconsistent past and is thus ruled out by logic,³ or that time travel is extremely improbable. Other time travel paradoxes include the so-called *predestination*

²*Pace* Monton (2009). We believe that Monton’s argument fails. If one rules out discontinuous worldlines and similarly unphysical constructs, then CTCs are arguably the only Lorentz-invariant way of implementing time travel.

³In a *dialectic logic*, i.e. a logic in which contradictions can be true, and perhaps in other paraconsistent logics, such contradiction need not imply the impossibility of time travel. A possible reply to the grandfather paradox is thus the rejection of classical logic. This price is considered too high in this article, particularly also because the contradiction can be resolved by other means, as will be argued shortly.

and *ontological* paradoxes. A paradox of predestination arises when the protagonist brings about an event exactly by trying to prevent it. These paradoxes are not confined to scenarios involving time travel, although they add to the entertainment value of the latter. Just imagine a time traveller traveling into her own past in an attempt to prevent the conception of her father, whose actions instead kindle the romance between her grandparents. The related *ontological paradox* can be exemplified by the story of the unpainted painting. One day, an older version of myself knocks on my door, presenting a wonderful painting to me. I keep the tableau until I have saved enough money to be able to afford a time machine. I then use the time machine to travel back in time to revisit my younger self, taking the painting along. I ring the doorbell of my earlier apartment, and deliver the painting to my younger self. Who has painted the picture? It seems as if nobody did since there is no cause of the painting. All the events on the CTC have just the sort of garden-variety causes as events not transpiring on CTCs do. The causal loop *as a whole*, however, does not seem to have an originating cause. For all these reasons, the popular argument goes, causal loops cannot exist.

Lewis (1976) has argued that although such scenarios contravene our causal intuitions, it is not in principle impossible that uncaused and thus unexplainable events in fact occur. According to Lewis, there are such unexplainable events or facts such as the existence of God, the big bang, or the decay of a tritium atom. True. Who would have expected that time travel scenarios will be easily reconcilable with our causal intuitions anyway? The fact that phenomena transpiring in a time-travel universe violate our causal intuitions, however, is no proof of the impossibility of such a world. Analogously, predestination paradoxes can be rejected as grounds for believing that time travel is impossible: although they undoubtedly exude irony, the very fact that it was the time traveller who enabled her grandparents' union is not in any way logically problematic. What is important as far as logic is concerned is that the time traveller has timelessly been conceived at some point during the year before her birth and has not been "added" or "removed" later. If it occurred, it occurred; if it didn't, it didn't. So despite their persuasiveness, the ontological and the predestination paradoxes don't go far in ruling out time travel.

The grandfather paradox cannot be dismissed so easily. Grandpa cannot simultaneously sire and not sire the parent of the time traveller. The central point is that the grandfather paradox does not rule out time travel *simpliciter*, but only inconsistent scenarios. In fact, all self-contradictory scenarios are forbidden, regardless of whether they involve time travel or not. Various options can be pursued in attempts to resolve the grandfather paradox. Apart from the costly rejection of bivalent logic, one can, following Jack Meiland (1974), postulate a two-dimensional model of time such that every moment entertains its own past which is distinct from the times that preceded that moment. According to this proposal, at a given moment there are two branches, one containing the actual events that preceded it, and the other representing an alternative past into which time travel can lead. If one travels back in time, then, one doesn't arrive at a time that preceded the departure, but rather at a time in the past of the moment when one departed. Time, on this understanding, is represented by a two-dimensional plane rather than a one-dimensional line. As Lewis (1976, 68), we do not find this resolution particularly attractive, primarily because the time traveller would on this conception never be able to revisit the very past moment when Grandpa first met Grandma. She would only be able to reach a "copy" of this moment on the past line of the moment of when the time machine is switched on. The event reached would thus be different from the one steeped in history that the intrepid traveller intended as the goal of her journey. Whatever travel this is, it is not the time travel characterized above.

An obvious, but rarely seriously entertained option tries to make sense of time travel by allowing the universe to bifurcate each time consistency would otherwise be violated. The instant the time traveller arrives in her past, the spacetime splits into two "sheets." (Unlike Meiland's proposal the

branches are “created” by time travel, they are not already in existence.)⁴ This branching does not happen in time or space alone, but in the overall causal structure of the spacetime in which the journey takes place. In particular, the causal future of the event where the traveller arrives must permit “two-valuedness.” In the case of such a “multiverse,” the adventurous traveller not only journeys in time, but also to a branch distinct from the one in which she departed. A multiverse with more than one actual past history does timelessly contain the killing of Grandfather, but only in one of the branches.⁵ Interaction between the co-existing branches is solely possible by time travel, which does arguably not deserve to qualify for time travel as it is not a journey back in the traveller’s “own” time. But the threat of inconsistency is surely banned if history along any given branch is consistent. This would for example mean that everybody’s worldlines have an unambiguous beginning and end points in all branches (see Fig. 1).

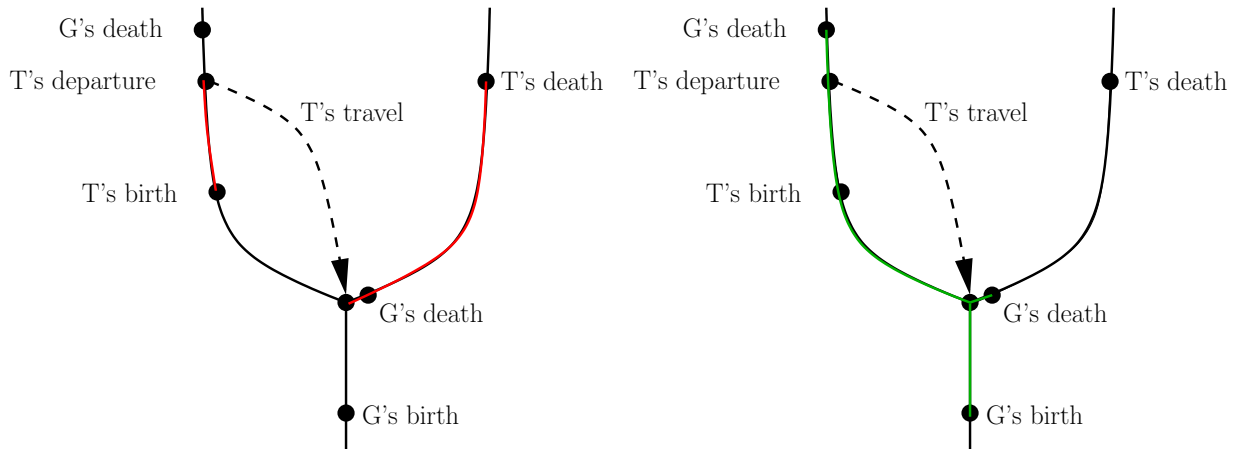


Figure 1: The worldlines of the time traveller T (red) and of Grandpa G (green) according to the multiverse proposal. Note that both figures are of the same multiverse; they are just highlighting different worldlines.

Does such travel in a multiverse change the past? Only in the sense that through the traveling activity, more and more branches of past histories seem to pop into existence. If this is the picture, then time traveling necessitates an inflation of branches as it becomes more popular. But since if it is possible to change the past, we run into the same difficulties as with the grandfather paradox again, these branches must in fact eternally co-exist with the sheet we are actually living in. Thus, if time travel is physically possible, then there will be an infinitude of branches corresponding to all the possible ways in which time travel could occur. Thus, there will be an infinity of actual past histories of the multiverse timelessly containing all time traveling activity. Even though such a construction does not live up to an ideal of metaphysical austerity, logic does not preclude it. However, in order to accommodate multi-valued fields in physics—which would be necessary in such a multiverse—, a radical rewriting of the laws of physics would be. Although topology offers manifolds which could potentially deal with multi-valuedness,⁶ these new types of laws would also have to tolerate it. But we do not know of a dynamical theory which could deliver this.

⁴A further contrast between the proposals is that on Meiland’s view the time traveller will not have complete freedom as to how to affect the past since, presumably, both pasts must lead to the same present moment located at the bifurcation point. This constraint seems to be absent in scenarios with branching structures into the future, at least if one grants the causal fork asymmetry (cf. Horwich 1987, 97-99).

⁵Lewis (1976, 80) briefly outlines this therapy.

⁶Cf. Visser (1996, 250-255). The concerned manifolds have to be non-Hausdorff in order to permit branching, as discussed in Douglas (1997). For a thorough critique of branching spacetimes, cf. Earman (2008).

We concur with John Earman (1995) (and, unsurprisingly, with Earman et al. (2009)) that the grandfather paradox only illustrates the fact that time-travel stories, just like any other story, must satisfy certain consistency constraints (CCs) that ensure the absence of contradictions. In other words, only one history of the universe is to be told, and this history had better be consistent. GR mandates that spacetimes satisfy what Earman dubbed a *global-to-local property*, i.e. if a set of tensor fields satisfy the laws of GR globally on the entire spacetime, then they do so locally in every region of spacetime.⁷ This property is shared by spacetimes with CTCs. The reverse *local-to-global property* would imply that any local solution could be extended to a global solution of the field equations. But this property need not hold in spacetimes with CTCs: situations that are admissible according to the local dynamical laws may lead to inconsistencies when evolved through a region containing CTCs. CCs are imposed to prevent such inconsistencies. Friedman et al. (1990) encode the demand that CCs are operative in their *principle of self-consistency*, which “states that *the only solutions to the laws of physics that can occur locally in the real Universe are those which are globally self-consistent.*”⁸ This principle guarantees the validity of the local-to-global property, at the cost of introducing non-trivial CCs.

How should we think of CCs? We can think of them as consisting of restrictions imposed on the initial data of, say, a matter field for point mass particles at a given point. Assume a single particle that moves along an inertial worldline in accordance with the dynamical laws that apply for the particle and assume further that the worldline is a CTC. The CCs would then have to restrict the choice of the initial velocity of the particle such that its trajectory smoothly joins itself after one loop. More generally, however, the CCs for any macroscopic object involving more complex physical processes would become very complicated indeed if spelt out explicitly. Consider a more concrete example involving macroscopic objects, such as a spacecraft venturing out to explore deep space only to discover that it in fact traces out a CTC. Here, the spacecraft would have to go over into its earlier self smoothly, including restoring the “original” engine temperature and settings of all board computers, refuelling to exactly the same amount of propellant, and so on and so forth. If the scenario included humans, it would become trickier still. The time traveller would have to rejoin exactly his worldline, wearing the same clothes, with the same shave, with each hair precisely in the same position, with his heart beat cycle exactly coinciding, his memory reset to the state when he entered the CTC etc. The world is rich in variety and complexity, and surely such strong constraints appear to be in conflict with our experience. This suggests that CTCs can only occur in very limited regions, if at all. Often, these regions of causality violations are found beyond horizons of epistemic accessibility of an earth-bound observer in realistic spacetimes. Hence, if taken as an objection against the possibility of CTCs, the difficulty of accommodating complex scenarios has little theoretical force. But it surely shatters the prospect of sending humans on a journey into their own past.

Since CCs seem to mandate what time travellers can and cannot do once they have arrived in their own past, the CCs’ insistence that there is only one past and that this past cannot be changed appears to give rise to a kind of *modal paradox*. Either John Connor’s mother is killed in 1984 or she isn’t. In case she survives, the deadliest Terminator with the highest firepower cannot successfully assassinate her. This inability stands in a stark contrast to the homicidal capacities that we would normally ascribe to an armed and highly trained cyborg. The modal paradox arises because the terminator can strike down Connor’s mother—he has the requisite weapons, training of many years, and a meticulous plan, etc.—but simultaneously he cannot do it as Sarah Connor

⁷Cf. Earman (1995, 173) for a more mathematically rigorous account.

⁸Friedman et al. (1990, 1916f), emphasis in original. For more advocacy of CCs, see Malament (1985b, 98f) and Earman (1995, passim). They both see the emergence of CCs as the one and only lesson to be learnt from the grandfather paradox.

actually survived 1984 and the Terminator would thus violate CCs were he to successfully kill her. Lewis (1976) has resolved the looming modal inconsistency by arguing that ‘can’ is ambivalently used here and that the contradiction only arises as a result of a impermissible equivocation. ‘Can’ is always relative to a set of facts. If the set contains the fact that Sarah has survived 1984, then the terminator will not be able kill her (in that year). If this fact is not included, however, then of course he can. The contradiction is only apparent and Lewis concludes that time travel into one’s own past is not logically impossible.

Thus, the paradoxes invoked do not establish that logic precludes time travel, although they exhibit how they constrain the sort of scenarios that can occur. Although logic does not prohibit it, time travel still faces stiff resistance from many philosophers. The resistance typically comes in one of two flavours: either it turns on the alleged improbability of time travel or on an argument barring the possibility of backward-in-time causation. Let us address both complaints in turn.

The first philosophical objection originated in Paul Horwich (1987) and was arguably given the most succinct expression in Frank Arntzenius (2006). This popular objection admits that the logical arguments fail to establish the *impossibility* of time travel, but insists that what they show is its *improbability*. The main reason for this improbability is the iron grip that consistency constraints exert on possible scenarios involving time travel. Imagine the well-equipped, highly trained, lethally determined Terminator who consistently fails to kill Connor’s mother again and again, in ever more contrived ways. The first time the trigger is jammed, another time a rare software bug occurs, the third time the Terminator slides on banana peel, the fourth time Reese shows up in time to spectacularly save her life, etc. The more often the Terminator attempts the murder and fails, the more improbable the ever more finely tuned explanations of his failure appear to us.

Apart from philosophical objections that have been raised against this argument,⁹ which do not concern us here, this line of reasoning also faces difficulties both concerning the exact formulation and interpretation of the probabilistic thesis to be defended as well as the interpretation of the consistency constraints. What exactly could be meant by the claim that scenarios involving CTCs are “improbable”? If the relevant probabilities should be interpreted as objective, then we should be in the position to define an adequate event space with a principled, well-defined measure. If we were handed such a space, we could then proceed to isolate the subspace of those events which include time travel. But this is hardly possible in a principled and fully general manner. If the generality is restricted to GR, the event space would arguably consist of generally-relativistic spacetime models. In this context, the claim that time travel is highly “improbable” could then be given concrete meaning by identifying the models of GR that contain CTCs and then arguing that these models-cum-CTCs represent only a very small fraction, perhaps of measure zero, of all the physically possible or physically realistic models of GR in the sense of some natural measure defined on the space of all generally-relativistic models. Such a construction would allow us to conclude that CTCs are indeed very rare in physically realistic set-ups. It would, however, presuppose significantly more knowledge about the space of all solutions to the Einstein field equations than we currently have. Although there is some exploratory material in this direction in the literature, it is certainly neither in Horwich (1987) nor in Arntzenius (2006). The latter invites us to consider the issue in the context of spacetimes where “later” CTCs may impose bizarre constraints on data in “earlier” regions of spacetime. We will return to the problem of finding objective probabilities in the setting of a spacetime theory in Section 5.

Alternatively, the probabilities involved can be interpreted as subjective. Of course, it is plausible that systematically tested betting behaviour of maximally rational human agents would ascribe

⁹Cf. e.g. Smith (1997).

a low probability to time-travel scenarios. But since we are unfamiliar with such phenomena, our untutored intuitions will not serve as reliable indicators in situations with CTCs. Our resulting betting behaviour may thus yield nothing more than additional confirmation that time-travel scenarios appear bizarre to us. This is a potentially interesting psychological point, but hardly qualifies as a serious statement about the possibility or probability of time travel. So either way, regardless of whether we interpret the probabilities involved objectively or subjectively, we are at a loss. Either we propose a novel interpretation of probability that does not run into these difficulties, or we settle for a non-probabilistic likelihood in the claim that time travel is “improbable.”

A second difficulty for Horwich’s argument arises when we regard the consistency constraints as laws of nature, as we can with some justification. Whether or not consistency constraints really are laws will undoubtedly depend upon our analysis of laws of nature, as Earman (1995) has argued. In case the consistency constraints turn out to qualify as laws, it would be amiss to infer the improbability of time travel from their existence. After all, we do not conclude from the fact that Newton’s law of universal gravitation is (at least approximately) a law of nature that a long sequence of bodies that have been released near the surface of the Earth and moved *toward* the Earth, rather than, say, away from it, is highly improbable. Of course, we can extend the set of relevant possible worlds, or of possible worlds simpliciter, in such a way that despite the nomological status of Newton’s law of gravitation in the actual world, there exists a subset of worlds in which objects move away from the centre of the Earth and that this subset is much larger than the subset of worlds in which objects fly toward Earth’s centre. Such an extension, however, would involve such modal luxations that the argument would lose much of its force. Of course, all of this would in no way imply that time travel must occur with a fairly high probability, but only that from the necessity of consistency constraints, one cannot infer its improbability.

One man’s *modus ponens* is another woman’s *modus tollens*, a defender of the improbability might object. Thus, she might simply point out that these considerations drawing on the nomological status of consistency constraints only shows, if anything, that we ought not to think of them as fully deserving of promotion to lawhood. The move, thus, is to exactly deny the analogy to the case of Newton’s law of universal gravitation. True, our point only commands any force *if* consistency constraints are awarded nomological status. But if we have independent reasons to so award them, then nothing about the probability or improbability of time travel can be inferred from the fact that consistency constraints obtain. If we are right, then Horwich’s claim must at least be strongly qualified or amended.

The second philosophical motivation for resisting time travel maintains that the presumed impossibility of backward causation disallows time travel. While most philosophers today would accept that time travel is not ruled out on grounds of paradoxes of consistency, they argue that it necessarily involves backward causation and since backward causation is conceptually impossible, they believe, so is time travel. Both premises of this argument, however, can defensibly be rejected. The reason why time travel as we conceive it does not involve backward causation, at least not locally, will become clear in the next section when we sharpen the concepts by reconfiguring the issues in the context of spacetime physics. Most philosophers, however, beg to differ as they insist that although a time traveller cannot *change* her past, she must still *affect* it. A causal link between antecedent conditions prior to departure and consequent conditions upon the earlier arrival is required to ascertain the personal identity and thus the persistence of the time traveller. But if such a causal relation is necessary for genuine time travel, they argue, there must be backward causation, i.e. causal relations where the effect precedes the cause. This argument, however, does not succeed if causation is conceptualized as a purely local phenomenon, connecting only events of adjacent spacetime regions or, more precisely, if causal propagation occurs only along smooth curves in spacetime. We will return to this in §3. The second premise—that backward causation is

incoherent or impossible—is genuinely metaphysical and shall be dealt with here. We do not wish to commit ourselves to the metaphysical possibility of backward causation here, but we want to elucidate the dialectical grounds on which backward causation is ruled out of bounds.

The basic idea behind the ruling is captured by the so-called *bilking argument*.¹⁰ Consider the following experimental “set-up.” We plan an experiment in which we attempt to produce (prevent) the subsequent cause C whenever we have previously observed the absence (presence) of the potential effect E . The experiment is then performed repeatedly, with sufficient runs to gain statistically meaningful results. It will be found that either one of two possibilities obtains: Either (i) E often transpires despite the absence of C , and the occurrence of C is often unaccompanied by E ; or (ii) whenever E does not occur, our attempts to bring about C consistently fail, and whenever E occurs, we cannot prevent the subsequent occurrence of C . In case (ii), our ability to produce C depends upon the previous presence of E . Advocates of the bilking argument insist that in this situation, we ought to interpret E as a necessary causal antecedent condition for the occurrence of C , rather than a consequent condition. The causal relation between E and C obtains, but should be taken to hold in the opposite direction: E is the cause, not the effect, of C . In case (i), on the other hand, the hypothesis of backward causation is simply false, as the factors do not stand in any causal relation. In either of the two cases, we don’t have backward causation.

Does the bilking argument rule out the possibility of backward causation? It is not so clear that it does. First, the experimental design may not be implementable in all situations of potential backward causation. Price (1984) has argued that while it seems reasonable to assume that backward causation is impossible in instances that could in fact be bilked, there is no guarantee that this is always possible. The conditions for bilking fail to obtain, for instance, if we cannot discover whether a supposed earlier effect has in fact occurred. Price (1984) has forcefully argued that this may indeed be the case for quantum-physical systems when the observation of the candidate effect can only be achieved at the price of disturbing the system in such a way that is itself causally relevant for the occurrence of E . Alternatively, it is at least conceivable that it may be nomologically impossible to determine whether E has indeed occurred prior to C . Second, even if the experimental design can be implemented in all relevant situations, the argument may not succeed because bilking may be frustrated not by a reversed causal relation between E and C , but simply by cosmic coincidences. We can run very long experimental series in an attempt to minimize the probability of such coincidences, but we never seem able to fully rule out their possibility.

Without delving into these matters here, we conclude that the case against backward causation is not closed. We would like to repeat, however, that the next section will show how the attack against the possibility of time travel based on the bilking argument against backward causation amounts to an *ignoratio elenchi*, i.e. misses the point, because there is a perfectly respectable sense in which time travel as it arises in the context of spacetime theories does not involve backward causation.

We conclude the section by admitting that the discussed paradoxes reveal that unusual consistency constraints are indeed required in time travel spacetimes. We do not, however, grant that the presence of such consistency constraints indicates the “improbability” of time travel scenarios or even offers the basis for rejecting time travel altogether. Most importantly, we conclude that time travel is neither logically nor metaphysically impossible, as all arguments attempting to establish inconsistency have failed and the philosophical considerations adduced against the possibility or probability of time travel are inconclusive at best. But even if logic and metaphysics do not rule

¹⁰For a classic formulation, cf. Mellor (1981). The bilking argument has first been formulated by Black (1956). Dummett (1964) reacts to Black, defending backward causation. The argument as presented here is grossly simplified and neglects additional factors such as the informational state of the epistemic agents involved, the possibly statistical nature of certain causal relations, etc.

out time travel, physics might. Let us thus now turn to the question of whether physics permits time travel.

3 Causal Structure of Relativistic Spacetimes

Formulating our question – whether physics permits time travel – precisely requires an understanding of the causal structure of relativistic spacetimes. The study of causal structure developed as physicists freed themselves from the study of particular solutions and began to prove theorems that would hold generically, most importantly the singularity theorems. The local constraints imposed by the dynamical laws of GR are compatible with solutions with a wide variety of different global properties, which are a rich source of counterexamples to proposed theorems. The ideas surveyed briefly in this section make it possible to characterize solutions based on their “large-scale” or global features, and to bar counter-examples by imposing conditions on the causal structure. We will see that the causality conditions can be thought of as specifying the extent to which a spacetime deviates from the causal structure of Minkowski spacetime.

We begin by reviewing definitions.¹¹ A *general-relativistic spacetime* is an ordered pair $\langle \mathcal{M}, g_{ab} \rangle$, where \mathcal{M} is a connected four-dimensional differentiable manifold without boundary, and g_{ab} is a Lorentz-signature metric defined everywhere on \mathcal{M} , such that g_{ab} satisfies Einstein’s field equations $R_{ab} - \frac{1}{2}Rg_{ab} + \Lambda g_{ab} = 8\pi T_{ab}$ for some energy-momentum tensor T_{ab} .¹² The metric g_{ab} fixes a light cone structure in the tangent space \mathcal{M}_p at each point $p \in \mathcal{M}$. A tangent vector $\xi^a \in \mathcal{M}_p$ is classified as timelike, null, or spacelike, according to whether $g_{ab}\xi^a\xi^b > 0, = 0$, or < 0 (respectively).¹³ Geometrically, the null vectors “form the light cone” in the tangent space with the timelike vectors lying inside and the spacelike vectors lying outside the cone. The classification of tangent vectors extends naturally to curves: a *timelike curve* is a continuous map of an interval of \mathbb{R} into \mathcal{M} , such that its tangent vector is everywhere timelike. A spacetime is *time orientable* if and only if there exists a continuous, nowhere-vanishing vector field, which makes a globally consistent designation of one lobe of the null cone at every point as “future” (in the which-is-which sense) possible. In a time-orientable spacetime, one can then define *future (past)-directed* timelike curves as those whose tangent vectors fall in the future (past) lobe of the light cone at each point. Future-directed *causal* curves have tangent vectors that fall *on or inside* the future lobe of the light cone at each point.

We can now precisely characterize the set of points in \mathcal{M} that are causally connected to a given point $p \in \mathcal{M}$. The *chronological future* $I^+(p)$ is defined as the set of all points in \mathcal{M} that can be reached from p by (non-trivial) future-directed timelike curves; the *causal future* $J^+(p)$ includes all points that can be reached by a future-directed causal curve. (The past sets $I^-(p)$ and $J^-(p)$ can be defined analogously, and the definitions extend straightforwardly to sets $S \in \mathcal{M}$ rather than points.) More intuitively, $I^+(p)$ includes the points in \mathcal{M} that can be reached by a “signal” emitted at the point p traveling below the speed of light. (The causal future $J^+(p)$ includes, in addition to those in $I^+(p)$, points connected to p by null geodesics; although light signals propagate along

¹¹We can recommend nothing better than Geroch and Horowitz (1979) for a clear and self-contained introduction to the global structure of relativistic spacetimes; see also Hawking and Ellis (1973) or Wald (1984) for textbook treatments, or Earman (1995) for a more philosophically oriented discussion.

¹²The *energy-momentum tensor* T_{ab} is a functional of the matter fields, their covariant derivatives, and the metric g_{ab} that satisfies the following conditions: (i) T_{ab} is a symmetric tensor, (ii) $T_{ab} = 0$ in some open region $\mathcal{U} \subset \mathcal{M}$ just in case the matter fields vanish in \mathcal{U} , and (iii) $T_{;b}^{ab} = 0$. Our discussion here is “kinematical,” in the sense that we do not yet impose any further constraints on the stress-energy tensor, such as the energy conditions discussed in §8 below. The symmetric tensor R_{ab} , called the *Riemann tensor*, is a functional of the metric, R is the *Riemann scalar*, and Λ is the cosmological constant. Throughout this article, we use natural units, i.e. $c = G = 1$.

¹³In this article we use a (1,3) signature, which means that the metric assigns a length +1 to one of the four orthonormal basis vectors that can be defined in each tangent space, and -1 to the other three.

such curves, using $J^\pm(p)$ rather than $I^\pm(p)$ in what follows leads to unnecessary complications.) The boundary of the chronological future, $\dot{I}^+(S)$, is the set of points lying in the closure of $I^+(S)$ but not its interior.¹⁴ It is easy to see that $\dot{I}^+(S)$ must consist of null surfaces (apart from the set S itself).¹⁵ In fact, $\dot{I}^+(S)$ is an achronal boundary, i.e. no point can be connected to another point in the surface via a timelike curve. Any given point $q \in \dot{I}^+(S)$ lies either in the closure of S or in a null geodesic lying in the boundary, called a generator of the boundary. In Minkowski spacetime, the generators of the boundary can have *future* endpoints on the boundary, beyond which they pass into $I^+(S)$, but their *past* endpoints all lie on the set S . However, this does not hold in general, given the possibility that null geodesics making up part of the boundary may encounter “missing points” rather than reaching S (as illustrated in Figure 2).

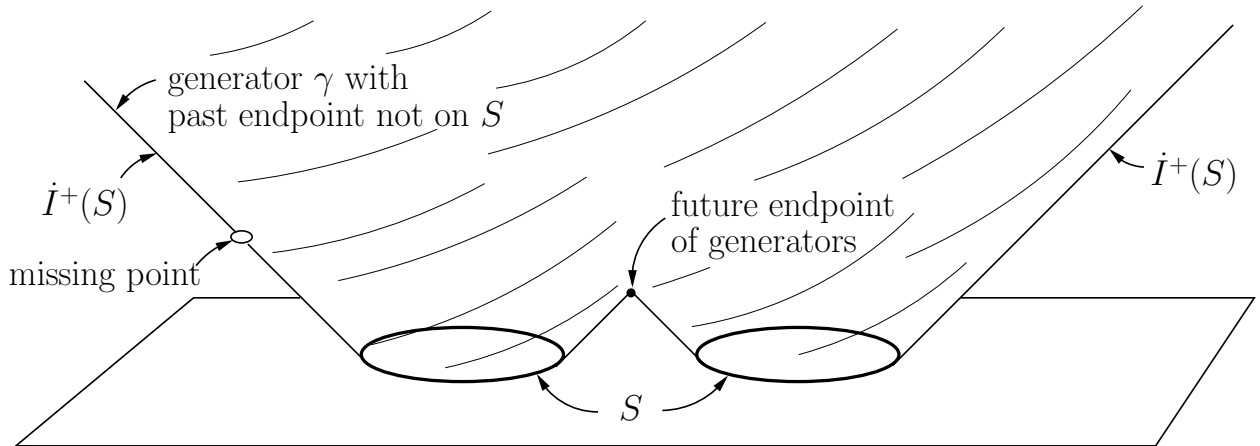


Figure 2: The generators of the boundary may have endpoints not in S .

More generally, the properties of the sets $I^\pm(p)$ in Minkowski spacetime hold locally in any general-relativistic spacetime, but to insure that they also hold *globally* one must impose further “causality conditions.” These conditions form a hierarchy in terms of strength. The lowest condition in the hierarchy, *chronology*, rules out the existence of CTCs. The existence of a CTC passing through a point p implies that $p \in I^+(p)$; chronology requires that there are no such points. One can construct “artificial” examples of relativistic spacetimes that fail this condition, such as the spacetime defined by “rolling up” 1 + 1 dimensional Minkowski spacetime (see figure ??). Although this example is obviously quite artificial, there are a number of more interesting chronology-violating spacetimes, such as Gödel spacetime (discussed in the next section).¹⁶ Even this simple example serves to illustrate the point that *locally* chronology violating spacetimes do not involve backward causation or anything of the sort. The description of physical processes in a region of rolled-up Minkowski spacetime will not differ from that in ordinary Minkowski spacetime, except that the presence of CTCs may lead to consistency constraints (we return to this topic below). Chronology is a weak condition and does not imply other properties that one might hope that the chronological

¹⁴The closure of a set S , denoted \bar{S} is the smallest closed set containing S , whereas the interior is the largest open set contained in S .

¹⁵Suppose the boundary is timelike. For an arbitrary point p just outside the boundary $\dot{I}^+(S)$, there will be points in the interior of $I^+(S)$ that can be connected to it by a timelike curve – and hence p lies in $I^+(S)$, contrary to our assumption. Similarly, suppose $\dot{I}^+(S)$ is spacelike and consider a point p lying to the future of the boundary. But there are timelike curves connecting p to points in $I^+(S)$, again contradicting the assumption. For further information regarding the properties of the boundary, see, e.g., Hawking and Ellis (1973, Prop. 6.3.1).

¹⁶Discussions of causal structure often use simple “artificial” constructions, based on the conviction that the features they illustrate show up in more complicated guise in spacetimes that are more physically reasonable.

future and past satisfy. For example, Minkowski spacetime is *past* and *future distinguishing* in the sense that different points have different chronological pasts and futures (that is, for every $p, q \in \mathcal{M}$, $I^-(p) = I^-(q) \rightarrow p = q$ and $I^+(p) = I^+(q) \rightarrow p = q$). It is easy to construct spacetimes in which chronology holds that fail to be past and future distinguishing, and it is also the case that the disjunction (past *or* future distinguishing) does not imply the conjunction.

Before moving up the hierarchy, consider the following question: what conditions do we need to impose in order to reconstruct a spacetime based on causal relations encoded in the sets $I^\pm(p)$ for all $p \in \mathcal{M}$? In other words, given two spacetimes $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}', g'_{ab} \rangle$ and a mapping between them $\phi : \mathcal{M} \rightarrow \mathcal{M}'$ that preserves the causal structure, in that $q \in I^+(p) \leftrightarrow \phi(q) \in I^+(\phi(p))$, what can we further claim regarding their full spacetime structures – including topology and geometry? Intuitively, if the causality conditions fail then requiring that such a map exists is less informative; for example, one can define such a map between two chronology-violating spacetimes with quite different topological and geometric structures.¹⁷ But at what point on the hierarchy do the sets $I^\pm(p)$ encode sufficiently rich information about the spacetime structure to make it possible to reconstruct $\langle \mathcal{M}, g_{ab} \rangle$ from $\langle \mathcal{M}', g'_{ab} \rangle$? Malament (1977) proved that if the spacetimes are past *and* future distinguishing, then $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}', g'_{ab} \rangle$ have the same geometrical structure up to a conformal factor.¹⁸ The import of this result is that any philosophical program to give a reductive analysis of spacetime structure in terms of causal relations has a fighting chance only in past and future distinguishing spacetimes.

At the top of the hierarchy of causality conditions we find *global hyperbolicity*. Define the *future domain of dependence* $D^+(\Sigma)$ for a global time slice Σ to be the set of points in \mathcal{M} such that every past inextendible causal curve through p intersects Σ (with the obvious analogous definition for $D^-(\Sigma)$). The “slice” Σ is a spacelike hypersurfaces with no edges. (The *edge* of an achronal surface S is the set of points p such that every open neighborhood O of p includes points in $I^+(p)$ and $I^-(p)$ that can be connected by a timelike curve that does not cross S .) Thus, a global time slice Σ is an inextendible, smooth spacelike hypersurface which trisects \mathcal{M} : Σ itself, as well as the “past” and the “future” of Σ . The future boundary of $D^+(\Sigma)$ is called the *future Cauchy horizon* $H^+(\Sigma)$.¹⁹ The *domain of dependence* $D(\Sigma)$ is the union of the past and future domains of dependence. Global hyperbolicity requires that the spacetime possesses a *Cauchy surface*, that is a slice Σ such that $D(\Sigma)$ is the entire spacetime.²⁰ Global hyperbolicity implies that the manifold \mathcal{M} is topologically $\Sigma \times \mathbb{R}$. It also insures that the generators of the boundary $\dot{I}^+(S)$ share the properties of boundaries in Minkowski spacetime, in particular that $\dot{I}^+(S)$ consists of the future-directed null geodesics with past endpoints on S – it rules out the possibility that there are incomplete null geodesics lying in the boundary of the chronological future (and past). A spacetime which possesses a Cauchy surface is safe for determinism, in that initial data specified on the Cauchy surface determine a unique solution to Einstein’s field equations throughout the spacetime. Furthermore, there are

¹⁷Given that the $I^\pm(p)$ are open sets, for a suitably well-behaved spacetime these can be used to define a topology (called the *Alexandrov topology*) equivalent to that of the manifold topology. In the presence of closed timelike curves or other acausalities, a topology defined using $I^\pm(p)$ is too coarse, lacking open sets corresponding to every open set of the manifold topology. For example, there is a trivial map preserving causal structure between Gödel spacetime (discussed below) and a four-dimensional “rolled up” Minkowski spacetime (since in both cases $\forall p(I^\pm(p) = \mathcal{M})$), despite their quite different topological and metrical structures.

¹⁸Two spacetimes $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}', g'_{ab} \rangle$ are said to have the *same geometrical structure up to a conformal factor* just in case $g_{ab} = \Omega^2 g'_{ab}$ where Ω is a smooth, strictly positive (and thus non-zero) function. A conformal transformation induced by Ω preserves local angles and ratios of magnitudes. Thus, the local light cone structure is preserved under conformal transformation.

¹⁹The *future boundary* of $D^+(\Sigma)$ is defined as $\overline{D^+(\Sigma)} - I^-(D^+(\Sigma))$, where the overbar denotes topological closure.

²⁰There are several different ways of defining global hyperbolicity that are provably equivalent; see, e.g., Hawking and Ellis (1973, pp. 206-212) for further discussion.

typically existence and uniqueness theorems showing that data specified on a partial Cauchy surface Σ uniquely determines a solution throughout the domain of dependence $D(\Sigma)$ for matter fields coupled to Einstein’s field equations. (See §5 for further discussion of the initial value problem in GR.)

We will be interested below in *chronology violating* spacetimes, i.e. spacetimes which contain CTCs. In some cases, such as the “rolled up” Minkowski cylinder mentioned above, a CTC passes through *every* point in the spacetime. But this is not generally the case, and we can define the *chronology violating region* $\mathcal{V} \subset \mathcal{M}$ as including all points $p \in \mathcal{M}$ such that a CTC passes through p ; if $\mathcal{V} \neq \emptyset$ then it is an open region. The local-to-global property introduced in §2 can now be expressed more rigorously: A general-relativistic spacetime $\langle \mathcal{M}, g_{ab} \rangle$ has the *local-to-global property* just in case for any open neighbourhood $\mathcal{O} \subset \mathcal{M}$, if $\langle \mathcal{O}, g_{ab}|_{\mathcal{O}} \rangle$ solves Einstein’s field equations for some (admissible) energy-momentum tensor $T_{ab}|_{\mathcal{O}}$, then $\langle \mathcal{M}, g_{ab} \rangle$ does as well for a T_{ab} that is identical to $T_{ab}|_{\mathcal{O}}$ in \mathcal{O} . Chronology violating spacetimes typically do not exemplify the local-to-global property. In other words, locally well-behaved solutions to Einstein’s field equation (perhaps coupled to some other dynamical equations) are not in general extendable to global solutions on all of \mathcal{M} . This is not surprising, of course, since in chronology-violating spacetimes there will in general be local solutions that do not satisfy the consistency constraints. However, it is quite surprising that there are several chronology-violating spacetimes that *do* exhibit the local-to-global property for some dynamical equations. The discovery of several systems with this property spurred interest in this spacetimes with CTCs, although there is still not a general characterization of which chronology-violating spacetimes admit well-posed initial value problems for an interesting set of dynamical equations.²¹

The causality conditions make it possible to classify spacetimes according to how much their causal structure diverges from various natural features of Minkowski spacetime, and hence to prove theorems for “well-behaved” spacetimes. The conditions themselves are not consequences of the dynamics – we have already mentioned simple cases of chronology violating spacetimes and we will see more interesting cases below. We have argued against the idea that something akin to causality conditions should be imposed as a further law of nature in order to avoid alleged time travel paradoxes. But granting this point obviously does not settle the question of what the existence of chronology violating spacetimes implies regarding the nature of time in GR. One response to this question, widespread in the physics literature until fairly recently, has been to simply dismiss chronology violating spacetimes as mathematical curiosities without physical relevance. Advocates of this line of thought are faced with the task of articulating a clear set of constraints on what qualifies as a “physically reasonable” spacetime such that all the chronology violating spacetimes are ruled out, a task that turns out to be surprisingly difficult (as we will see in §6 below). Admittedly, we are hard pressed to produce an example of a chronology violating spacetime whose success in modeling a particular physical phenomena depends upon the presence of CTCs. But chronology violating spacetimes need not be viable models of observed phenomena in order to be worthy of study, or to shed light on the conceptual structure of GR. In the next sections we turn to assessing the implications of the existence of chronology violating spacetimes.

4 Implications of Time Travel

Given that time travel cannot be straightforwardly ruled out as incoherent or logically impossible, we now face the following difficult questions: In what sense is time travel physically possible, and

²¹See Friedman 2004 for a survey of the initial value problem in spacetimes with CTCs and references to earlier results.

what does this imply regarding the nature of time? More precisely, what are the *novel* consequences of time travel, i.e. ones that do not follow already from more familiar aspects of special or general relativity? As a first step towards answering these questions, we will consider Gödel’s (in)famous argument for the ideality of time.²²

Gödel (1949) was the first to clearly describe a relativistic spacetime with CTCs.²³ Gödel’s stated aim in discovering this spacetime was to rehabilitate an argument for the ideality of time from special relativity within the context of GR. In special relativity, Gödel asserts that the ideality of time follows directly from the relativity of simultaneity. He takes as a necessary condition for the existence of an objective lapse of time the possibility of decomposing spacetime into of a sequence of “nows” – namely, that it has the structure $\mathbb{R} \times \Sigma$, where \mathbb{R} corresponds to “time” and Σ are “instants,” three-dimensional collections of simultaneous events. But in special relativity the decomposition of the spacetime into “instants” is relative to an inertial observer rather than absolute; as Gödel puts it, “Each observer has his own set of ‘nows,’ and none of these various systems of layers can claim the prerogative of representing the objective lapse of time” (Gödel 1949, 558).

This conclusion does not seem to carry over to GR, because there is a natural way to privilege one set of “nows” in a cosmological setting. The privilege can be conferred on a sequence of “nows” defined with respect to the worldlines of galaxies or other large scale structures. It is natural to require the surfaces of simultaneity to be orthogonal to the worldlines of the objects taken to define the “cosmologically preferred frame.” The question is then whether one can extend local surfaces of simultaneity satisfying this requirement to a global foliation for a given set of curves. For the FLRW cosmological models, as noted above, the answer is yes. These models have a natural foliation, a unique way of globally decomposing spacetime into a one-dimensional “cosmic time” and three-dimensional surfaces Σ representing “instants,” orthogonal to the worldlines of freely falling bodies. (Cosmic time in this case would correspond to the proper time measured by an observer at rest with respect to this privileged frame.) Thus Gödel’s necessary condition for an objective lapse of time is satisfied in the FLRW cosmological models, and in this sense the pre-relativistic concept of absolute time can be recovered.

But in Gödel’s spacetime one cannot introduce such a foliation. The spacetime represents a “rotating universe,” in which matter is in a state of uniform rigid rotation.²⁴ Due to this rotation there is no way to define a privileged frame with global “instants,” as there is in the FLRW models. An analogy due to Malament (1995) illustrates the reason for this. One can slice through a collection of parallel fibers with a single plane that is orthogonal to them all, but if the fibers are twisted into a rope there is no way to cut through the rope while remaining orthogonal to each fiber. (The “twist” of the fibers is analogous to the rotation of worldlines in Gödel’s model.) The construction of global “instants” described above can be carried out if and only if there is no “twist” (or rotation) of the worldlines used to define the cosmologically privileged frame. Demonstrating that such rotating models exist by finding an explicit spacetime model solving Einstein’s field equations was clearly

²²The following papers, which we draw on below, discuss aspects of Gödel’s argument: Stein (1970), Malament (1985b), Savitt (1994), Earman (1995), Dorato (2002), Belot (2005). Yourgau (2006) offers an accessible account of Gödel’s argument and its impact, as well as his interactions with Einstein, and Ellis (1996) discusses the impact of Gödel’s paper.

²³Although von Stockum (1937) discovered a solution describing an infinite rotating cylinder that also contains CTCs through every point, this feature of the solution was not discussed in print, to the best of our knowledge, prior to Tipler (1974). Gödel does not cite von Stockum’s work. Others had noted the possibility of the existence of CTCs without finding an exact solution exemplifying the property (see, e.g., Weyl (1921), p. 249).

²⁴More precisely, in Gödel’s universe a congruence of timelike geodesics has non-zero twist and vanishing shear. Defining rotation for extended bodies in general relativity turns out to be a surprisingly delicate matter (see, especially, Malament 2002).

Gödel’s main aim. But the welcome discovery that in his rotating universe there is a CTC passing through every point further bolstered his argument for the ideality of time.²⁵ It is noteworthy that many chronology violating spacetimes resemble Gödel’s solution in the following sense: they contain rotating masses and CTCs wind around the masses against the orientation of the rotation.²⁶

What, then, is Gödel’s argument? The crucial problem is how to get from discoveries regarding the nature of time in this *specific* spacetime to a conclusion about the nature of time *in general*. Gödel could avoid this problem if his spacetime, or a spacetime with similar features, were a viable candidate for representing the structure of the observed universe. Then his results would obviously have a bearing on the nature of time in *our* universe. Gödel apparently took this possibility quite seriously, and subsequently discovered a class of rotating models that incorporate the observed expansion of the universe (Gödel 1952). In these models, one can construct suitable “instants” as long as the rate of rotation is sufficiently low, and recent empirical work places quite low upper limits on the rate of cosmic rotation.²⁷ Gödel goes on to argue that even if his model (or models with similar features) fails to represent the actual universe, its mere existence has general implications (p. 562):²⁸

The mere compatibility with the laws of nature of worlds in which there is no distinguished absolute time, and, therefore, no objective lapse of time can exist, throws some light on the meaning of time also in those worlds in which an absolute time *can* be defined. For, if someone asserts that this absolute time is lapsing, he accepts as a consequence that, whether or not an objective lapse of time exists ... depends on the particular way in which matter and its motion are arranged in the world. This is not a straightforward contradiction; nevertheless, a philosophical view leading to such consequences can hardly be considered as satisfactory.

Despite disagreement among recent commentators regarding exactly how to read Gödel’s argument, there is consensus that even this modest conclusion is not warranted. The dynamical connection between spacetime geometry and the distribution of matter encoded in Einstein’s field equations insures that, in some sense, many claims regarding spacetime geometry depend on “how matter and its motion are arranged.” Nearly any discussion of the FLRW models highlights several questions regarding the overall shape of spacetime – e.g., whether time is bounded or unbounded and what

²⁵Malament observed that the existence of CTCs is not mentioned in three of the five preparatory manuscripts for Gödel (1949a or b), and it appears that Gödel discovered this feature in the course of studying the solution. In addition, in lecture notes on rotating universes (from 1949) Gödel emphasizes that he initially focused on rotation and its connection to the existence of global time slices in discovering the solution. See Malament (1995) and Stein (1995, 227-9).

²⁶Cf. Andréka et al. (2008). That rotation may be responsible for the formation of CTCs is also suggested by Bonnor’s (2001) result that stationary axially symmetric solutions of Einstein’s field equations describing two spinning massive bodies under certain circumstances include a non-vanishing region containing CTCs.

²⁷These instants are not surfaces orthogonal to timelike geodesics, as there is still rotation present, but Gödel (1952) establishes that surfaces of constant matter density can be used to define a foliation that satisfies his requirements for an objective lapse of time. For recent empirical limits on global rotation based on the cosmic microwave background radiation, see, for example, Kogut et al. (1997).

²⁸As Sheldon Smith pointed out to us, if this is taken to be Gödel’s main argument then it is not clear why the mere existence of Minkowski spacetime, regarded as a vacuum solution of the field equations, does not suffice. Why did Gödel need to go to the effort of discovering the rotating model granted that there is no distinguished absolute time in Minkowski spacetime? Although we do not find a clear answer to this in Gödel (1949a or b), we offer two tentative remarks. First, Gödel may have objected to classifying Minkowski spacetime as physically reasonable because it is a vacuum spacetime. Second, and more importantly, Gödel took the prospect of discovering a rotating and expanding model consistent with observations more seriously than most commentators allow. This suggests that the argument in the quoted passage is a fall-back position, and that Gödel put more weight on the claim that he had discovered a viable model for the observed universe that lacks an objective lapse of time.

is the appropriate spatial geometry for “instants” – that depend on apparently contingent properties such as the value of the average matter density. What exactly is unsatisfactory about this? What does the mere possibility of spacetimes with different geometries imply regarding geometrical structure in general? Earman (1995, Appendix to Chapter 6) challenges the implicit modal step in Gödel’s argument. How can we justify this step on Gödel’s behalf, and elucidate what is unsatisfactory about objective time lapse in general, without lapsing back into pre-GR intuitions?

Perhaps the argument relies on an implicit modal assumption that lapsing, in the sense described above, must be an essential property of time. Then (given that $\diamond(\neg P) \leftrightarrow \neg(\Box P)$), the demonstration that $\diamond(\neg P)$ (where P is the existence of an objective lapse of time) via finding the Gödel spacetime would be decisive. But what is the basis for this claim about the essential nature of time, and how can it be defended without relying on pre-relativistic intuitions? Earman (1995) considers this and several replies that might be offered on Gödel’s behalf, only to reject each one. Savitt (1994) defends a line of thought (drawing on Yourgau 1991) that is more of a variation on Gödelian themes than a textual exegesis. On Savitt’s line, Gödel’s argument rests not on essentialist claims regarding the nature of time but instead on a claim of local indistinguishability. Suppose that it is physically possible for beings like us to exist in a Gödel spacetime, and (1) that these denizens have the “same experience of time” as we do. Assume further that (2) the only basis for our claim that objective time exists in our universe is the direct experience of time. Then the existence of the Gödel universe is a defeater for our claim to have established objective time lapse on the basis of our experience, because (for all we know) we could be in the indistinguishable situation – inhabiting a Gödel universe in which there is no such lapse. While this variation does not require a modal step as suspect as the original version, neither (1) nor (2) are obviously true – and it is unclear how they can be established without begging the question.²⁹

A different approach spelled out by Belot (2005) offers a *methodological* rather than *metaphysical* response to Earman’s challenge. Belot concedes to Earman’s challenge given a “natural-historical” construal of Gödel’s argument, according to which the nature of time can be established based on empirical study of “how matter and its motion are arranged.” On this reading, time in *our* universe is characterized by the appropriate spacetime of GR that is the best model for observations – and the mere existence of alternative spacetimes is irrelevant. But on a “law-structural” construal questions regarding the nature of time focus on the laws of nature rather than on contingent features of a particular solution. Belot makes a case that a law-structural construal of the question is more progressive methodologically, in that it fosters deeper insights into our theories and aids in the development of new theories.³⁰ If we grant that understanding the laws may require study of bizarre cases such as Gödel’s spacetime alongside more realistic solutions then we have the start of a response to Earman’s challenge.

It is only a start, because this suggested reading remains somewhat sketchy without an account of “laws of nature,” which is needed to delineate the two construals more sharply. Even if we had a generally accepted account of the laws of nature, the application of “laws” to cosmology is controversial: how can we distinguish nomic necessities from contingencies in this context, granting the uniqueness of the universe? Setting this issue aside, Earman’s challenge can be reiterated by asking which spacetimes should be taken as revealing important properties of the laws. Why should Gödel spacetime, in particular, be taken to reveal something about the nature of time encoded in the laws of GR? Suppose we expect that only a subset of the spacetimes deemed physically possible within classical GR will also be physically possible according to the as-yet-undiscovered theory of

²⁹See Belot (2005) and Dorato (2002) for further discussion.

³⁰Belot finds inspiration for this position in several brief remarks regarding the nature of scientific progress in manuscript precursors to Gödel (1949); however, he does not take these considerations to be decisive (see p. 275, f.n. 52).

quantum gravity. How would we argue that Gödel spacetime should fall within that subset, and that it should be taken to reveal a fundamental feature of the laws of GR that will carry over to quantum gravity? The features Gödel used to establish the lack of absolute time in his model are often taken to support a negative answer to this question that does not appear to be *ad hoc*. Many approaches to quantum gravity simply rule out spacetimes with CTCs *ab initio* based on the technical framework adopted.³¹ As we will discuss below, much of the physics literature on spacetimes with CTCs seeks clear physical grounds to rule them beyond the pale; insight into the laws of a future theory of quantum gravity would come from showing why the laws do *not* allow CTCs. But we agree with Belot that what is more unsatisfying regarding Gödel’s argument, even on the “law-structural” construal, is that an argument by counter-example does little to illuminate deeper connections between the nature of time and the laws of the theory.

Assessing the implications of Gödel’s spacetime clearly turns on rather delicate issues regarding modality and the laws of nature. Perhaps our failure to articulate a clear Gödelian argument indicates that the properties of such bizarre spacetimes can be safely ignored when we investigate the nature of time in GR? Maudlin (2007) advocates a dismissive response to CTCs, which would otherwise pose a threat to his metaphysical account of the passage of time: “It is notable in this case that the equations [Einstein’s field equations] do not force the existence of CTCs in this sense: for any initial conditions one can specify, there is a global solution for that initial condition that does not have CTCs.” He anticipates a critic’s response that his metaphysical account of passage boldly stipulates that the nature of time is not compatible with the existence of CTCs, and replies: “...But is it not equally bold to claim insight into the nature of time that shows time travel to be possible *if we grant that it is not actual and also that the laws of physics, operating from conditions that we take to be possible, do not require it*” (Maudlin 2007, 190). These assertions would follow from the proof of the following form: CTCs do not arise from “physically possible” initial states under dynamical evolution according to Einstein’s equations. Below we will consider a more precise formulation of this “chronology protection conjecture” (in §6). But at this point we wish to emphasize that this is still a *conjecture*, and that there are a number of subtleties that come into play in even formulating a clear statement amenable to proof or disproof.³² Perhaps a claim like Maudlin’s, suitably disambiguated, will prove to be correct, but part of the interest of the question is precisely due to the interesting technical questions that remain open (see, by contrast, Stein 1970). In any case, Maudlin’s remarks usefully indicate a fruitful way of addressing the importance of solutions with exotic causal structure. Arguments by counterexample – displaying a solution to Einstein’s field equations with exotic causal structure – are unsatisfying because it is usually not clear how the solution in question relates to solutions used to model physical systems. Rather than considering such solutions, we can ask instead what Einstein’s field equations imply for the dynamical evolution of some class of initial data we decide to treat as “physically possible.” This statement of the question is closely tied to the initial value formulation of GR, as we will see below. (Note that we can only address the existence of chronology-violating spacetimes indirectly on this approach, given that they lack surfaces upon which initial data can be specified.) We would avoid controversial questions regarding modalities in cosmology, and instead focus on whether it is possible according to GR to manipulate matter and energy in a local region such that, *contra*

³¹Gödel’s solution might be ruled out due to the symmetries of the solution, as Belot notes: symmetric solutions pose technical obstacles to some approaches to quantization, and it seems precarious to base assertions regarding features of quantum gravity on properties of special, symmetric solutions. But this argument seems too strong, in that it also would rule out the FLRW models, which are currently accepted as the best classical descriptions of the large-scale structure of the universe.

³²And we should note that absent some further qualifications, Maudlin’s first claim is false – as established by Manchak’s theorem discussed in the next section.

Maudlin, CTCs are the inevitable result. In more vivid language, is it possible in principle to build a time machine? Formulating this idea precisely is the task of the next section.

5 Time Machines

In the usual parlance of science fiction authors, time machines are simply devices that enable time travel in roughly Lewis’s sense (discussed in §2) — the time elapsed during the journey does not match some appropriately external measure of the time interval between departure and arrival. We have already argued in favor of one departure from this usage, in that we define time travel in terms of the existence of CTCs. Time travel is possible in our sense only in chronology-violating spacetimes, and we will leave aside the issue of whether any observer or particle in fact instantiates time travel by following a CTC. But in a departure that will be more disappointing for science fiction fans, we will define a “time machine” as a device which produces CTCs where none would have existed otherwise. This is more demanding than merely requiring the existence of CTCs. In Gödel spacetime, for example, one cannot claim that the CTCs are produced by local manipulation of matter and energy within some finite region, as we require for the existence of a time machine. Our immediate goal is to show how, following Earman et al. (2009), to flesh out this idea in terms of spacetime geometry. Science fiction fans will be disappointed by this analysis because, as we will see, our definition rules out the possibility of using a time machine to travel into one’s own past.

Our starting point is the idea that a “time machine” is a region of spacetime in which local manipulation of the fields can produce CTCs. As a first step, we turn to the initial value formulation of GR as a way of clarifying the sense in which manipulating fields in a finite region can be said to determine their values elsewhere. A physical theory has an initial value formulation provided that one can prove that there exists a unique solution for a given set of appropriate initial data — intuitively, given a specification of the state of the system at one time, the dynamical laws determine a unique state at other times. A theory’s initial value formulation must satisfy further conditions to be viable physically (and qualify as “well-posed”): ‘small changes’ in the initial data lead to ‘small changes’ in the evolved states, and changes to the initial data in a region Σ should not effect causally disconnected regions. For GR, the dynamical equations are Einstein’s field equations, coupled with additional equations for the matter fields. GR possesses a well-posed initial value formulation (up to gauge freedom).³³ However, this formulation applies only to globally hyperbolic spacetimes, which form a proper subset of all spacetimes that solve Einstein’s field equations. This restriction seems natural, however, since only globally hyperbolic spacetimes admit a foliation by Cauchy surfaces. And a Cauchy surface, as stated above, is a spacelike slice Σ such that $D(\Sigma)$ is the entire spacetime. Thus, even without studying the initial value problem in its full mathematical detail, it ought to be clear that spacetimes admitting Cauchy surfaces are good candidates for a general-relativistic initial value formulation. What this also means, however, is that there is a large class of spacetimes, viz. those which fail to admit a Cauchy surface, that lack a natural dynamical interpretation. For instance, Gödel spacetime has no three-dimensional, spacelike surfaces without boundary, and thus, *a fortiori*, no Cauchy surfaces.

Let us thus restrict our attention to globally hyperbolic spacetimes, for the moment. These spacetimes always admit a foliation by Cauchy surfaces Σ_t and thus a global time function t , i.e. a

³³The qualification is essential, for if one neglects gauge freedom then the initial data appear to *underdetermine* the dynamical evolution — the solution is only fixed up to diffeomorphism. Furthermore, the initial data for GR must satisfy constraints. Once a gauge condition is imposed, Einstein’s field equations take the form of quasi-linear, second order, hyperbolic partial differential equations, which admit a well-posed local initial value formulation. See Hawking and Ellis (1973, Ch. 7), Wald (1984, Ch. 10), and Rendall (2008) for discussions of the initial value formulation of GR.

function such that each surface of constant t is a Cauchy surface of the spacetime.³⁴ The spacetime then has topology $\mathbb{R} \times \Sigma$, where the topology of the three-spaces Σ_t is arbitrary but must be the same for all Σ_t .³⁵ Conversely, however, not every spacetime with topology $\mathbb{R} \times \Sigma$ affords a global time function or a Cauchy surface—at least not in the narrower sense as defined above. As a matter of fact, the spacetime may be foliated using a “flow of space” rather than a “flow in time.” The topology of the spacetime is in both cases $\mathbb{R} \times \Sigma$, yet only the latter case would be assuredly amenable to the introduction of a global time function.

The analysis of time machines bears a close relationship to the initial value problem. Let us explicate this relationship by conceptualizing time machines starting out from some globally hyperbolic spacetime $\langle \mathcal{M}, g_{ab} \rangle$. Its Cauchy surfaces Σ_t trisect \mathcal{M} : Σ_t itself, and the ‘past’ and the ‘future’ of Σ_t .³⁶ Consider the chronological past $I^-(\Sigma)$ of Σ as fixed. It is of considerable importance for dynamical formulations of GR to determine the ‘causal stability’ of the dynamical evolution of globally hyperbolic spacetimes. The natural approach to an analysis of this issue is to consider globally hyperbolic spacetimes and study the causal structure of their maximal extensions. The original globally hyperbolic spacetime $\langle \mathcal{M}, g_{ab} \rangle$ can thus be thought of as the past and the present of some time slice Σ .³⁷ As a part of a globally hyperbolic spacetime, $I^-(\Sigma)$ does not contain CTCs. The important foundational issue is how much assurance does the theory give us that any dynamical evolution of $\langle \mathcal{M}, g_{ab} \rangle$, i.e. any maximal extension of $\langle \mathcal{M}, g_{ab} \rangle$ in accordance with the dynamical equations, is causally well-behaved. An analysis of whether CTCs can arise is part of a systematic study of this foundational issue. We will discuss how it fits into the wider problem in the next Section.

The question of whether time machines can be had in a spacetime theory thus amounts to asking whether suitable maximal extensions of a globally hyperbolic spacetime $\langle \mathcal{M}, g_{ab} \rangle$ contain CTCs. A spacetime $\langle \mathcal{M}, g_{ab} \rangle$ can be *extended* if it is isometric to a proper subset of another spacetime.³⁸ An extension of a spacetime is *maximal* if it is *inextendible*, i.e. it is not isometric to a proper subset of another spacetime. The initial state of a physical system prior to the operation of the time machine is defined on a subset of Σ . In this setting, asking of a spacetime whether a time machine can be operated in it amounts to the question of whether the spacetime admits a time slice Σ such that $I^-(\Sigma)$ does not contain CTCs and $\langle \mathcal{M}, g_{ab} \rangle$ can be analytically extended to a spacetime $\langle \mathcal{M}', g'_{ab} \rangle$ in a way that gives rise to a non-vanishing region $\mathcal{V} \subset \mathcal{M}'$ consisting of points through which there exists a CTC. One can distinguish at least three cases: either none, some, or all of the suitable maximal extensions of $\langle \mathcal{M}, g_{ab} \rangle$ contain CTCs. In the first case, $\langle \mathcal{M}, g_{ab} \rangle$ is maximally causally

³⁴We thus assume that the spacetimes at stake afford a time orientation, and that this time orientation is encoded in the global time function.

³⁵At least for spatially compact Hausdorff spacetimes $\langle \mathcal{M}, g_{ab} \rangle$; this is essentially Geroch’s Theorem (Geroch 1967) which states that for a compact spacetime $\langle \mathcal{M}, g_{ab} \rangle$ whose boundary is the disjoint union of two closed spacelike three-manifolds, \mathcal{S} and \mathcal{S}' , \mathcal{S} and \mathcal{S}' are diffeomorphic if $\langle \mathcal{M}, g_{ab} \rangle$ admits a time orientation and does not contain closed timelike curves. Intuitively, one can think of mapping points of \mathcal{S} into \mathcal{S}' using a congruence of timelike curves; for the topology to differ, at least some of these curves must fail to connect the boundaries, instead forming CTCs confined to the compact region bounded by \mathcal{S} and \mathcal{S}' . The theorem does not apply to spatially open spacetimes. For a penetrating discussion of topology change in general and of Geroch’s Theorem in particular, see Callender and Weingard (2000).

³⁶For ease of notation, we drop the index ‘ t ’ in what follows.

³⁷The immediate question that arises is what could guarantee that Σ is the ‘latest’ of all time slices, i.e. how can we assume that some part of the original globally hyperbolic spacetime isn’t to the future of Σ ? In general, we would surely expect that $I^+(\Sigma) \neq \emptyset$. However, even if $I^+(\Sigma) \neq \emptyset$ for a given Σ , it may be that there is no ‘later’ Cauchy surface Σ' , where ‘later’ is defined by the total ordering relation induced by the global time function t . Thus, we are interested in the one Cauchy surface $\Sigma_\tau \subset \mathcal{M}$ such that for all values t of the global time function, $t \leq \tau$. Or at least in one reasonably close to it.

³⁸That is, it can be extended if there exists a spacetime $\langle \mathcal{M}', g'_{ab} \rangle$, $\mathcal{M} \subsetneq \mathcal{M}'$, and an isometric embedding $\phi: \mathcal{M} \rightarrow \mathcal{M}'$ such that $\forall p \in \mathcal{M}, \phi^*(g_{ab}(\phi^{-1}(p))) = g'_{ab}(p)$.

robust and no time machine can be operated. Let’s discuss the two other cases in turn, starting with the latter. Thus, is it possible that all suitable maximal extensions of a globally hyperbolic spacetime contain CTCs?

Before we answer this question, two difficulties become immediately apparent. First, as should be obvious, time machines as conceptualized in this way will not be amenable to an avid science fiction lover’s desire to return from Σ into the causal past $J^-(\Sigma)$ of Σ . In our present set-up, time travel is confined to the future of Σ . This is the price we pay in order to make the analysis more relevant to foundational concerns. And we are happy to pay it.

Second, and more importantly, in what sense could the emergence of a non-vanishing region \mathcal{V} be appropriately said to be the ‘result of’ the operation of the time machine? Within the domain of dependence of a surface Σ , the initial value formulation gives clear content to claims such as “wiggling the value of a field on Σ is causally responsible for a corresponding wiggle elsewhere in $D(\Sigma)$.” But by construction, \mathcal{V} must lie outside of $D^+(\Sigma)$, since the past-inextendible, past-oriented curves through points in \mathcal{V} don’t intersect Σ (cf. Fig. 3).³⁹ This means that not all causal

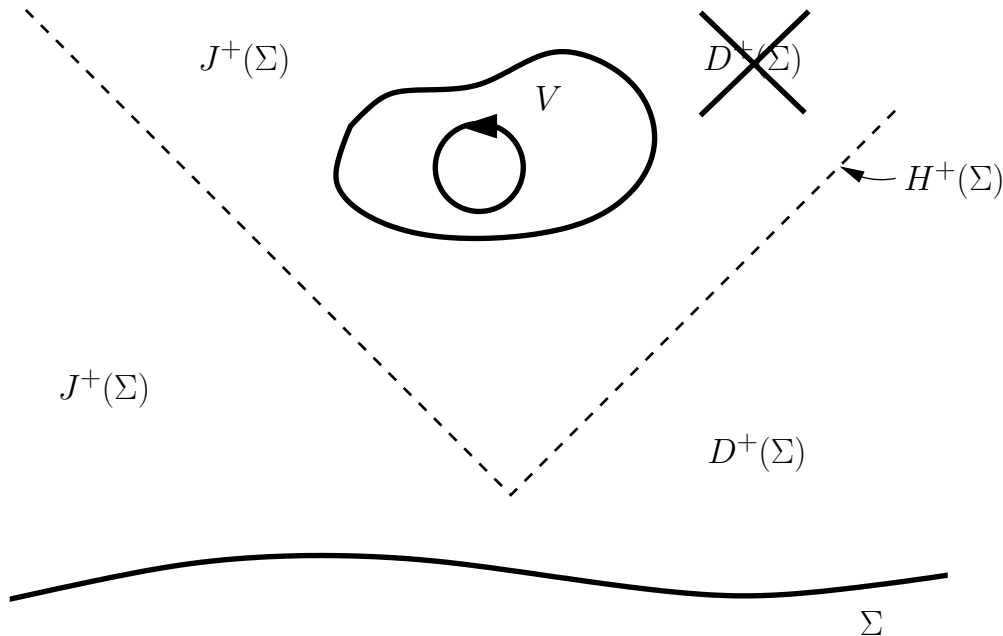


Figure 3: The causal future $J^+(\Sigma)$ contains a ‘chronology violating region’ $\mathcal{V} \neq \emptyset$, i.e. the set of all points through which there exists a CTC.

influences on the region \mathcal{V} stem from the conditions set on Σ . Some, perhaps decisive, influences on \mathcal{V} may come from outside of Σ . It also means that the original spacetime manifold \mathcal{M} that we were seeking to extend may be thought of, without loss of generality, as the region $I^-(\Sigma) \cup D(\Sigma)$.⁴⁰

The predicament is the following: in order to elucidate what it means to be “causally responsible” for CTCs, we have to abandon our best understanding of causal responsibility in GR – that

³⁹This means, of course, that Σ is no longer a Cauchy surface of $\langle \mathcal{M}', g'_{ab} \rangle$, i.e. Σ is no longer a spacelike hypersurface which every inextendible non-spacelike curve intersects exactly once. The Cauchy surfaces of $\langle \mathcal{M}, g_{ab} \rangle$ become *partial Cauchy surfaces* of $\langle \mathcal{M}', g'_{ab} \rangle$, i.e. spacelike hypersurfaces which no inextendible non-spacelike curve intersects *more than once*. Non-spacelike curves in \mathcal{V} , for instance, will not intersect Σ .

⁴⁰This alleviates the worry expressed in footnote 37. For a more thorough discussion of these points, and for the details of the construction, cf. Earman et al. (2009).

provided by the initial value formulation.⁴¹ That said, we claim that one can still give content to the idea of a time machine as follows. Although the conditions set on Σ have no prayer of being fully causally responsible for the emergence of a region \mathcal{V} , the third case where all maximal extensions of the original spacetime $I^-(\Sigma) \cup D(\Sigma)$ contain a non-vanishing \mathcal{V} is the next best thing to full causal responsibility. We say that in this case, the spacetime satisfies the

Condition 1 (Potency Condition) *Every suitable smooth, maximal extension of $I^-(\Sigma) \cup D(\Sigma)$ contains CTCs.*

The question, of course, is which smooth, maximal extensions ought to be deemed ‘suitable.’ The stronger the restrictions on which extensions qualify as suitable, and thus the fewer extensions are required to contain CTCs in order for the Potency Condition to be satisfied, the weaker the potency. A natural condition would be to restrict the Potency Condition to the extensions that satisfy the dynamical equations of the theory at stake.⁴² As we have argued in Earman et al. (2009), where we have also offered a more detailed discussion of which extensions suitably fall under the purview of the Potency Condition, this imposes a severe limitation on the analysis of time machines in spacetime theories. The causal structure as set up here is more general in that it underwrites not only GR, but also semi-classical quantum theories of gravity. For the remainder of this section, however, one can assume that the satisfaction of Einstein’s field equations is a suitability condition.

More fundamentally, however, and as also argued in Earman et al. (2009), we impose a slightly modified version of Geroch’s (1977) definition of *hole freeness*. This condition disbars artificial cut-and-paste manoeuvres in a principled manner.⁴³

Definition 1 (Hole freeness) *A spacetime $\langle \mathcal{M}, g_{ab} \rangle$ is said to be hole free just in case for any spacelike $\Sigma \subset \mathcal{M}$, there is no isometric imbedding $\iota : D^+(\Sigma) \rightarrow \mathcal{M}'$ into a spacetime $\langle \mathcal{M}', g'_{ab} \rangle$ such that $\iota(D^+(\Sigma))$ is a proper subset of $D^+(\iota(\Sigma))$ (cf. Fig. 4).*

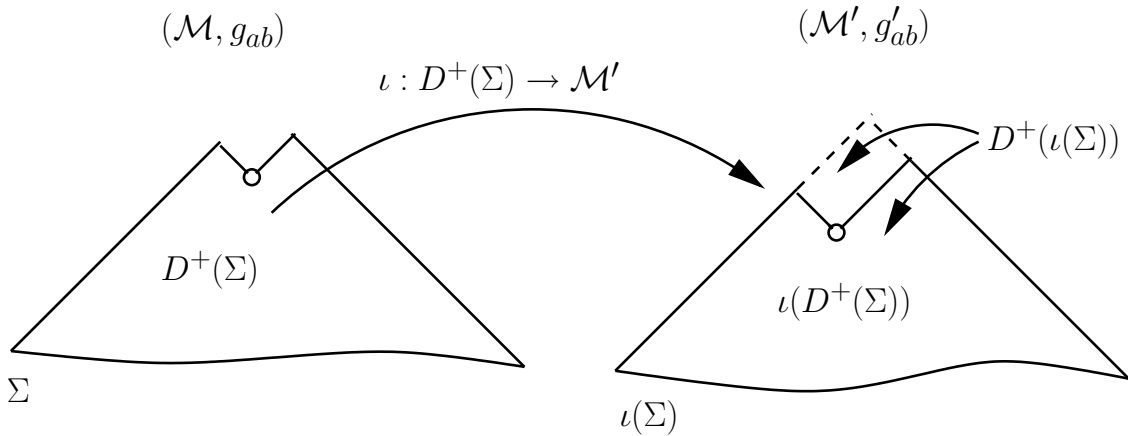


Figure 4: An illustration of the definition of hole freeness.

The demand for hole freeness has some bite, as can be seen from considering a theorem by Krasnikov (2002):

⁴¹This predicament was described clearly in Robert Geroch’s lecture at the Second International Conference on Spacetime Ontology (Montreal, June 2006).

⁴²Additionally, one might demand that the extensions comply with energy conditions, discussed in more detail below.

⁴³For a recent discussion on whether hole freeness is a physically plausible condition, cf. Manchak (2009a).

Theorem 1 (Krasnikov) *Any spacetime $\langle \mathcal{U}, g_{ab} \rangle$ has a maximal extension $\langle \mathcal{M}_{max}, g'_{ab} \rangle$ such that all closed causal (and, a fortiori, timelike) curves in \mathcal{M}_{max} (if they exist there) are confined to the chronological past of \mathcal{U} .*

The construction that Krasnikov gives in his proof of Theorem 1 allows any local conditions on the metric, such as the satisfaction of Einstein’s field equations or of energy conditions, to be carried over to \mathcal{M}_{max} . If it weren’t for the (global) condition of hole freeness, Theorem 1 would imply that there are no spacetimes that satisfy the Potency Condition. But the proof no longer goes through once the demand for hole freeness is imposed, so the question remains open.

Does any general-relativistic spacetime satisfy the Potency Condition amended with demand of hole freeness (and of complying to Einstein’s field equations)? A recent result by Manchak (2009b) answers this question in the affirmative, thereby offering a counterexample to the conjecture that Krasnikov’s theorem also holds for hole free spacetimes.

Theorem 2 (Manchak) *There exists an ESW time machine, where an ESW time machine is a hole-free spacetime satisfying the Potency Condition.*

As Manchak points out, this existence theorem establishes that we face a trilemma, in that some initial conditions force us to give up at least one of the following: either (i) spacetime is inextendible, (ii) spacetime is hole free, or (iii) spacetime does not contain CTCs. We wish to qualify this trilemma.⁴⁴ It is possible, arguably even likely, that the general-relativistic spacetime that most accurately describes the large-scale structure of our actual universe is inextendible, hole free, and entirely free of CTCs. The initial conditions which are *known* to force us into this trilemma, encoded in Misner spacetime, are not physically realistic. [Characterizing more precisely what qualifies as physically realistic initial data, and then demonstrating that the trilemma is avoided, is related to the censorship theorems discussed in the next section.](#) Having said that, however, there seems to be a vast class of spacetimes $I^-(\Sigma) \cup D(\Sigma)$ such that *some*, but not all, suitable smooth, maximal extensions contain CTCs. It is to this case—the second as listed above—that we now turn.

GR is not a deterministic theory *simpliciter*. As with other theories, GR is deterministic insofar as existence and uniqueness theorems can be proven for its dynamical equations. But the standard initial value formulation of general relativity only fixes dynamical evolution up to the Cauchy horizon of the spacetime (up to diffeomorphisms), and not beyond. As far as the evolution beyond the Cauchy horizon $H^+(\Sigma)$ is concerned, the data on Σ , together with the dynamical equations, does not uniquely determine the evolution into $J^+(H^+(\Sigma))$. This failure of determinism has motivated work on the censorship theorems discussed in the next section, but it is what makes it possible to find multiple, inequivalent extensions beyond $H^+(\Sigma)$ — including extensions with CTCs.

While it is certainly the case that the causal stability of spacetimes that can be extended beyond their Cauchy horizon to include CTCs is limited,⁴⁵ this does not imply that a time machine can be operated in any clear sense. First, these time machines would not operate perfectly in that they would sometimes fail to produce a non-vanishing region of CTCs. But why would one want to speak of a time machine if a certain fraction of suitable extensions contains CTCs while the rest doesn’t and it is a matter of pure chance as to whether the would-be time machine operators will end up living in a spacetime that evolves to contain CTCs or not. We could only appropriately speak of a time machine if the local manipulation of energy and matter distribution would somehow *increase the chance* that there will be a CTC-containing region to the future of the Cauchy horizon. We will call such a device an *incremental time machine*.

⁴⁴The first—very minor—qualification is that there is really a fourth option: the spacetime is not smooth.

⁴⁵As will be discussed below, this is a vast class of spacetimes.

The obvious manner in which talk of chance can be made respectable is in terms of probabilities. An incremental time machine would then be a device that would, by means of local operations, increase the probability that there emerge CTCs to the future of $H^+(\Sigma)$. This would be given a clear meaning if we had a probability measure defined over the set of suitable extensions. Thus, an incremental time machine would be operative just in case the probability of there being CTCs to the future of $H^+(\Sigma)$ conditional on the local manipulation performed by the incremental time machine is higher than the probability of there being CTCs to the future of $H^+(\Sigma)$ not so conditionalized, where the probabilities are given by the measure defined over the set of suitable extensions. If this can be done, we say that the spacetimes satisfies the

Condition 2 (Mitigated Potency Condition) *The operation of an incremental time machine increases the measure of those extensions of $I^-(\Sigma) \cup D(\Sigma)$ containing CTCs among the set of all suitable smooth, maximal, and hole-free extensions of $I^-(\Sigma) \cup D(\Sigma)$.*

It is far from trivial, however, to make clear sense of what this really means. So don't be fooled by the appearance of clarity and well-definedness, if there was any, in how we set up the question. The foremost, and we believe deeply problematic, issue is how to cash out the references to 'agents,' sentient or not, who 'manipulate' the energy and matter content of the universe to 'influence' or 'change' the future in a way such as to 'bring about' or 'produce,' at least with partial success, CTCs in the region beyond the Cauchy horizon. These agents could not be free in the sense of disposing of genuinely alternative action among which they can choose since they operate in a globally hyperbolic region of spacetime. As we have seen above, the dynamics of GR are deterministic up to diffeomorphism in globally hyperbolic regions.⁴⁶

True, free will may not require there to be such genuine alternatives, but even if that is granted, the problem of our counterfactual formulation remains. In our analysis of (strict) time machines, we circumvented this issue, as was appropriate there, since the analysis reduced to an investigation of whether all suitable extensions of a given, particular globally hyperbolic spacetime contained CTCs. Sidestepping the problem is not longer an option if we want to make sense of what it could mean to 'increase the measure of extensions' with certain properties. The notion of 'increasing' straddles us with counterfactual discourse that is difficult to parse out. The best way to escape the morass of agency, free will, manipulation, and counterfactuals, it seems to us, is to migrate our discussion to two problems in the neighbourhood of the original one: Can we get a principled handle on defining probability measures over the set of suitable extensions of $I^-(\Sigma) \cup D(\Sigma)$; and can we gain some understanding concerning the physical mechanisms that might be responsible for the emergence of causally unusual structures such as CTCs? We will not broach the second of these issues here.⁴⁷ And we have only very little to say concerning the first.

The most principled way of addressing the first issue of introducing probability measures over the set of suitable extensions is to start out from the set of all admissible spacetimes of the theory, define a probability measure over this set and conditionalize on the subset of spacetimes in which the spacetime-to-be-extended at stake can be isomorphically embedded. This would result in a probability distribution over the set of suitable extensions. With such a probability distribution at hand, we could then determine the relative frequency of extensions-cum-CTCs in terms of their measure in the space of suitable extensions. This sounds all very principled and rigorous, but we face unresolved, and perhaps unresolvable, difficulties at every turn of this path.

⁴⁶It's not as if the existence of diffeomorphic alternatives opens the door to free will if genuine alternatives of action are necessary for free will. The reason for this is that diffeomorphically related spacetimes are usually considered as physically identical variants, an interpretation that we accept.

⁴⁷For a step into this direction, see Andr eka et al. (2008). Cf. also footnote 26.

First, at least for GR, the set of admissible spacetimes, i.e. the space of solutions to Einstein’s field equations, is not known. Second, and *a fortiori*, we cannot define a probability measure over this set. But even if we pretend that we know the set, or at least some significant subset of it, it is not trivial to endow it with a canonical measure and not much of help can be found in the literature. Most of the attempts to define such measures over sets of solutions focus on causally well-behaved spacetimes such as the Friedmann-Lemaître-Robertson-Walker cosmological models.⁴⁸ These measures have primarily been designed to deal with the ‘flatness problem’ in standard cosmology in an attempt to avoid inflationary scenarios. Extant results in this field, however, are hardly of much use to our present purposes since they only extend to a particular parameter family of comparatively well-understood spacetimes. Without a more general measure defined over larger classes at hand, the prospect of this research programme seems daunting, if not downright hopeless.

There is an alternative way to obtain a sense of how generic maximal extensions containing CTCs are for given initial spacetimes. Rather than measuring extensions with CTCs, we might be able to count them. Presumably, a theorem analogous to Theorem 1 may be found which shows that it is always possible to find an extension which respects certain local conditions while displaying CTCs. Perhaps we could then establish a theorem of ‘parallel existence’ according to which there exists, for every causally virtuous extension, an extension with CTCs. Take, for instance, any ‘clean’ extension received with the help of Theorem 1. It seems that any of these extensions could be infected e.g. by a Deutsch-Politzer gate, thus producing CTCs.⁴⁹ In all fairness, such limited theorems could at best serve to strengthen our intuitions. In order to obtain more conclusive statements concerning the genericity of causal and acausal extensions, one would have to establish theorems asserting the open density of one of the two families of extensions, causal or acausal, thus showing that it is of measure one, while its complement is ‘nowhere dense’ and therefore of measure zero. Such a proof, however, would again require a well-defined measure on the space of extensions.

There is thus little hope that much more can be said about the second case where some, but not all, extensions contain CTCs. Before we close this section, a brief word concerning the first case, the one where none of the suitable extensions harbours CTCs and the initial spacetime is maximally causally robust. It seems as if hole freeness is not required of the admissible extensions, then there always exists a smooth, maximal extension with CTCs that can be gained by the scissors-and-glue method.⁵⁰ Trivially, those spacetimes which already are maximal and thus admit of no non-trivial extension beyond the Cauchy horizon will not be extendible in a way that includes a non-vanishing region \mathcal{V} . Thus, the conjecture has to be reformulated as claiming that for all spacetimes *that permit a non-trivial extension beyond their Cauchy horizon*, there exists a suitable smooth, maximal extension containing CTCs. Furthermore, once one adds the constraint of hole freeness, as we did above, it seems as if this is no longer possible in general. It is thus an open question whether there always exists a *hole-free*, smooth, maximal extension of $I^-(\Sigma) \cup D(\Sigma)$ containing CTCs.

6 Censorship Theorems

Given this understanding of what a TM requires in terms of spacetime structure, we can now return to the question of whether TMs are physically possible according to GR. The discussion above

⁴⁸Cf. Hawking and Page (1988); Cho and Kantowski (1994); and Coule (1995).

⁴⁹Deutsch 1991, Politzer 1992.

⁵⁰Presumably including the satisfaction of local conditions such as the dynamical equations and energy conditions. We wish to thank David Malament and John Manchak for pointing this out to us.

shows that TMs are physically possible in the weak sense that there are spacetime geometries that instantiate a TM. But we have set aside until now the question of whether such spacetimes, or more generally spacetimes with CTCs or other causal pathologies, are physically possible in the stronger sense that there are solutions to Einstein’s field equations exhibiting these features that are “physically reasonable.” There are several conditions one might impose to delineate the subset of spacetimes that qualify as reasonable:

1. *Causality Conditions*: treat one of the causality conditions (e.g., global hyperbolicity) as a law of nature not derived from the field equations.
2. *Conditions on Source Terms*: impose energy conditions on the matter fields, or limit consideration to T_{ab} derived from “fundamental fields.”
3. *Generic*: rule out “special” spacetimes (e.g., those of measure zero) as possible models for real systems.
4. *Quantum Considerations*: impose conditions that the spacetime must satisfy to admit a QFT, or to be the classical limit of a quantum gravity solution.

The first option is sometimes motivated by the paradox-mongering we have criticized above. There is a further objection to simply imposing causality constraints, namely that one sets aside the possibility that they might be enforced by the dynamics combined with other conditions. As Hawking puts it (in Hawking and Penrose 1996, 10),

... [M]y viewpoint is that one shouldn’t assume [global hyperbolicity] because that may be ruling out something that gravity is trying to tell us. Rather, one should deduce that certain regions of spacetime are globally hyperbolic from other physically reasonable assumptions.

Attempts to decipher what gravity is trying to tell us have led to what we will call generally “censorship hypotheses,” and in some cases censorship theorems. Such a theorem shows that specific features such as causal pathologies either do not develop under dynamical evolution from suitable initial conditions, or, failing that, that these features are “censored” – i.e. hidden safely behind the event horizons of black holes. The name derives from Penrose’s “cosmic censorship hypothesis” (Penrose 1969, 1979), which in slogan form holds that nature abhors a naked singularity. Turning this slogan into a precise claim amenable to proof is no easy task, and the proper formulation and status of cosmic censorship remains one of the central open problems in classical GR. Here we will briefly survey how conditions (2) and (3) come into play in attempts to state clearly and then prove censorship theorems (leaving 4 until the next sections), and focus on “chronology protection” results which aim to demonstrate the impossibility of TMs.

Penrose formulated the cosmic censorship hypothesis on the heels of his ground-breaking work (along with Hawking and Geroch) on the singularity theorems. These theorems show that a large class of physically reasonable spacetimes, relevant for cosmology and for the gravitational collapse of stars, must be singular, in the sense of geodesic incompleteness.⁵¹ However, although all globally hyperbolic spacetimes resemble one another, each singular spacetime is singular in its own way, and the singularity theorems reveal little about the nature of the singularity. A successful proof of cosmic censorship would rule out all but the relatively benign singularities. Specifically, one would hope to rule out *nakedly singular* spacetimes. These are defined, intuitively, as spacetimes containing

⁵¹A spacetime is *geodesically incomplete* just in case there exist geodesics in it which are inextendible in at least one direction (timelike, null, or spacelike) yet run only over a finite range of their affine parameter.

point(s) p such that $I^-(p)$ includes an entire inextendible timelike curve with finite proper length, representing the trajectory γ of a point particle that “falls into the singularity” – the singularity that γ encounters is “visible” from such points.⁵² Strong cosmic censorship asserts that there are no such points, and Penrose shows this holds iff a spacetime is globally hyperbolic. Strong censorship can be formulated in slightly different terms as follows: strong cosmic censorship holds if initial data specified on Σ (for a suitable hypersurface)⁵³ have a maximal Cauchy development $D(\Sigma)$ that is inextendible. There are counterexamples to the conjecture if there are not further qualifications, however. One such counterexample (based on numerical simulations by Choptuik and analytical work by Christodoulou) prompted Hawking to pay up on a famous bet in favor of cosmic censorship against Kip Thorne and John Preskill. (Hawking wagered in favor of the claim that “When any form of classical matter or field that is incapable of becoming singular in flat spacetime is coupled to general relativity via the classical Einstein equations, the result can never be a naked singularity,” at 2 to 1 odds.)⁵⁴

One response to such counterexamples is to give a more refined formulation of cosmic censorship, as Hawking did in immediately placing a new wager with Thorne and Preskill regarding the claim that: “When any form of classical matter or field that is incapable of becoming singular in flat spacetime is coupled to general relativity via the classical Einstein equations, then dynamical evolution from generic initial conditions (i.e., from an open set of initial data) can never produce a naked singularity (a past-incomplete null geodesic from [future null infinity] \mathcal{I}^+)”. The clarification of what is meant by a “naked singularity” is based on a distinction between “local observers” and “observers at infinity,” and requires that only the latter are safely shielded from the singularity by the event horizon of a black hole. This is usually called weak cosmic censorship.⁵⁵ The formulation of weak cosmic censorship relies on a precise notion of infinity developed in the study of gravitationally isolated systems. Roughly speaking, a spacetime is *asymptotically flat* at late times if it can be conformally embedded into a spacetime with a boundary \mathcal{I}^+ composed of the endpoints of null geodesics that propagate to arbitrarily large distances.⁵⁶ The black hole region \mathcal{B} is then defined as the region of the manifold from which light cannot escape to \mathcal{I}^+ – i.e., the complement of $I^-(\mathcal{I}^+)$ – and the event horizon is the boundary $\dot{\mathcal{B}}$ (see Figure 5). Inset: Definition of completeness of Scri We then have the following formulation of weak cosmic censorship: the maximal Cauchy development of *generic*, asymptotically flat initial data for Einstein’s field equations with *suitable* matter fields has a complete \mathcal{I}^+ . If weak cosmic censorship holds, it would imply that even if it is in principle possible to produce a singularity (even one that is not “benign” locally) by rearranging matter and energy within a finite region of spacetime, the effects of this singularity will not reach distant observers. The region outside the black hole is “safe for determinism” in the sense that the data on an appropriate slice Σ uniquely determines the evolution throughout the asymptotic region.

⁵²Here we follow Penrose’s characterization of cosmic censorship in terms of detectability (Penrose 1979); his formulation is based on the ideas of “indecomposable past sets” and “terminal indecomposable past sets,” a way of defining ideal or boundary points. See Earman (1995) and Geroch and Horowitz (1979) for discussions of alternative approaches to formulating cosmic censorship.

⁵³The qualifier is required to rule out surfaces such as an achronal surface Σ for which $\exists p : \Sigma \subset I^-(p)$ in Minkowski spacetime; for such a surface, $D(\Sigma)$ is not the entire manifold, but this just reflects the poor choice of Σ .

⁵⁴The bets are posted outside Thorne’s office at Caltech, and are discussed in Thorne (2002).

⁵⁵There are other ways of formulating a weaker condition (Earman 1995, 74-75), such as adding further requirements related to the curve γ in the definition of a naked singularity – e.g., that it is a geodesic curve, or that curvature invariants blow up along γ .

⁵⁶In slightly more detail, we require that there is a spacetime $(\tilde{\mathcal{M}}, \tilde{g}_{ab})$, the conformal completion of (\mathcal{M}, g_{ab}) , consisting of $\mathcal{M} \cup \mathcal{I}$ (where \mathcal{I} is the asymptotic region consisting of a past and a future part, \mathcal{I}^+ and \mathcal{I}^- respectively, along with spatial infinity i^0), and a conformal isometry such that $\tilde{g}_{ab} = \Omega^2 g_{ab}$ on \mathcal{M} ; see, e.g., Wald 1984, Chapter 11.

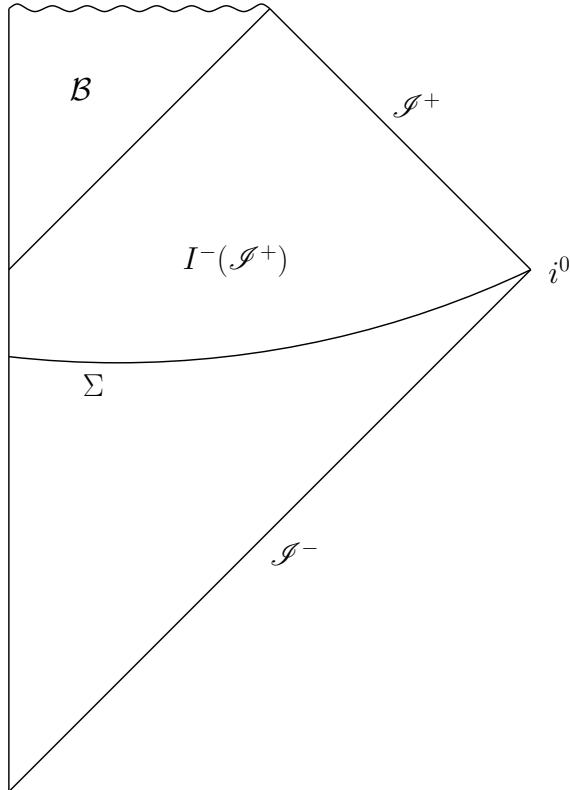


Figure 5: A black hole.

Proving weak cosmic censorship requires first giving some precise content to the qualifications inserted in the above formulation (regarding *generic* initial data, and *suitable* matter fields). The initial wager above explicitly limited suitable matter fields to those that do not themselves develop singularities in flat spacetime. This restriction is motivated by the desire to separate singularities due to the matter fields themselves from those due to gravity, but it rules out fluids because shocks and other singularities do occur in flat spacetime.

The restriction that is more significant for our purposes is an implicit imposition of energy conditions (ECs). These are restrictions on T_{ab} , the stress-energy tensor appearing on the right hand side of Einstein’s field equations, sometimes taken to characterize “reasonable” classical fields. Einstein’s field equations can be solved for specific sources – that is, a specific choice regarding the matter distribution – but one can also consider solutions that obtain as long as the stress-energy tensor satisfies certain conditions. (Without such restrictions, Einstein’s field equations can be taken to *define* the stress-energy tensor for a given g_{ab} .) The *dominant energy condition* (DEC) holds if the energy-momentum flow measured by any observer at any point is a timelike or null vector; intuitively, this requires that energy-momentum propagates on or within the light cone. Technically, the DEC requires that for all pairs of future-directed, unit timelike vectors ξ^a, ζ^a , $T_{ab}\xi^a\zeta^b \geq 0$. If we require only that the inequality holds when $\zeta^a = \xi^a$, we have the *weak energy condition* (WEC); intuitively, this requires that there are no negative energy densities according to any observer. Finally, the *strong energy condition* (SEC) requires that $T_{ab}\xi^a\xi^b \geq \frac{1}{2}\text{Tr}(T_{ab})$ for every unit timelike ξ^a . (This terminology is misleading, in that the SEC does *not* entail the WEC.)

What is the status of these conditions? They were once taken as defining properties for “reasonable” fields, and they effectively guarantee that gravity is an “attractive force” — in the sense that

geodesics converge in regions with a non-zero stress-energy tensor. In contrast, geodesics *diverge* in regions with non-zero EC-violating fields, other things being equal. Due to this stark contrast, the ECs clearly play an essential role in various results in classical GR. Recently physicists have seriously considered EC-violating fields in a variety of situations that seem to call for such repulsive behavior, such as inflationary cosmology and in modeling “dark energy” thought to drive the observed accelerated expansion of the universe. One can formulate classical theories with a scalar field which violate one or more of the ECs but are arguably “physically reasonable.”⁵⁷ In addition, we will see in the next section that these inequalities are fundamentally at odds with the kinematics of quantum field theory.

The second qualification in the formulation above allows that there may be counter-examples based on highly “special” initial data sets (e.g., with a high degree of symmetry) that lead to a naked singularity. The refined bet raises the bar: any counter-example to cosmic censorship must work for an open set of initial data, which requires that “nearby” initial data would also lead to naked singularities. Specifying what qualifies as “special” or “generic” initial conditions requires introducing a measure on the space of initial data. We concur with Wald (1998)’s assessment that choosing an appropriate measure or topology demands greater insight into the dynamics of GR than we currently possess, so any proposed definitions of “special” initial data are provisional.

A successful proof of strong cosmic censorship would immediately rule out TMs as defined above, because it would establish that “suitable” initial data lead to an inextendible spacetime — with no Cauchy horizon, and no possibility of CTCs developing beyond it. Weak cosmic censorship, on the other hand, would only establish that asymptotic observers would be shielded from causal pathologies, including CTCs, by an event horizon. There are related censorship theorems whose proof may shed some light on these conjectures. Hawking (1992)’s “chronology protection conjecture” is more specific than either version of cosmic censorship, as it states that the laws of physics do not allow CTCs to be created. But before sketching Hawking’s argument, we will briefly discuss one result that has been proven: the topological censorship theorem.

Why is it the case that physical spacetime appears to have a surprisingly simple topology? Nothing within GR rules out attributing enchantingly baroque topological structures to spacetime. A topological space is simply connected if any closed loop can be smoothly contracted to a point. A torus is an example of a multiply connected space; the loops going around the central hole and those looping around the ring itself cannot be contracted to a point. The basic requirements GR places on an underlying manifold in order to qualify as a spacetime do not rule out multiply connected topologies, and even requiring that a given three-manifold can serve as a “reasonable” initial data surface does not place constraints on the topology. Witt (1986) showed that it is possible to specify vacuum initial data for *arbitrary* closed or asymptotically flat three-manifolds. Given this tolerance for varied topology, one may wonder why the GR models describing observed, macroscopic regions are simply connected.

The topological censorship theorem shows that in GR any “naked topology” is hidden in much the same fashion as naked singularities would be hidden according to weak cosmic censorship. Multiply-connected topologies are hidden in that any causal curve connecting past and future null infinity in an asymptotically flat spacetime (\mathcal{I}^- and \mathcal{I}^+) cannot pass through a non-simply connected region (Friedman, Schleich, and Witt 1993; cf. Galloway 1995). (Suppose that there is a curve γ from \mathcal{I}^- to \mathcal{I}^+ that “threads the topology”, in the sense that it cannot be deformed into (i.e. is not homotopic to) a curve lying entirely in the asymptotic region, which is simply

⁵⁷Barceló and Visser (2002) argue that all of the energy conditions are “dead” or “moribund,” based in part on theories with scalar fields. Mattingly (2001), perhaps the only paper focused on energy conditions in the philosophy of physics literature, endorses a similar position. We will discuss the status of the ECs in more detail in the next section.

connected. The proof proceeds by showing that any such curve would have to thread a strongly outer trapped surface, but the properties of such a surface imply that γ cannot reach \mathcal{S}^+ .⁵⁸ The theorem was inspired by a result of Gannon (1975), which shows (using the same methods as the Hawking-Penrose singularity theorems) that incomplete geodesics occur to the future of a non-simply connected Cauchy surface in an asymptotically flat spacetime. Gannon’s proof implies that the incomplete geodesics arise as a result of the topology of the Cauchy surface, leading to the natural question whether event horizons develop as well – and the topological censorship theorem establishes that they do. Any observer lucky enough to enter a region of spacetime with multiply-connected topology would be trapped behind an event horizon just as surely as an astronomer on the unfortunate mission of exploring the interior of a black hole.

Topological censorship is closely connected to the discussion of time machines due to its implications for traversable wormholes. The discovery by Kip Thorne and co-workers that a traversable wormhole could be used as a time machine provided that the two mouths of the wormhole are in relative motion inspired much of the current literature on TMs.⁵⁹ Topological censorship might appear to rule out the use of traversable wormholes as TMs, but the theorem applies only granted two strong assumptions. The first assumption is that the averaged null energy condition holds. This is weaker than the pointwise energy conditions stated above: it allows that at some points $T_{ab}\xi^a\xi^b$ is negative (where ξ^a is a null vector), but requires that the average value of $T_{ab}\xi^a\xi^b$ along a null geodesic (where ξ^a are the null tangent vectors) is non-negative. The importance of this assumption indicates that maintaining a wormhole requires a form of “exotic” matter for which the averaged null energy condition fails to hold. Second, the theorem requires a strong causality condition – global hyperbolicity. Thus the theorem might be best thought of as elucidating the consequences of global hyperbolicity for the topology of the asymptotic region and for event horizons.⁶⁰ Obviously this second assumption undercuts the usefulness of this result as a no-go theorem for TMs, as the assumption rules out CTCs by *fiat*.

Hawking (1992) aimed to establish such a no-go theorem for TM spacetimes that does not depend on imposing a causality condition. In general terms one would like to prove a theorem of the following form:

Conjecture 1 ((Classical) Chronology Protection) *Given initial data, satisfying “-----,” specified on a surface Σ , there exists a solution of Einstein’s field equations $\langle \mathcal{M}, g_{ab} \rangle$ (unique up to diffeomorphism), with properties “-----,” that does not contain CTCs.*

⁵⁸More precisely (following Friedman, Schleich, and Witt 1993), the theorem states that given an asymptotically flat, globally hyperbolic spacetime satisfying the averaged null energy condition, every causal curve from \mathcal{I}^- to \mathcal{S}^+ is homotopic to γ_0 , a causal curve lying in the simply-connected region U of \mathcal{S} . The proof of the theorem relies primarily upon one lemma, namely that none of the generators of $J^+(\tau)$ for an outer-trapped surface τ intersect \mathcal{S}^+ . If we suppose that one of the generators of $J^+(\tau)$, say ξ , intersects \mathcal{S}^+ , which implies that τ intersects $J^-(\mathcal{S})$, then ξ would stretch from τ to \mathcal{S}^+ with infinite affine parameter length. ξ may only remain a generator of the boundary if it has no conjugate points (i.e., a point where it intersects an infinitesimally neighboring member of a congruence of null geodesics). However, the energy condition guarantees that the matter-energy distribution in the spacetime will serve to focus null geodesics. Given that the expansion is initially non-positive (which is the case because τ is an outer-trapped surface), this focusing must lead to a conjugate point within finite affine parameter length. Thus, we have a contradiction, and the original hypothesis that ξ intersects \mathcal{S}^+ is false. This implies that γ cannot both be homotopically inequivalent to γ_0 , as that would require threading an outer-trapped surface τ , and connect \mathcal{S}^- to \mathcal{S}^+ .

⁵⁹These ideas were introduced in Morris and Thorne (1988) and Morris, Thorne and Yurtsever (1988), see also Thorne’s engaging account of how this line of research unfolded in Thorne (1994), Chapter 14.

⁶⁰Gannon shows a weaker condition is all that is in fact required (...need to track reference down...). Note that the main use of topological censorship in the physics literature is proving results regarding the topology of black holes; see Friedman and Higuchi (2006) for further discussion and references.

The importance of the result depends on exactly what goes into the two blanks: the result would be more decisive to the extent that there are unambiguous, well-motivated ways to fill in both blanks, such that the precisely formulated claim is amenable to proof. Filling in the blanks requires facing up to the same challenges we saw above in the discussion of strong and weak cosmic censorship: what is the status of the energy conditions one might impose in order to give Einstein’s field equations some bite? How should we formulate the requirement that the resulting solution is not “special” within the space of solutions? It will thus come as no surprise that debate regarding results that have been established focuses on these questions.

Rather than attacking the general conjecture directly, one might instead formulate a necessary condition for a TM spacetime and then show that this condition is incompatible with some other requirement (such as the energy conditions). Hawking (1992) argued that a suitable necessary condition for a TM spacetime is the existence of a *compactly generated Cauchy horizon*, and then showed that the existence of such a horizon entails that the null energy condition must be violated. A Cauchy horizon is compactly generated if the null geodesics that are generators of the surface enter and remain within a compact set. (This condition is meant to rule out influences “coming from infinity” as having some impact on the creation of CTCs in the region \mathcal{V} .) Hawking’s argument then proceeds roughly as follows. The generators of the Cauchy horizon are null geodesics segments, which do not have past endpoints.⁶¹ However, in order to enter and remain within a compact set (towards the past) the generators must converge. Imposing the null energy condition along with the assumption that the geodesics encounter some non-zero energy density (sometimes called the generic condition) leads to a contradiction, because the presence of matter satisfying the energy condition implies, via the Raychaudhuri equation, that the generators must converge and have past endpoints. Thus the existence of a compactly generated Cauchy horizon requires violation of the null energy condition or the generic condition on the Cauchy horizon.

One response to Hawking’s result is to question whether the presence of a compactly generated Cauchy horizon is an appropriate necessary condition (see Earman et al. 2009 for further discussion). We have argued above in favor of the Potency Condition as an appropriate definition of a TM spacetime, and it is not the case that the Potency Condition fails in cases where the Cauchy horizon fails to be compactly-generated (cf. Ori 2007). Even though our preferred definition of TM is thus broader than Hawking’s, his results might apply more broadly as well, to spacetimes with a Cauchy horizon. It is plausible that the existence of a Cauchy horizon, compactly generated or not, will be accompanied by violations of energy conditions in classical GR (cf. the earlier results of Tipler 1976, 1977), although we do not know of any theorems that establish the general claim. There are also some results indicating that Cauchy horizons may be “special” in the sense of being measure zero given a reasonable measure assigned to the space of solutions of Einstein’s field equations. For example, Moncrief and Isenberg (1983) prove that in a particular case (namely, granted that the Cauchy horizon is analytic and ruled by closed null geodesics) there are symmetries in the neighborhood of the Cauchy horizon. A more general result along these lines would support the strong cosmic censorship by showing that the existence of Cauchy horizons does not hold generically, for open sets of initial data rather than just for specific cases.

But the more intriguing question is whether we should trust the physical description of a Cauchy horizon offered by classical GR. Hawking (1992) hoped to formulate a quantum chronology

⁶¹A *past endpoint* of a curve $\gamma(s)$ is a point p such that for every neighborhood O of p there exists s' such that $\gamma(s) \in O$ for all $s < s'$ (where the parameter s increases with time along the curve). Note that although there is at most one past endpoint of a curve, the curve may “continually approach” p without there being a value of s such that $\gamma(s) = p$. The proof that the generators have no past endpoints uses the same techniques as proofs of the properties of other achronal surfaces such as $\dot{I}^+(S)$; the existence of a past endpoint in $H^+(\Sigma)$ is incompatible with the properties of such a surface.

protection conjecture based on a combination of ideas from quantum theory and classical GR. Quantum theory provides encouragement to time travel fans because quantum fields do not satisfy the point-wise energy conditions. This opens the possibility of treating results implying the violation of energy conditions as simply characterizing the fact that quantum effects will become important at the Cauchy horizon. In any case, Hawking sought a stronger no-go theorem based on semi-classical quantum gravity. This is a hybrid theory which incorporates the effects of quantum fields as sources within classical GR, without attempting to quantize spacetime geometry itself. The goal is to calculate the so-called “backreaction” of quantum fields as a perturbation to a classical spacetime, by putting the quantity $\langle \xi | T_{ab} | \xi \rangle$ (the expectation value of the renormalized stress energy tensor for the quantum state ξ) into Einstein’s field equations. Hawking (1992) argued that, *contra* Morris and Thorne (1991)’s earlier results, the divergence of this quantity near the Cauchy horizon enforces chronology protection: the backreaction effects effectively prevent the formation of CTCs. The ensuing debate led to a theorem by Kay, Radzikowski and Wald (1997) which shows that the quantity $\langle \xi | T_{ab} | \xi \rangle$ is not well-defined at all points of a compactly generated Cauchy horizon.⁶² This result may be taken to support Hawking’s conjecture, in that it clarifies the pathologies associated with compactly generated Cauchy horizons. However, one might instead read the result as showing that Cauchy horizons lie outside the domain of applicability of semi-classical quantum gravity. Even if we might have hoped to prove the chronology protection conjecture using a hybrid theory such as semi-classical quantum gravity, it seems that the full-fledged quantum gravity cannot be avoided. This is not so much a shortcoming of current results as a reason for interest in this topic: determining whether chronology protection holds, and if so *why* it holds, may provide some insight into a theory of quantum gravity.

7 Energy Conditions in QFT

Before briefly commenting on TMs in full quantum gravity, will consider the status of energy conditions in QFT. In GR the term appearing on the right-hand side of the field equations (the Einstein tensor G_{ab} , constructed out of the metric and its first and second derivatives) characterizes spacetime geometry, whereas the stress-energy tensor on the left-hand side (T_{ab}) contains information about the source of the gravitational field. Einstein frequently expressed his dissatisfaction with the need to put the stress-energy tensor into the field equations by hand, by choosing a particular matter model. Einstein (1936, 335) described the field equations as a building with two wings: the left-hand side built of fine marble, and the right-hand side built of low-grade wood. Presumably Einstein’s unified field theory project, if successful, would have produced a building constructed entirely of marble, in which fields act as their own sources and there is no need for independent matter models. But one need not share Einstein’s goal of unification to have reason to avoid relying too heavily on low-grade wood, in the sense of proving results that hold only for particular stress-energy tensors. In order to study dynamical evolution according to Einstein’s field equations at a more general level, it is natural to consider properties shared by stress-energy tensors for different types of matter. The energy conditions allow a more general approach and they are crucial assumptions in major results in classical GR such as the singularity theorems and positive mass

⁶²They demonstrate, roughly speaking, that for a compactly generated Cauchy horizon there is a non-empty set of base points, which are past terminal accumulation points for some null geodesic generator γ – intuitively, γ continually re-enters any given neighborhood of the point p . In any neighborhood of such points, there are points that can be connected by a null curve “globally” (within the full spacetime) even though they cannot be connected by a null curve “locally” (i.e. within the neighborhood). This conflict between local and global senses of null related undermines the standard prescription for defining $\langle \xi | T_{ab} | \xi \rangle$. See Earman et al. 2009 for further discussion, and Visser 2003, Fewster and Higuchi 2008 for more detailed reviews and further references.

theorems.⁶³ However, the point-wise energy conditions described above *all* fail for quantum fields. What does this failure imply regarding results such as the singularity theorems and the status of spacetimes with exotic causal structure?

As we mentioned above, the failure of point-wise energy conditions in QFT provides some encouragement, in the sense that quantum fields may provide the “exotic matter” needed to violate the energy conditions and, for example, maintain a traversable wormhole long enough to convert it into a time machine. The point-wise energy conditions fail in QFT because the energy density necessarily admits arbitrarily negative expectation values (as shown by Epstein, Glaser, and Jaffe 1965) at a point. These negative energy densities occur even though the overall energy is positive, and they arise due to quantum coherence effects. Exploiting quantum fields as a kind of “exotic matter” requires understanding the failure of the energy conditions in more detail. For example, does QFT allow one to have quantum states with large negative energy densities not just at a single point but over an extended region of spacetime? And what would be required in order to create macroscopic wormholes or other exotic structures, as opposed to “micro-scopic” (Planck scale) exotic structures? Recent work has demonstrated that QFT *does* place constraints on negative energy densities, in the form of “non-local” energy conditions (specifying limits on energy densities over spacetime regions rather than points). The precise nature of these constraints and their implications for debates regarding TMs are still being debated, but the current results provide some evidence that the energy conditions enforced by QFT will be sufficient to rule out the exotic structures incompatible with the point-wise energy conditions in classical GR (within the domain of applicability of semi-classical quantum gravity).⁶⁴

Ford (1978) originally proposed non-local energy conditions (which he calls “quantum inequalities”) based on thermodynamical considerations: he argued that negative energy densities without further constraints would lead to a violation of the second law of thermodynamics. If one could manipulate the quantum fields appropriately, it would be possible in principle to use negative energy fluxes to lower both the temperature and entropy of a hot body. Ford argued that if the magnitude of the flux is small enough (in particular, $\Delta E \geq -\hbar/\Delta t$), given the time scale during which it was transferred to the body Δt , the effect of the flux on the body’s energy would be smaller than the uncertainty in the body’s energy. Thus avoiding a conflict with thermodynamics requires placing constraints on the negative energies and fluxes allowed by QFT. Ford and Roman subsequently derived a number of quantum inequalities for different cases. These results generally take the form of showing that there is a (state-dependent) lower bound on the negative energy density “smeared” over a spacetime region, such that the bound on the energy varies inversely with the size of the region. As far as we know, the results obtained so far are still a patchwork quilt covering a variety of different cases – different choices of fields, flat vs. curved spacetimes, and etc (see, e.g., Fewster 2005). But they are very suggestive that while QFT allows for negative energy densities, the resulting violations of the point-wise energy conditions will not be sufficient to undermine the results of classical GR.

8 From Classical to Quantum Gravity

Turning to full quantum theories of gravity now, we would like to take a brief look at three rather different approaches: causal sets, loop quantum gravity, and string theory. Causal sets is an iconoclastic approach to formulating a quantum theory of gravity that does not rely on known

⁶³The positive mass theorems establish that the total energy associated with an isolated system is positive (see, e.g., Wald 1984, Chapter 11).

⁶⁴See, e.g., Flanagan and Wald 1996, Roman 2004, Ford ..., Fewster 2005, and Friedman and Higuchi 2006.

physics as a vantage point, instead trying to arrive at such a theory *ab initio* (Bombelli et al. (1987); Bombelli et al. (2003)). The causal sets approach postulates a fundamentally discrete spacetime structure that satisfies a few simple conditions and tries to establish that in the classical limit, the continuous spacetimes of GR can be recovered.⁶⁵ More particularly, the approach demands that the elements of the fundamental spacetime exhibit the structure of a *causal set*. Causal sets \mathcal{C} are endowed with a binary relation \prec such that for all $a, b, c \in \mathcal{C}$, (i) $a \prec b$ and $b \prec c$ imply $a \prec c$ (transitivity), (ii) $a \not\prec a$ (acyclicity), and (iii) all *past sets* $\mathcal{P}(a) \doteq \{b : b \preceq a\}$ are finite. Condition (ii) amounts to ruling out CTCs by stipulation, at least at the fundamental level. Although it is not yet clear how the theory relates these discrete structures to the continuous spacetimes of GR, since it fundamentally encodes the causal structure in the manner specified above, it cannot give rise to continuous spacetimes containing CTCs. Malament’s (1977) theorem, mentioned above, establishes that if $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}', g'_{ab} \rangle$ are both past and future distinguishing spacetimes, and if there exists a bijection f between \mathcal{M} and \mathcal{M}' such that both f and f^{-1} preserve the causal precedence relations, then f must be homeomorphism, i.e. a topological isomorphism between the manifolds. This implies, unsurprisingly, that an approach encoding only the causal structure cannot allow closed causal curves. This means that such an approach does not command the resources to recover the metric structure of classical GR in its full generality. The spacetimes that can be captured in the continuum limit by a causal-set approach thus represent a proper subset of those admitted by the Einstein equations, excluding those with CTCs. In fact, the causal sets approach is wedded to a commitment to the first way of delineating physically reasonable spacetimes of the previous section, i.e. the one based on causality conditions. And we maintain, as above, that there are good reasons to prefer that causality conditions such as global hyperbolicity be deduced from independently motivated assumptions, rather than stipulated by hand.

To be sure, if a particular approach to quantum gravity turns out to offer a successful, or even only a viable, quantum theory of gravity, then such success or viability would **trump our objections to imposing causality conditions *a priori***. Unfortunately, it looks like the causal sets are far from delivering this. More promising, arguably, is another approach to quantum gravity: loop quantum gravity (Rovelli 2004, Thiemann 2007). Loop quantum gravity (LQG) attempts a canonical quantization of a Hamiltonian formulation of GR.⁶⁶ As a research program with the ambition of delivering a full quantum theory of gravity, and only of gravity, LQG has been brought to a completion yet. Most importantly, the dynamics of the theory—and thus the physical Hilbert space—, and the relationship to classical spacetimes theories remain ill-understood. Nevertheless, LQG is considered by many a very promising research program and is certainly the frontrunner of approaches starting out from classical GR.

In an earlier essay (Earman et al. (2009)), we stated, by way of conclusion, that LQG, like causal sets, simply ignores the possibility of CTCs as the canonical quantization procedure requires the spacetimes subjected to it to be globally hyperbolic, except in that as long as the classical limits of LQG states is so poorly understood, we cannot exclude that CTCs might emerge in this limit. This is a real possibility since in some cases, e.g. such as classically singular spacetimes, the classical spacetime structure does not even approximate the well-defined corresponding quantum state. We would like to add a further reason for hesitation. The above remarks tacitly assume that LQG aspires to describe the global structure of quantum spacetime. This need not be so: one might just as well think of the theory as offering descriptions of much more local features of quantum

⁶⁵The causal sets approach is still regarded as a classical theory so far, as it fails to provide a proper quantum dynamics. To our knowledge, the classical probabilities involved in the dynamical evolution according to the causal sets theory have so far not been replaced by truly quantum dynamics, including e.g. transition amplitudes.

⁶⁶For GR as a Hamiltonian system with constraints, cf. Thiemann (2007, Ch. 1) and Wüthrich (2006, Ch. 4).

spacetime, such as the spatial volume of a finite chunk of spacetime in a laboratory.⁶⁷ Of course, all these chunks of spacetime will be assumed globally hyperbolic. However, a spacetime patched together from globally hyperbolic spacetimes need not be globally hyperbolic itself.⁶⁸ If conceived in this way, therefore, LQG might well permit time travel. It should be noted, however, that this may mean that LQG cannot be a fundamental quantum theory of spacetime as it doesn't account for the global structure of spacetime. Unless one thinks that the global structure of spacetime emerges from or is supervenient on the fundamental structure of patched together chunks of quantum spacetime, and barring the possibility of CTCs emerging in the classical limit of a loop quantum gravitational state, either LQG cannot be a fundamental quantum theory of spacetime or it rules out time travel.

The third, and by far most researched, approach to quantum gravity is string theory (Zwiebach 2004). String theory takes as its vantage point, both historically and systematically, the standard model based on conventional QFT. It exists at two separate levels. At the perturbative level, on the one hand, string theory consists of a set of well-developed mathematical techniques which define the string perturbation expansion over a given background spacetime. On the other hand, attempts at formulating the elusive non-perturbative theory, supposed to be capable of generating the perturbation expansion, have not succeeded so far. Such a theory, conventionally named M-theory, for 'membrane,' 'matrix,' or 'mystery' theory, currently consists of but incipient formulations using non-perturbative compactifications of higher dimensional theories based on so-called duality symmetries, i.e. symmetries relating strong coupling limits in one string theory to a weak coupling limit in another (dual) string theory.

To the best of our knowledge, there are so far no results in non-perturbative M-theory pertaining to the possibility of time travel. There are, however, a series of pertinent findings in supersymmetric gravity, a close relative of string theory. If supersymmetric gravity turned out to permit time travel in the sense of CTCs, then string theory will be hard pressed to eschew its possibility.⁶⁹ As far as we are aware, the story begins ten years ago when Gibbons and Herdeiro (1999) asked whether supersymmetry allowed CTCs. The straightforward answer is that it does, at least in that there are supersymmetric solutions of flat space with periodically identified time coordinate analogous to the rolled-up Minkowski spacetime in GR. Solutions of this type, however, are not topologically simply-connected and the CTCs can thus be avoided by passing to a covering spacetime. In many cases of supersymmetric solutions with CTCs, this move is not possible since the relevant supersymmetric spacetimes are topologically trivial. It turns out that there are at least two important types of supersymmetric solutions containing CTCs: a supersymmetric cousin of Gödel spacetime and the so-called BMPV black hole spacetime. Let us look at these in turn.

Gauntlett et al. (2003) have shown that there exists a solution of five-dimensional supersymmetric gravity that is very similar to Gödel spacetime of GR in that it also describes a topologically trivial, rotating, and homogenous—and thus not asymptotically flat—universe. Almost by return of mail, however, Petr Hořava and collaborators have argued that holography acts as a form of chronology protection in the case of this Gödel-like spacetime in that the CTCs are either hidden behind 'holographic screens', and thus inaccessible for timelike observers, or that they are broken up into causally non-circular pieces (Boyda et al. 2003).

⁶⁷We are indebted to Carlo Rovelli for arguing, in private conversations, for the validity of this approach to one of us (C.W.).

⁶⁸For a simple example, just think of a rolled up slice of Minkowski spacetime, which contains CTCs and is thus not globally hyperbolic, as a carpet glued together from globally hyperbolic, diamonds-shaped tiles.

⁶⁹Many of the following results have been gained in five-dimensional supersymmetric gravity, rather than its higher-dimensional relatives. Five-dimensional supergravity is an approximation to higher-dimensional string theory. It should be noted that all solutions in the five-dimensional case can easily be amended to be solutions for ten- and eleven-dimensional supergravity (Gauntlett et al. 2003, 4590).

The second important supersymmetric spacetime containing CTCs is the so-called BMPV black hole solution.⁷⁰ BMPV black holes are the supersymmetric counterparts of the Kerr-Newman black holes of GR: They are charged, rotating black holes in simply connected, asymptotically flat spacetime. And similarly to the Kerr-Newman case in GR, as Gibbons and Herdeiro (1999) have shown, it can be maximally analytically extended to contain a region with CTCs.⁷¹ Below the critical value of angular momentum, we find a black hole with an event horizon, and CTCs in a region screened off by this horizon from the asymptotic observer (Gauntlett et al. 2003, 4589). Not only can the asymptotic observer not see the CTCs in this case, but they are inaccessible to her in that they are hidden behind a ‘velocity-of-light surface,’ i.e. a surface which can only be passed by accelerating beyond the speed of light. If the angular momentum is above the critical value, however, the black hole is shielded in the sense that geodesics from the asymptotic region cannot pass into the black hole (Gibbons and Herdeiro (1999)). The solution is thus geodesically complete. In this case, we find ‘naked’ CTCs outside the event horizon.⁷² As Gibbons and Herdeiro (1999) show, no cosmic censorship seems to be able to rule out this case: This hyper-critical solution represents a geodesically complete, simply connected, asymptotically flat, non-singular, time-orientable, supersymmetric spacetime with finite mass that satisfies the dominant energy condition.

Gibbons and Herdeiro note in their analysis, however, that this hyper-critical solution describes a situation where the CTCs have existed ‘forever,’ i.e. it seems not amenable to an implementation of a time machine in our sense. The reason for this is that passing from the ‘under-rotating’ case to the ‘over-rotating’ one seems to require an infinite amount of energy. As Dyson (2004) has shown in her analysis of BMPV black holes made out of gravitational waves and D-branes, i.e. hypersurfaces in ten-dimensional spacetime, speeding up the rotation of an ‘under-rotating’ BMPV black hole in order to produce naked CTCs leads to the formation of a shell of gravitons with the D-branes enclosed inside the black hole. This mechanism, which is akin to the ‘enhanced mechanism’ that string theorists use to block a class of naked singularities, precludes the system from speeding up beyond the critical value.

Although this may rule out the practicability of time travel in BMPV spacetimes, Gauntlett et al. (2003) have more good news for the aspiring time traveller. In their classification of all supersymmetric solutions of minimal supergravity in five dimensions, they find that CTCs generically appear in physically important classes of solutions. In fact, they complain that it is difficult to find any new solutions of five-dimensional supersymmetric gravity that do *not* contain either CTCs or singularities. While all of these results are somewhat provisional in that we do not know whether they translate to the full, non-perturbative M-theory, the possibility of time travel and perhaps of time machines will likely stay with the foundations of physics for some time to come.

9 Conclusions

In conclusion, let us return to the question posed in the introduction: what does the existence of solutions to Einstein’s field equations with exotic causal structure imply regarding the nature of space and time according to GR; or, more generally, whether physics permits such exotic causal structures, and if so, what does this permission mean for the nature of space and time? Our first focus has been on the implications of time travel, defined in terms of the existence of CTCs.

⁷⁰After the initials of Breckenridge et al. (1997).

⁷¹Strictly speaking, Gibbons and Herdeiro show this for the *extremal* case, i.e. black holes whose angular momentum is equal to its mass (in natural units). But the result is thought to generalize to the non-extremal case.

⁷²It has been suggested that the emergence of ‘naked’ CTCs may be the result of the breakdown of unitarity (Herdeiro 2000).

Many philosophers have attempted to dismiss this question as illegitimate, on the grounds that a variety of paradoxes establish the logical impossibility, metaphysical impossibility, or improbability of time travel in this sense. We found these arguments wanting, although they do usefully illustrate the importance of consistency constraints in spacetimes with CTCs. It may come as a shock to discover [that](#) the consistent time-travel scenarios are not just the stuff of fiction: there are several chronology-violating spacetimes that exhibit the local-to-global property described in §3 for appropriate choices of fields. However shocking the existence of these solutions may be, we assert that there is no footing to reject them due to alleged paradoxes, and no basis for imposing a causality condition insuring “tame” causal structure as an *a priori* constraint.

Setting aside objections based on the paradoxes, attempting to answer our question leads into a tangle of interconnected issues in philosophy of science and the foundations of GR. We hope to have at least clearly identified some of these issues and illustrated how their resolution contributes to an answer. First, consider cosmological models such as Gödel’s that are not viable models for the structure of the observed universe. Assessing the importance of these models turns on difficult questions of modality applied to cosmology. Even if we grant that GR provides the best guide to what is physically possible in cosmology, the existence of models like Gödel’s does not directly undermine the use of special structures such as the preferred foliation in the FLRW models without a questionable modal argument or claim that such models reveal something significant about the laws of GR. Thus, if we were only considering cosmological models with exotic causal structure, it would be difficult to answer Maudlin’s challenge. Maudlin claims that metaphysicians can safely dismiss exotic spacetimes because dynamical evolution according to Einstein’s field equations does not force CTCs to arise from possible initial data. But this assertion presumes a resolution of a second open issue, the cosmic censorship conjecture or [\(some form of\)](#) the weaker chronology protection conjecture. Given a proof of the cosmic censorship conjecture, one could clearly demarcate situations in which Einstein’s field equations coupled to source equations satisfying constraints such as the energy conditions generically lead to globally hyperbolic spacetimes from situations in which dynamical evolution leads to Cauchy horizons, and the possibility of extensions beyond them containing CTCs. There are still significant obstacles to a proof of cosmic censorship due to our lack of understanding of the space of solutions to GR. [Similarly, a proof of a sufficiently powerful chronology protection conjecture imposing some principled conditions on a spacetime’s properties would underwrite Maudlin’s claims. Alas, this second issue remain open to date, not least because it is far from obvious how the blanks in Conjecture 1 concerning suitable initial data and physically reasonable spacetimes ought to be filled in.](#)

A third issue concerns the impact of incorporating quantum effects. Does the space of solutions of semi-classical quantum gravity, or even full quantum gravity, include TM solutions or solutions with CTCs? [Thus, our investigation went beyond a mere analysis of the foundations of GR, in at least two respects. First, we have turned to semi-classical quantum gravity and listed how the quantum can be more permissive in tolerating the violation of energy conditions and thus be more lax about the suitability of the matter sector. Although no one really takes semi-classical theories seriously as competitors for final theories of quantum gravity, important lessons of how spacetime and quantum matter interact may be gleaned from them. Second, in a brief survey of three approaches to full quantum gravity, causal set theory, loop quantum gravity, and string theory, we have found that string theory in particular seems to nourish the hopes of aspiring time travellers, while one shouldn’t be too hasty in ruling time travel out in the case of loop quantum gravity. These results are very preliminary and much remains to be seen, not the least of which is whether any of the mentioned theories can offer a full quantum theory of gravity. But we hope that the reader walks away from this article with a firm sense that these foundational analyses in GR, semi-classical, and full quantum gravity constitute important attempts at both understanding](#)

the classical theory, as well as illuminating the path towards a quantum theory of gravity.

Acknowledgements

We wish to thank the Southern California Reading Group in the Philosophy of Physics for valuable feedback. C.W. gratefully acknowledges support for this project by the Swiss National Science Foundation (Project “Properties and Relations”, grant 100011-113688).

References

- [1] Andréka, H., Németi, I., and Wüthrich, C. (2008), ‘A twist in the geometry of rotating black holes: seeking the cause of acausality’, *General Relativity and Gravitation* 40: 1809-1823.
- [2] Arntzenius, F. (2006), ‘Time travel: double your fun’, *Philosophy Compass* 1: 599-616.
- [3] Arntzenius, F., and Maudlin, T. (2005), ‘Time travel and modern physics’, in E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/time-travel-phys/>.
- [4] Barceló, C. and Visser, M. (2002), ‘Twilight for the energy conditions?’, *International Journal of Modern Physics D* 11: 1553-1560.
- [5] Belot, G. (2005), ‘Dust, Time, and Symmetry’, *British Journal for the Philosophy of Science* 56: 255-91.
- [6] Black, M. (1956), ‘Why cannot an effect precede its cause?’, *Analysis* 16: 49-58.
- [7] Bombelli, L., Lee, J., Meyer, D., and Sorkin, R.D. (1987), ‘Space-time and a causal set’, *Physical Review Letters* 59: 521524.
- [8] Bonnor, W. B. (2001), ‘The interactions between two classical spinning particles’, *Classical and Quantum Gravity* 18: 1381-1388.
- [9] Boyda, E.K., Ganguli, S., Hořava, P., and Varadarajan, U. (2003), ‘Holographic protection of chronology in universes of the Gödel type’, *Physical Review D* 67: 106003.
- [10] Breckenridge, J.C., Myers, R.C., Peet, A.W., and Vafa, C. (1997), ‘D-branes and spinning black holes’, *Physics Letters B* 391: 93-98.
- [11] Brightwell, G., Dowker, H.F., Garcia, R.S., Henson, J., and Sorkin, R.D. (2003), ‘“Observables” in causal set cosmology’, *Physical Review D* 67: 08403.
- [12] Callender, C. and Weingard, R. (2000), ‘Topology change and the unity of space’, *Studies in History and Philosophy of Modern Physics* 31: 227-246.
- [13] Deutsch, D. (1991), ‘Quantum mechanics near closed timelike lines’, *Physical Review D* 44: 3197-3217.
- [14] Dorato, M. (2002), ‘On becoming, cosmic time, and rotating universes’, in C. Callender (ed.), *Time, Reality, and Existence* (Cambridge University Press), 253-76.
- [15] Douglas, R. (1997) ‘Stochastically branching spacetime topology’, in S. Savitt (ed.), *Time’s Arrow Today* (Cambridge University Press), 173-190.
- [16] Dummett, M. (1964), ‘Bringing about the past’, *Philosophical Review* 73: 338-359.
- [17] Dyson, L. (2004), ‘Chronology protection in string theory’, *Journal of High Energy Physics* 03(2004): 024.
- [18] Earman, J. (1995), *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes* (Oxford University Press).
- [19] Earman, J. (2008), ‘Pruning some branches from ‘branching spacetimes’ ’, in Dieks, D. (ed.), *The Ontology of Spacetime II* (Elsevier).

- [20] Earman, J., Smeenk, C., and Wüthrich, C. (2009), ‘Do the laws of physics forbid the operation of time machines?’, *Synthese*, 169: 91-124.
- [21] Earman, J. and Wüthrich, C. (2004), ‘Time machines’, in E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/time-machine/>.
- [22] Einstein, A. (1936), ‘Physik und Realität’, *Journal of the Franklin Institute* 221: 313-337. Translated by S. Bargmann as “Physics and reality”, in Einstein, A., *Ideas and Opinions* (Crown Publishers), 290323.
- [23] Ellis, G. F. R., ‘Contributions of K. Gödel to Relativity and Cosmology’, in P. Hájek (ed.), *Gödel ’96: Logical Foundations of Mathematics, Computer Science and Physics – Kurt Gödel’s Legacy*, (Berlin: Springer-Verlag), 34-49.
- [24] Fewster, C.J. (2005), ‘Energy inequalities in quantum field theory’, math-ph/0501073.
- [25] [Fewster, C.J. and Higuchi, A. \(2008\)–I can’t find this paper...](#)
- [26] Flanagan, E. and Wald, R (1996), ‘Does backreaction enforce the averaged null energy condition in semiclassical gravity?’, *Physical Review D* 54: 6233-6283.
- [27] Ford, L.H. (1978), *Proceedings of the Royal Society London A* 364: 227.
- [28] Friedman, J.L. (2004), ‘The Cauchy problem on spacetimes that are not globally hyperbolic’, in Chrusciel, P.T. and Friedrich, H. (eds.), *The Einstein Equations and the Large Scale Behavior of Gravitational Fields: 50 Years of the Cauchy Problem in General Relativity* (Birkhäuser), 331-346.
- [29] Friedman, J.L. and Higuchi, A. (2006), ‘Topological censorship and chronology protection’, *Annalen der Physik* 15: 109-128.
- [30] Friedman, J.L., Morris, M.S., Novikov, I.D., Echeverria, F., Klinkhammer, G., Thorne, K.S., and Yurtsever, U. (1990), ‘Cauchy problem in spacetimes with closed timelike curves’, *Phys. Rev. D* 42: 1915-1930.
- [31] Friedman, J.L., Schleich, K., and Witt, D.M. (1993), ‘Topological censorship’, *Physical Review Letters* 71: 1486-1489.
- [32] Galloway, G.J. (1995), ‘On the topology of the domain of outer communication’, *Classical and Quantum Gravity* 12: L99-101.
- [33] Gannon, D. (1975), ‘Singularities in nonsimply connected space-times’, *J. Math. Phys.* 16: 2364-2367.
- [34] Gauntlett, J.P., Gutowski, J.B., Hull, C.M., Pakis, S., and Reall, H.S. (2003), ‘All supersymmetric solutions of minimal supergravity in five dimensions’, *Classical and Quantum Gravity* 20: 4587-4634.
- [35] Geroch, R. (1967), ‘Topology in general relativity’, *Journal of Mathematical Physics* 8: 782-786.
- [36] Geroch, R. (1977), ‘Prediction in general relativity’, in J. Earman, C. Glymour, and J. Stachel (eds.), *Foundation of Spacetime Theories, Minnesota Studies in the Philosophy of Science VIII* (University of Minnesota Press), 81-93.

- [37] Geroch, R., and Horowitz, G. (1979), ‘Global structure of spacetimes’, in S. Hawking and W. Israel (eds.), *General Relativity: An Einstein Centenary Survey* (Cambridge University Press), 212-293.
- [38] Gibbons, G.W. and Herdeiro, C.A.R. (1999), ‘Supersymmetric rotating black holes and causality violation’, *Classical and Quantum Gravity* 16: 3619-3652.
- [39] Gödel, K. (1949a), ‘An example of a new type of cosmological solutions of Einstein’s field equations of gravitation’, *Review of Modern Physics* 21: 447-450.
- [40] Gödel, K. (1949b), ‘A remark about the relationship between relativity theory and idealistic philosophy’, in P.A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist* (Open Court), 557-562.
- [41] Gödel, K. (1952), ‘Rotating universes in general relativity theory’, in L.M. Graves et al. (eds.), *Proceeding of the International Congress of Mathematicians*, (American Mathematical Society), 175-181.
- [42] Hawking, S.W. (1992), ‘Chronology protection conjecture’, *Physical Review D* 46: 603-611.
- [43] Hawking, S.W., and Ellis, G.F.R., (1973), *The Large Scale Structure of Space-time* (Cambridge University Press).
- [44] Hawking, S.W., and Penrose, R., (1996), *The Nature of Space and Time*, (Princeton University Press).
- [45] Herdeiro, C.A., (2000), ‘Special properties of five dimensional BPS rotating black holes’ *Nuclear Physics B* 582: 363-392.
- [46] Horwich, P. (1987), *Asymmetries in Time: Problems in the Philosophy of Science*, (MIT Press).
- [47] Kay, B. S., Radzikowski, M. J., and Wald, R.M. (1997), ‘Quantum field theory on spacetimes with compactly generated cauchy horizons’, *Communications in Mathematical Physics* 183: 533-556.
- [48] Kim, S.-W. and Thorne, K.S. (1991), ‘Do vacuum fluctuations prevent the creation of closed timelike curves?’, *Physical Review D* 43: 3929-3949.
- [49] Kogut, A., Hinshaw, G., and Banday, A.J. (1997), ‘Limits to global rotation and shear from the COBE DMR four-year sky maps’, *Physical Review D* 55: 1901-5.
- [50] Krasnikov, S. (2002), ‘No time machines in classical general relativity’, *Classical and Quantum Gravity* 19: 4109-4129.
- [51] Lewis, D. (1976), ‘The paradoxes of time travel’, *American Philosophical Quarterly* 13: 145-152. Reprinted in *Philosophical Papers*, Volume II (Oxford University Press), 67-80.
- [52] Malament, D. (1985a), ‘Minimal acceleration requirements for “time travel” in Gödel spacetime’, *The Journal of Mathematical Physics* 26: 774-777.
- [53] Malament, D. (1985b), ‘“Time travel” in the Gödel universe’, in P.D. Asquith and P. Kitcher (eds.), *PSA 1984*, Vol. 2, East Lansing: Philosophy of Science Association, 91-100.

- [54] Malament, D. (1995), ‘Introductory note to *1949b’, in S. Feferman et al. (eds.), *Kurt Gödel: Collected Works, Volume III* (Oxford University Press), 261-269.
- [55] Malament, D. (2002), ‘A no-go theorem about rotation in relativity theory’, in Malament, D. (ed.), *Reading Natural Philosophy (Essays Dedicated to Howard Stein on His 70th Birthday)*, (Open Court Press), 267-293.
- [56] Manchak, J.B. (2009a), ‘Is spacetime hole-free?’, *General Relativity and Gravitation*, in print.
- [57] Manchak, J.B. (2009b), ‘On the existence of “time machines” in general relativity’, *Philosophy of Science*, in print.
- [58] Mattingly, J. (2001), ‘Singularities and scalar fields: Matter fields and general relativity’, *Philosophy of Science* 68: S395-S406.
- [59] Meiland, J.W. (1974), ‘A two-dimensional passage model of time for time travel’, *Philosophical Studies* 26: 153-173.
- [60] Mellor, D.H. (1981), *Real Time* (Cambridge University Press).
- [61] [Moncrief, V. and Isenberg, J. \(1983\), ‘Symmetries of cosmological Cauchy horizons’, *Communications in Mathematical Physics* 89: 387-413.](#)
- [62] Monton, B. (2009), ‘Time travel without causal loops’, *Philosophical Quarterly* 59: 54-67.
- [63] Morris, M.S. and Thorne, K.S. (1988), ‘Wormholes in spacetime and their use for interstellar travel: A tool for teaching general relativity’, *American Journal of Physics* 56: 395-412.
- [64] [Morris, M.S. and Thorne, K.S. \(1991\), ? And Yurtsever?](#)
- [65] Morris, M.S., Thorne, K.S. and Yurtsever, U. (1988), ‘Wormholes, Time Machines, and the Weak Energy Condition’, *Physical Review Letters* 61: 1446-1449.
- [66] Ori, A. (2007), ‘Formation of closed timelike curves in a composite vacuum/dust asymptotically flat spacetime’, *Physical Review D* 76: 044002.
- [67] Penrose, R. (1969), ‘Gravitational collapse: the role of general relativity’, *Rivista del Nuovo Cimento* 1: 252-276 (Numero speciale).
- [68] [Penrose, R. \(1979\),](#)
- [69] Politzer, H. D. (1992), ‘Simple quantum systems in spacetimes with closed timelike curves’, *Physical Review D* 46: 4470-4476.
- [70] Price, H. (1984), ‘The philosophy and physics of affecting the past’, *Synthese* 16: 299-323.
- [71] Rendall, A. (2008), *Partial Differential Equations in General Relativity*, (Oxford University Press).
- [72] Rovelli, C. (2004), *Quantum Gravity* (Cambridge University Press).
- [73] Savitt, S. (1994), ‘The replacement of time’, *Australasian Journal of Philosophy* 72: 463-474.
- [74] Smith, N.J.J. (1997), ‘Bananas enough for time travel?’, *British Journal for the Philosophy of Science* 48: 363-389.

- [75] Smith, N.J.J. (1998), ‘The problems of backward time travel’, *Endeavour* 22: 156-158.
- [76] Stein, H. (1970), ‘On the paradoxical time-structures of Gödel’, *Philosophy of Science*, 37: 589-601.
- [77] Stein, H. (1995), ‘Introductory note to *1946/9’, in S. Feferman et al. (eds.), *Kurt Gödel: Collected Works, Volume III* (Oxford University Press), 202-229.
- [78] Thiemann, T. (2007), *Modern Canonical Quantum General Relativity* (Cambridge University Press).
- [79] Thorne, K.S. (1994), *Black Holes and Time Warps: Einstein’s Outrageous Legacy* (W.W. Norton and Company).
- [80] Tipler, F.J. (1974), ‘Rotating cylinders and the possibility of global causality violation’, *Phys. Rev. D* 9: 2203-2206.
- [81] Tipler, F. J. (1976), ‘Causality violation in asymptotically flat space-times’, *Physical Review Letters* 37: 879-882.
- [82] Tipler, F. J. (1977), ‘Singularities and causality violation’, *Annals of Physics* 108: 1-36.
- [83] van Stockum, W.J. (1937), ‘The gravitational field of a distribution of particles rotating about an axis of symmetry’, *Proc. R. Edinb.* 57: 135-154.
- [84] Visser, M. (2003), ‘The quantum physics of chronology protection’, in G.W. Gibbons, E.P.S. Shellard, and S.J. Rankin (eds.), *The future of theoretical physics and cosmology: celebrating Stephen Hawking’s 60th birthday* (Cambridge University Press), 161-176.
- [85] Wald, R.M. (1984), *General Relativity*, (The University of Chicago Press).
- [86] Wald, R.M. (1998), ‘Gravitational collapse and cosmic censorship’, in B.R. Iyer and B. Bhawal (eds.), *Black Holes, Gravitational Radiation and the Universe: Essays in Honor of C. V. Vishveshwara* (Kluwer Academic), 69-85.
- [87] Weyl, H. (1921), *Space-Time-Matter*, translated by S. Brose, (Methuen).
- [88] Wheeler, J. and Feynman, R. (1949), ‘Classical electrodynamics in terms of direct interparticle action’, *Reviews of Modern Physics* 21: 425-434.
- [89] Witt, D.M. (1986), ‘Vacuum space-times that admit no maximal slice’, *Physical Review Letters* 57: 1386-1389.
- [90] Wüthrich, C. (1999), *On time machines in Kerr-Newman spacetimes*, Master thesis, University of Bern. <http://philosophy.ucsd.edu/faculty/wuthrich/pub/WuthrichChristian1999MScThesis.pdf>
- [91] Wüthrich, C. (2006), *Approaching the Planck Scale From a Generally Relativistic Point of View: A Philosophical Appraisal of Loop Quantum Gravity*, PhD thesis, University of Pittsburgh. <http://philosophy.ucsd.edu/faculty/wuthrich/pub/WuthrichChristianPhD2006Final.pdf>
- [92] Yourgrau, P. (1991), *The Disappearance of Time: Kurt Gödel and the Idealistic Tradition in Philosophy* (Cambridge University Press).
- [93] Zwiebach, B. (2004), *A First Course in String Theory* (Cambridge University Press).