Causality and Time in Non-smooth Lorentzian Geometry

Leonardo García Heveling

Causality and Time in Non-smooth Lorentzian Geometry

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 26 september 2023
om 16.30 uur precies

door

Leonardo García Heveling

geboren op 10 januari 1995
te Vilassar de Mar, Spanje

Promotor:
prof. dr. N.P. Landsman

Copromotor:
dr. A.Y. Burtscher

Manuscriptcomissie:

prof. dr. R. Loll	(voorzitter)
prof. dr. R.J. McCann	(University of Toronto, Canada)
dr. A. Sakovich	(Uppsala Universitet, Zweden)
prof. dr. M. Sánchez Caja	(Universidad de Granada, Spanje)
dr. S. Suhr	(Ruhr-Universität Bochum, Duitsland)

Contents

Chapter 1

Introduction

This thesis consists of a collection of research articles in the area of Lorentzian geometry, written in the years 2020–2023. Lorentzian geometry provides the framework for the physical theory of general relativity, which describes gravity through the geometry of spacetime. We start with a historical overview.

1.1 From Newton to Einstein

The first fundamental theory of gravity was due to Newton. In his theory, space is modelled by the Euclidean space $\mathbb{R}^3$. The gravitational force between two masses m and M at positions $\vec{x}$ and $\vec{y}$ is given by *Newton's universal law of gravitation*

$$\vec{F} = G\frac{mM}{|\vec{x} - \vec{y}|^2}, \tag{1.1}$$

where G is the gravitational constant. The time evolution of the position $\vec{x}$ of the mass m can then be obtained from *Newton's second law of motion*

$$m\frac{d^2\vec{x}}{dt^2} = \vec{F}. \tag{1.2}$$

In Newton's theory, it is implicit that any observer measures spatial distances in the same way, and experiences time passing in the same way. In particular, it can be established *objectively* (independently of the observer) whether two events are happening simultaneously, or if one occurs before the other.

The point of view implied by Newton's mechanics is incompatible with the physical principle that the speed of light is the same for any two observers, even if one of them is moving with respect to the other. This fact was established theoretically by Maxwell, who showed that his equations of electromagnetism allow wavelike solutions, interpreted as light waves, whose speed of propagation is a fundamental constant of the theory. In a successful attempt to merge Maxwell's equations with the laws of mechanics, Einstein proposed the theory of special relativity. Shortly after, Minkowski reformulated special relativity in a geometrical language, by merging space and time into what we now call Minkowski spacetime, which is the vector space $\mathbb{R}^4$ equipped with the indefinite scalar product

$$\eta = -c^2 dt^2 + dx^2 + dy^2 + dz^2, \tag{1.3}$$

where c is the speed of light. The set of vectors $v \in \mathbb{R}^4$ with $\eta(v, v) = 0$ is called the lightcone, and depicted in Figure 1.1.

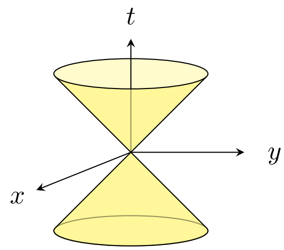


Figure 1.1: The lightcone.

The radical innovation of special relativity is that the transformations between different coordinate systems (corresponding to different inertial observers) are *Lorentz transformations*. The latter are precisely the linear maps leaving (1.3) invariant, and they mix the time and space coordinates. Therefore, the notion of "simultaneity" can no longer be objectively defined.

A consequence of special relativity is that Newton's law of gravitation must be superseded, among other reasons because the distance $|\vec{x} - \vec{y}|$ would be observer-dependent. It turns out that another leap of abstraction is needed, namely to substitute $\mathbb{R}^4$ by a manifold M with a (point-dependent) metric tensor g that replaces the Minkowski metric (1.3). This is precisely the premise of general relativity. The metric tensor g is of Lorentzian signature, meaning that it still looks like (1.3) in each tangent space (see also Figure 1.1). The *Einstein equations*

$$\mathrm{Ric} - \frac{1}{2} R g = \frac{8\pi G}{c^4} T \tag{1.4}$$

relate the Ricci curvature of the metric to the matter content of the Universe (encoded in the energy-momentum tensor T), replacing Newton's universal law of gravitation (1.1). The trajectories of unaccelerated point particles in spacetime are given by geodesics, replacing (1.2). As John Wheeler elegantly summarized [MTW73, p. 5],

"Spacetime tells matter how to move; Matter tells spacetime how to curve."

Precisely speaking, we call the pair (M, g), together with a continuous choice of future vs. past lightcone at every point, a spacetime (in this introduction, we also use the word spacetime to describe the physical entity). Points on M have the physical interpretation of events. Since cause-and-effect relationships are fundamental to science, it is natural to ask

When can a spacetime event p causally affect another event q?

In Newtonian mechanics, the answer would be "if p happens before q", but already in special relativity, this no longer makes sense. Instead, we need to enforce that causal interactions cannot propagate faster than the speed of light. In Minkowski spacetime (1.3), this means that an event can only influence its future solid lightcone, and can only be affected by its past solid lightcone (the upper and lower cones in Figure 1.1, respectively). The idea remains the same in general relativity, where Figure 1.1 still applies locally. Globally, however, the non-trivial topology and geometry of (M, g) allow for a much richer causal structure. In particular, the answer to the above question is

When p and q can be joined by a curve whose tangent vector lies in the future lightcone at every point.

This condition induces two relations on M, depending on whether we include the boundary of the lightcone (corresponding to trajectories at the speed of light, such as those of light rays themselves) or not (corresponding to trajectories of massive particles). The causal structure is one of the cornerstones of general relativity, and of this thesis in particular.

1.2 Non-smooth Lorentzian geometry

In the previous section, we described how manifolds with smooth Lorentzian metrics have been successfully used to model spacetime in general relativity. In some cases, however, it is necessary to take a broader view, and consider more general mathematical structures. This is the starting premise of this thesis, and here we give an overview and motivation of the kinds of structures considered. The discussion of the results is postponed to Section 1.4, after some detailed concepts have been introduced in Section 1.3.

1.2.1 Continuous Lorentzian metrics

In classical general relativity, it is usually assumed that the metric tensor g on a spacetime (M, g) is smooth. A partial differential equation, however, can also admit *weak solutions*, which in the case of the Einstein equations (1.4) correspond to non-smooth metric tensors. Many different regularity classes can be considered; in this thesis, the class C^0 of continuous metrics plays a prominent role. From a physical point of view, the motivation for considering weak solutions to (1.4) is twofold.

On the one hand, one may consider matter models where the energy-momentum tensor T is non-smooth, or even distributional, leading thus to a non-smooth solution of the Einstein equations [BL14; Chr95; GT87; JM19]. This can be useful, for instance, as an idealization: say, one wants to model a star by a constant density that abruptly drops to zero on the star's surface.

On the other hand, weak solutions to (1.4) can be considered also in a merely hypothetical way. In particular, the cosmic censorship conjectures make statements about generic solutions to the Einstein equations; Allowing weak solutions broadens the scope of these conjectures (see page 13 for more details).

The study of the causal structure of continuous spacetime metrics was initated by Chruściel and Grant [CG12], then further developed by Grant et al. [Gra+20], Sämann [Säm16] and in Chapters 2 and 3 of this thesis. Other works on contiuous metrics, in particular in relation to extendibility questions and the strong cosmic censorship conjecture, include [Chr99; DL17; GLS18; Lin20; Sbi18; Sbi22].

1.2.2 Synthetic Lorentzian geometry

Let us start by making an analogy to Riemannian geometry. A Riemannian manifold (M, g) has a positive-definite metric tensor g (signature $(+, ..., +)$), in other words, an inner product on each tangent space. This naturally endows it with the structure of a length space. In particular, the length of a piecewise smooth curve $\gamma\colon [a, b] \to M$ is given by

$$L_g(\gamma) := \int_a^b \sqrt{\left|g_{\gamma(s)}\left(\dot\gamma(s), \dot\gamma(s)\right)\right|} ds, \tag{1.5}$$

and the distance between two points p, q is defined as the infimum

$$d_g(p, q) := \inf_\gamma L_g(\gamma) \tag{1.6}$$

over all such curves γ that join p and q. Viewing Riemannian manifolds as length spaces is useful, for instance, because there exist good notions of when a sequence of length spaces converges to a limit space, such as Gromov–Hausdorff convergence. In general, even a sequence of manifolds does not converge to a manifold in the limit; think, for instance, of a sequence of rounded cones converging to a cone with a sharp tip, as in Figure 1.2.

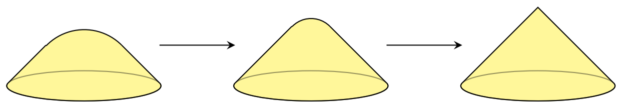


Figure 1.2: Convergence towards a non-smooth limit.

Developing a notion of convergence of spacetimes is highly desired, and currently an active field of research (see [Sor18], where Yau is credited as the first to propose this). In particular, it can be used to make "stability" questions precise, since it gives meaning to the idea of two spacetimes being "close" or "similar" to each other. Consider, for example, the case of cosmology, which studies the large scale properties of the Universe. In this setting, it is often assumed that the distribution of matter in the Universe is homogeneous. Clearly, this can only be true in an approximate sense, since we know that matter is actually clustered into planets, galaxies, etc. The idea, of course, is that "from far away, it looks homogeneous". Making this statement mathematically precise is still a matter of current research.

With this motivation in mind, and also simply in order to have a unified, axiomatic approach to (non-smooth) Lorentzian geometry, one seeks to find a Lorentzian analogue of a length space. Two distinct philosophies stand out, which we now review; both of them are investigated in this thesis. In between, we have a brief digression about lower Ricci curvature bounds, which also plays an important role in this context.

Lorentzian length spaces

On a spacetime (M, g), the *time separation function* (also called *Lorentzian distance*) is defined as

$$d_g(p, q) := \sup_\gamma L_g(\gamma), \tag{1.7}$$

where L_g is given by (1.5), and the supremum is over all future-directed causal curves γ (i.e. $\dot\gamma$ is in the future lightcone) joining $p, q \in M$, with the convention that it is 0 if none exists.

The Lorentzian length $L_g(\gamma)$ has the physical interpretation of *proper time* of an observer with trajectory γ. In other words, an observer travelling from event p to event q along γ, experiences an amount $L_g(\gamma)$ of time in between. The time separation (1.7) is thus the longest possible time an observer can spend between events p and q. The fact that different curves connecting the same two events can have different lengths, gives rise to phenomena such as the twins paradox. It is not really a paradox, but rather just the observation that a person travelling in a spaceship might have aged less upon its return than their twin sibling who stayed on Earth.

Unlike its Riemannian analogue (1.6), the time separation function is not a distance in the usual sense. It is antisymmetric, it obeys the inverse triangle inequality, and $d_g(p, q) = 0$ does not imply $p = q$. Lorentzian length spaces in the sense of Kunzinger and Sämann [KS18] are essentially spaces (not necessarily manifolds) equipped with a causal relation and a time separation function with precisely these properties (see Definitions 4.5 and 4.14 for the exact details).

The study of Lorentzian length spaces is by now a very lively field. Their causal structure is been investigated in [ACS20; Rot22] and Chapter 4. Both sectional [KS18] and Ricci [CM20] curvature bounds have been introduced (further discussed below), and their consequences investigated [BOS22; BNR23; BR22; BS22; Ber+22]. Concrete constructions of Lorentzian length spaces are found in [AH+22; Ale+19; BOS23; BC22] and Chapter 5. Notions of Lorentzian Gromov–Hausdorff distance are also under development [BN04; MS22b; Mül22a; Nol04a; Nol04b], not necessarily using the exact same axioms as Kunzinger and Sämann, but following the same philosophy.

Ricci curvature bounds and optimal transport

Certain inequalities involving the Ricci tensor appear naturally in general relativity as *energy conditions*, for instance as assumptions in the celebrated singularity theorems of Hawking and Penrose [HE73, Ch. 8.2]. Essentially, energy conditions are used whenever one does not wish to work with the full Einstein equations (1.4), but still needs to impose restrictions on the Ricci curvature based on physical properties of matter in the Universe. We refer to [KS20] for a modern review on physical aspects of energy conditions.

McCann proved that the *strong energy condition*

$$\mathrm{Ric}(v, v) \geq 0 \text{ for all } v \text{ in the lightcone} \tag{1.8}$$

is equivalent to a convexity property of a certain entropy functional on the space of probability measures supported on the spacetime [McC20], using tools from optimal transport (see also [EM17]). This result thus characterizes Ricci curvature bounds without any reference to the Ricci tensor itself, using only the time separation (1.7) and the canonical volume element induced by g. McCann's approach is based on earlier work in the Riemannian setting [CEMS01; RS05], and even the Einstein equations (1.4) have by now been formulated in this language by Mondino and Suhr [MS22c].

By turning McCann's theorem into a definition, lower Ricci bounds in the optimal transport sense have since been extended by Cavalletti and Mondino to measured Lorentzian length spaces, where no Ricci tensor is available [CM20] (see also [Bra22; Bra23; CM22; MS22a]). Their work is in the spirit of an earlier generalization from Riemannian manifolds to metric spaces [LV09; Stu06]. We consider such synthetic lower Ricci bounds on spacetimes with continuous metrics (which also lack a well-defined Ricci tensor) in Chapter 3. Note that in addition to the strong energy condition, recently also the *null* energy condition (which requires (1.8) only on the boundary of the lightcone) has been characterized in this language by Ketterer [Ket23] (on smooth spacetimes) and McCann [McC23].

Null distance

Instead of developing an Lorentzian analogue of Gromov–Hausdorff convergence based on the time separation function, one can also equip a Lorentzian manifold with a metric space structure, and apply the existing convergence machinery. In other words, one is looking for a functor from the category of spacetimes to the category of metric spaces (both possibly with additional restrictions or structure). This categorical formulation of the problem is due to Müller [Mül22b], who also proposes such a functor. In this thesis, however, we work with an earlier construction due to Sormani and Vega [SV16].

While one can always find *some* Riemannian metric on a given manifold, and despite the fact that this is often a useful tool in Lorentzian geometry, it is not satisfying for the purpose at hand. Instead, what we are looking for is a positive-definite metric that manages to capture some (or all) of the Lorentzian information, such as the causal relationships between events, or the curvature of the Lorentzian metric.

A construction that achieves this is the null distance of Sormani and Vega [SV16]. Convergence of a certain class of spacetimes has been studied in this context by Allen and Burtscher [AB22], while the relationship of the null distance with the causal structure of spacetime is investigated in the paper [SS23] by Sakovich and Sormani and in Chapter 6 of this thesis.

1.2.3 Quantum gravity

Around the same time that general relativity was developed, another revolution was happening in physics: quantum mechanics. Classically, matter is assumed to follow deterministic trajectories given by an equation of motion, and every observable quantity can be computed for any point on the trajectory. In quantum mechanics, this is radically different: The state of matter is represented by a vector in a Hilbert space, and observable quantities are represented by operators on this Hilbert space. If an observation is made, the outcome is an eigenvalue of the corresponding operator. According to the Copenhagen interpretation, measurements are inherently random, although very accurate predictions can be made of probabilities of different outcomes. Moreover, the state of the physical system is altered by the measurement: The new state after the observation is an eigenvector corresponding to the observed eigenvalue. Note that there are still open questions and controversies, broadly referred to as the *measurement problem*, which lie beyond the scope of this thesis.

Recall that in general relativity, the matter content of the Universe is encoded in the energy-momentum tensor T (see (1.4)), which is defined pointwise (or distributionally, see Section 1.2.1) on the spacetime. This is in conflict with the quantum mechanical way of measuring observables like the energy and the momentum, described above. It turns out that resolving this conflict is very hard, and remains one of the major open problems in theoretical physics. Incidentally, one of the main difficulties is the *problem of time*, i.e., the fact that time in quantum physics is essentially an external parameter, while in general relativity, it arises dynamically as part of spacetime.

Theories of quantum gravity generally do away with the Einstein Equations (1.4), just as quantum mechanics does away with deterministic equations of motion for particles and fields. In this context, a possible replacement for an equation of motion is a *path integral*. This approach was pioneered in particle physics by Feynman [Fey48], who computed the probability of a particle moving from a location x to another location y by integrating over all classical paths, weighted by the action of the path. In path-integral quantum gravity, the idea is to consider 3-dimensional Riemannian manifolds, roughly playing the role of outcomes of an observation of space at a given moment in time. Then, one wants to find the probability of transitioning between two such 3-manifolds, say V_1, V_2, by computing a path-integral over all Lorentzian cobordisms with boundary $V_1 \sqcup V_2$ (see Figure 1.3), weighted by their Einstein–Hilbert action (whose Euler–Lagrange equations are the Einstein equations (1.4)).

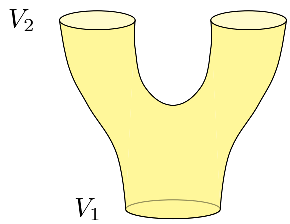


Figure 1.3: A cobordism.

So far, it has proven difficult to give precise meaning to the gravitational path integral. Somewhat analogously to how a Riemann integral is defined, a possible approach is to discretize or triangulate spacetime [AJL00; Bom+87; Wei82]. Both the interpretation that this is merely an approximation, or that spacetime is actually discrete at small scales, have been proposed. Note that in either case, a notion of “closeness” of spacetimes is desirable (spacetime here meant in a broad sense, see Section 1.2.2), as one wants to quantify how close our discrete/continuum approximation is to the physical continuum/discrete spacetime.

One of the features of path-integral quantum gravity is that, in principle, it is possible that the spatial topology changes, that is, to have a non-zero probability to transition between non-homeomorphic spaces V_1 and V_2 (as in Figure 1.3). Indeed, this possibility was suggested by Wheeler [Whe57], and Yodzis [Yod73] quotes Gell-Mann's totalitarian principle to support it:

"Anything which is not forbidden is compulsory."

We refer to Chapter 5 for a further discussion of this topic, where we study a class of topology changing spacetimes, the so-called Morse spacetimes. These are manifolds equipped with *degenerate* metric tensors, which are allowed to vanish at certain critical points. Regardless, one can define a causal structure that includes those critical points, and induces the structure of a Lorentzian length space.

Finally, let us briefly comment on Euclidean quantum gravity. This is a variation of the path-integral approach described above, where instead of Lorentzian metrics, one uses Riemannian ones. The motivation comes from the idea of Wick-rotation, where one constructs a Riemannian metric out of a Lorentzian one by "flipping the $-$ sign to a $+$ sign". This construction, however, is not uniquely defined. In Chapter 6, we use Wick rotation as a tool in some proofs, and along the way, discover that the null distance of Sormani and Vega is, in a sense, a generalization of it.

1.3 Causal theory

The main focus of this thesis is causal theory, that is, the study of the possible cause-and-effect relationships between spacetime events. In this section, we state the basic notions of causal theory for smooth Lorentzian manifolds, adding some comments about how these notions inspire or relate to the more general frameworks considered in the upcoming chapters. Note that for all purposes, C^2 metric tensors are as good as smooth ones (in fact, all the way down to local Lipschitz continuity, many properties are preserved).

Definition 1.1. A *smooth spacetime* is a triple (M, g, X) consisting of a smooth manifold M, a smooth Lorentzian metric tensor $g\colon TM \otimes TM \to \mathbb{R}$ of signature $(-, +, \dots, +)$, and a continuous vector field X such that $g(X, X) < 0$, called the time-orientation.

The time orientation allows us to distinguish between the future and past lightcone at every point (see Figure 1.1). By abuse of notation, we will often suppress it. Points on a spacetime have the physical interpretation of *events*, and certain curves correspond to the trajectories of observers or signals. The fundamental principle that nothing can travel faster than the speed of light is enforced through the following conditions on the tangent vector to the trajectory.

Definition 1.2. Let $I \subseteq \mathbb{R}$ be an interval. A curve $\gamma\colon I \to M$ with $\dot\gamma \neq 0$ is called

(i) *timelike* if $g(\dot\gamma, \dot\gamma) < 0$,

(ii) *null* if $g(\dot\gamma, \dot\gamma) = 0$,

(iii) *causal* if $g(\dot\gamma, \dot\gamma) \leq 0$.

Any such curve is *future-directed* if $g(\dot\gamma, X) < 0$ and *past-directed* if $g(\dot\gamma, X) > 0$.

Null curves correspond to trajectories precisely at the speed of light, timelike curves to strictly slower ones (such as those of massive particles). We are being deliberately vague about the regularity class of curves to consider; This issue will be discussed in depth in Chapter 2. The notions of causal and timelike curves induce the *causal structure*, a pair of relations defined as follows: We write $p \leq q$ if $p = q$ or there exists a future directed causal curve from p to q (in that case $p < q$), and $p \ll q$ if there exists a future-directed timelike curve from p to q.

$$\begin{aligned} J^+(p) &:= \{q \in M \mid p \leq q\}, & I^+(p) &:= \{q \in M \mid p \ll q\}, \\ J^-(p) &:= \{q \in M \mid q \leq p\}, & I^-(p) &:= \{q \in M \mid q \ll p\}. \end{aligned}$$

It was shown by Hawking [Haw14, Lem. 19] (see also [HKM76, Thm. 5]) that the causal structure determines the metric tensor g up to a conformal transformation, that is, up to multiplication by a positive function.

The three following tools lie at the core of almost every proof in causal theory and are proven using local arguments in normal neighbourhoods.

- **Push-up property.** $p \leq q \ll r$ or $p \ll q \leq r \implies p \ll r$.
- **Limit curve theorem.** If a sequence of causal curves converges locally uniformly, then the limit curve is also causal[1].
- **Openness of the chronological relation.** $\ll$ is an open subset of $M \times M$.

One might even postulate the above as axioms. Indeed, this was first done in the 1967 article by Kronheimer and Penrose [KP67], where a *causal space* is (essentially) defined as a set X equipped with two relations $\ll, \leq$ satisfying the push-up property. Kunzinger and Sämann [KS18] extended this for their Lorentzian length spaces by adding a topology and suitable axioms which imply the limit curve theorem and openness of $\ll$. More generally, the idea of studying abstract "causal-like" relations, in addition to or instead of the usual $\ll, \leq$, permeates the whole thesis.

The causal relation on an arbitrary spacetime can be quite ill-behaved, both from the physical and the mathematical perspective. This led to the development of various *causality conditions*, which are additional restrictions on spacetimes in order to have a nicer causal structure. The different causality conditions are often ordered from strongest to weakest, forming the *causal ladder* on Figure 1.4 (see also [BEE96; Car71; Min19a]). In the remainder of this section, we review those steps in the causal ladder that will be most relevant in the upcoming chapters.

[1] The limit curve theorem admits many variants (see [Min08a]), all of which involve the above statement combined with an application of the Ascoli–Arzelà theorem. The latter provides sufficient conditions for a sequence of causal curves to have a convergent subsequence.

Globally hyperbolic
$\Downarrow$
Causally simple
$\Downarrow$
Causally continuous
$\Downarrow$
Stably causal (K-causal)
$\Downarrow$
Strongly causal
$\Downarrow$
Distinguishing
$\Downarrow$
Non-totally imprisoning
$\Downarrow$
Causal
$\Downarrow$
Chronological

Figure 1.4: The causal ladder.

1.3.1 Causality

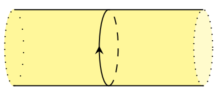


Figure 1.5: A closed timelike curve.

A spacetime is said to be *causal* if the relation $\leq$ is antisymmetric, or, equivalently, if there are no closed causal curves. Physically, causality represents the impossibility of time travel. Without imposing causality, it is not possible to establish cause-and-effect relationships, which are vital for science. Moreover, the possibility of time travel leads to unacceptable consequences, such as the grandparents paradox, i.e., that someone could travel to the past and kill their own grandparents, but then this person should have never been born in the first place. Causality is therefore assumed in almost all contexts. An example of a non-causal spacetime is the cylinder $S^1 \times \mathbb{R}$ with metric $-d\theta^2 + dx^2$ and time orientation ∂_θ, where for instance the curve $s \mapsto (s, 0)$ is a closed causal (even timelike) curve, depicted in Figure 1.5.

1.3.2 Stable causality and K-causality

A spacetime (M, g) is said to be *stably causal* if for every Lorentzian metric $\tilde{g}$ that is C^0-close to g, the spacetime $(M, \tilde{g})$ is causal. Thus (M, g) itself is also causal. In physics, one never expects to be able to have a completely accurate description of every real-life situation. Rather, one studies idealized cases that approximate real life. This fact has to be accounted for by defining concepts in a way that is stable under perturbations; otherwise one runs into so-called fine-tuning problems.

Stable causality was introduced by Hawking [Haw68], who showed that it is equivalent to the existence of a time function, defined as follows (see also later, more rigorous proofs [CGM16; Min09; Min10]).

Definition 1.3. A continuous function $\tau\colon M \to \mathbb{R}$ is called a *time function* if one of the following equivalent conditions hold:

(i) τ is strictly increasing along every future-directed causal curve.

(ii) $p < q \implies \tau(p) < \tau(q)$.

A smooth function $\tau\colon M \to \mathbb{R}$ is called a *temporal function* if

(iii) $d\tau$ is past-directed timelike.

Temporal functions are a special case of time function, condition (iii) being the infinitesimal version of (i). It was shown by Bernal and Sánchez that every smooth spacetime admitting a time function also admits a temporal function [BS05].

Stable causality is also known to be equivalent to another condition called *K-causality*. The latter is defined by antisymmetry of the relation $K^+ \subseteq M \times M$, which in turn is defined as the smallest closed and transitive relation containing $\ll$ (equivalently, containing $\leq$). The K^+ relation was by introduced in Sorkin and Woolgar [SW96], and later extensively studied by Minguzzi [Min09; Min10]. The K^+ relation also appears in this thesis, namely in Chapters 2, 4 and 6. The fact that K^+ is defined intrinsically using only on the chronological relation $\ll$ and the topology on $M \times M$ makes it more adaptable to a low-regularity context.

1.3.3 Causal continuity

Causal continuity is the condition that the set-valued maps $p \mapsto I^{\pm}(p)$ be continuous. This can be made precise in various ways, for instance by choosing an admissible (see Section 4.4.1) measure μ on M, and requiring the real-valued maps $t^{\pm}\colon p \mapsto \pm\mu\left(I^{\pm}(p)\right)$ to be continuous. By transitivity of $\ll$, the maps $t^{\pm}$ are always increasing along future-directed causal curves, and hence $t^{\pm}$ are time functions whenever they are continuous. Causal continuity thus implies stable causality.

Hawking and Sachs [HS74] introduced causal continuity, arguing that it is necessary, from a philosophical point of view, in order to do physics. For if an observer could suddenly observe a large region of spacetime that was hidden just an instant earlier, this would ruin the possibility of making meaningful physical predictions. Moreover, Sorkin [Sor89] conjectures that causal continuity is necessary in order to have well-behaved quantum field theories on the spacetime (see Conjecture 5.1).

1.3.4 Global hyperbolicity

Global hyperbolicity is the strongest condition on the causal ladder. It is defined by requiring that (M, g) is causal[2] and has compact *causal diamonds* $J(p, q) :=$

[2] This is a modern, streamlined definition due to Bernal and Sánchez [BS07]. Recently, Hounnonkpe and Minguzzi [HM19] showed that the assumption of (M, g) being causal is redundant when M is non-compact and of dimension greater than 2.

$J^+(p) \cap J^-(q)$ for all $p, q \in M$. Despite being the most restrictive causality condition, there are reasons to believe that physically realistic spacetimes should be globally hyperbolic.

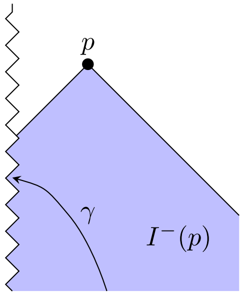


Figure 1.6: A locally naked singularity, represented by the zig-zag line.

Penrose argued that global hyperbolicity is equivalent to the absence of locally naked singularities [Pen79, Sec. 12.3.2]. Singularities correspond to the breakdown of spacetime, for instance at the Big Bang or inside a black hole. They are rather tricky to define precisely (see for instance [Ger68] for a discussion). A singularity is said to be locally naked if it is possible for a nearby observer to send a signal and watch it fall into the singularity in finite proper time. In this context, it means that there exists an *inextendible* causal curve γ, corresponding to the signal, lying entirely in the past $I^-(p)$ of some point p, corresponding to the observer at a given instant (see Figure 1.6). According to Penrose, such a situation is unphysical and cannot happen, because it would ruin predictability. This statement is known as the strong cosmic censorship conjecture and is equivalent to saying that physically realistic spacetimes are globally hyperbolic. Indeed, it is easy to see that the causal diamond $J^+(q, p)$, for any q on γ, is non-compact (one can find a sequence of points along γ that approach the singularity, hence it has no convergent subsequences).

Global hyperbolicity has strong geometrical implications. We highlight the following theorem, proven successively by Geroch [Ger70] (smooth g, continuous Σ and τ), Bernal and Sánchez [BS05] (everything smooth, τ temporal), and Sämann [Säm16] (continuous g, smooth or continuous Σ and τ). There also exist versions for cone structures [BS18; FS12; Min19b], and we prove a version for Lorentzian length spaces in Chapter 4.

Theorem 1.4. *Let (M, g) be a spacetime. The following are equivalent:*

(i) (M, g) *is globally hyperbolic.*

(ii) (M, g) *contains a Cauchy surface Σ. That is, a subset $\Sigma \subset M$ such that every inextendible causal curve intersects Σ exactly once.*

(iii) (M, g) *admits a Cauchy time function τ. That is, a time function $\tau \colon M \to \mathbb{R}$ such that $\tau^{-1}(C) \subset M$ is a Cauchy surface for every $C \in \mathbb{R}$.*

Cauchy surfaces are, as the name suggests, embedded codimension-1 hypersurfaces. Using the flow of the time orientation vector field, one can see that two Cauchy surfaces in a given spacetime are always homeomorphic (diffeomorphic if both are smooth), and M decomposes into $\mathbb{R} \times \Sigma$. When τ is temporal, also the metric splits orthogonally into

$$g = -\beta d\tau^2 + \bar{g}, \tag{1.9}$$

where β is a function and $\bar{g}$ restricts to a Riemannian metric on each τ-level set.

Cauchy surfaces are of special importance in connection to the Einstein equations (1.4), since Σ with its induced metric, extrinsic curvature tensor and energy-momentum tensor satisfying the constraint equations provides an admissible initial data set for the Einstein equations. More generally, global hyperbolicity is intimately linked to the well-posedness of the Initial Value Problem (IVP) of hyperbolic Partial Differential Equations (PDEs) through the following two results, due to Leray [Ler53] (who was, in fact, the first to define global hyperbolicity) and Choquet-Bruhat and Geroch [CBG69], respectively:

- The IVP for the wave equation on a fixed background spacetime (M, g) is well-posed if and only if (M, g) is globally hyperbolic.
- Given an initial data set for the Einstein equations, there exists a unique Maximal Globally Hyperbolic Development (MGHD) (M, g) of said initial data.

Here, the term "maximal" means inextendible within the class of globally hyperbolic solutions; the MGHD might still admit a non-globally hyperbolic extension. Christodoulou [Chr99] turned the strong cosmic censorship conjecture of Penrose (which is of physical nature, see above), into the mathematical conjecture that the MGHD of generic, admissible initial data is inextendible. Thus the statement can be read as "the globally hyperbolic development is all there is, thus spacetime is globally hyperbolic". Here we mean, of course, inextendible in the class of all spacetimes, not only globally hyperbolic ones. Moreover, it is customary to allow extensions with metrics g of low-regularity, in particular, continuous metrics with locally square integrable Christoffel symbols. The assumption on the Christoffel symbols is required both so that g can be interpreted as a weak solution to the Einstein equations, and because without it, the conjecture is known to be false [DL17]. Regardless, said regularity class still allows for much wilder behaviour than, say, Lipschitz continuous metrics, such as the bubbling phenomena investigated in Chapters 2 and 3.

1.4 Outline of results and discussion

In this section, we give an outline of the results in each chapter. We also contextualize them in the current literature, including references that appeared after the publication of the preprint or article the chapter is based on.

Chapters 2 and 3: Continuous Lorentzian metrics

In Chapters 2 and 3, we study the causality theory of spacetimes with continuous metrics. Recall from Section 1.2.1 that these appear in general relativity as weak solutions to the Einstein equations. Chruściel and Grant [CG12] found that the push-up property (see page 9) fails for continuous metrics, leading to most of known causal theory to break down. In particular, the sets $J^+(p) \setminus I^+(p)$ can contain regions of non-empty interior, called *causal bubbles*, which do not appear when the metric is smooth (see Figure 1.7). Later, Grant et al. [Gra+20] discovered that the openness of $\ll$ holds or fails depending on the regularity class where one defines timelike curves, something which does not matter for smooth metrics [Chr20, Cor. 2.4.11].

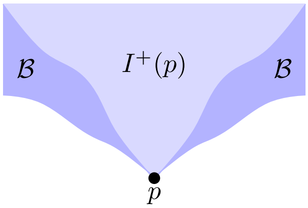


Figure 1.7: A causal bubble $\mathcal{B}$. Here $J^+(p) = I^+(p) \cup \mathcal{B}$.

Given these difficulties, it appears natural to try and replace the definition of $\leq$ via causal curves (or that of $\ll$ via timelike ones) by an axiomatic notion, which ideally should recover the usual relations in the smooth case, and be more robust in low-regularity. Indeed, already in 1996 and prior to the discovery of causal bubbles, Sorkin and Woolgar suggested such an approach via their K^+-relation [SW96]. However, the combination of $\ll$ and K^+ does not satisfy the push-up property either, as K^+ is in general very large.

Recall that K^+ is the smallest closed and transitive relation containing $\ll$. While closedness is a nice property, $\leq$ is usually not closed, even on smooth spacetimes. In fact, causal spacetimes with closed $\leq$ are called *causally simple*, which is the second highest level on the causal ladder (see Figure 1.4). Something that *does* always hold is weak local causal closedness (Definition 4.10), which is in fact crucial in order to establish the limit curve theorem. On the other hand, Minguzzi [Min08b] defines a largest relation d^+ such that the push-up property (w.r.t. $\ll$) holds. Recall also that Kronheimer and Penrose [KP67] had established a series of axioms for a pair of two relations to form a "causal space", notably including the push-up property. In this spirit, we ask ourselves whether there exists a pair of relations on spacetimes with continuous metrics such that:

- One of them is the usual notion with timelike or causal curves, since we don't want just any pair of relations on M, completely unrelated to g.

- They satisfy openness of the smaller relation, the push-up property, and a limit curve theorem (page 9).

The answer is **negative**: We construct examples where the maximal relation satisfying one property is strictly smaller than the minimal relation satisfying another one, and hence no relation with all desired properties can exist.

In spite of the possible causal pathologies described above, Theorem 1.4 still holds for continuous metrics, leading to the question of whether global hyperbolicity prevents such pathologies [Säm16]. Again, the answer is **negative**, as is shown by our example in Chapter 3, where additionally the metric splits into (1.9), something which for continuous metrics does not necessarily follow from global hyperbolicity. We also argue, with the same example, that lower Ricci curvature bounds do not prevent causal pathologies, either. Given that the metric is only continuous, synthetic notions of curvature bounds have to be considered (see Section 1.2.2).

Chapter 4: Time functions on Lorentzian length spaces

In Chapter 4, we turn to the framework of Lorentzian length spaces in the sense of Kunzinger and Sämann [KS18]. These include some spacetimes with continuous metrics (in fact, precisely those without causal pathologies), in addition to more general, non-manifold spaces equipped with a chronological and causal order and a notion of Lorentzian distance. In this setting, one can define the causality conditions on the causal ladder, and wonder whether the implications between them still hold. Indeed, Kunzinger and Sämann [KS18] considered some steps of the ladder, and later Aké Hau, Cabrera Pacheco and Solis [ACS20] brought it to the full extent shown in Figure 1.4.

Given that various causality conditions on the causal ladder can be characterized by the existence of (certain types of) time functions, the next natural question is whether also these characterizations carry over to Lorentzian length spaces. Indeed, we prove in this thesis that the existence of a time function is equivalent to K-causality. Notice that K-causality is well-defined for Lorentzian length spaces, while it is not straightforward to translate stable causality to this setting (because we do not have a metric tensor that we can perturb).

Theorem 4.1. *Suppose X is a second countable, locally compact Lorentzian length space. Then X is K-causal if and only if X admits a time function.*

The existence of a time function had already been established by Minguzzi for very general topological ordered spaces [Min10]. The converse (that a time function implies K-causality), also shown by Minguzzi in the manifold setting [Min10], requires the limit curve theorems. Hence our proof makes critical use of the structure of a Lorentzian (pre-)length space, as does the proof of the following non-smooth analogue to Theorem 1.4.

Theorem 4.3. *Let X be a second countable Lorentzian length space with a proper metric structure. Then the following are equivalent:*

(i) X is globally hyperbolic,

(ii) X is non-totally imprisoning and the set of causal curves between any two points is compact,

(iii) X admits a Cauchy set,

(iv) X admits a Cauchy time function.

Here Cauchy set is defined in the same way as a Cauchy surface is in the manifold case, but one cannot say that it is a surface when the ambient space is not a manifold. Then again, McCann and Sämann [MS22a] have recently introduced an analogue of the Hausdorff dimension for Lorentzian length spaces. It remains open if or when Cauchy sets have codimension-1 in their sense. Another important difference to the smooth setting is that a given globally hyperbolic Lorentzian length space can admit non-homeomorphic Cauchy sets (as in Figure 1.3).

Note that in some works, especially in connection with lower Ricci bounds [CM20], the (in principle) stronger causality condition $\mathcal{K}$-global hyperbolicity is used instead of global hyperbolicity. Concretely, $\mathcal{K}$-global hyperbolicity strengthens the requirement of compactness of causal diamonds $J(p,q)$ to causal emeralds $J(K,S)$ for all $K,S \subset X$ compact. In fact, Minguzzi [Min23] has argued that $\mathcal{K}$-global hyperbolicity is more natural in the length space setting, but has also shown that under the assumptions of Theorem 4.3, both conditions are equivalent [Min23, Cor. 3.8]. In that same article, it is also shown that global hyperbolicity for Lorentzian length spaces can be equivalently defined using causality (as in Section 1.3.4), rather than non-total imprisonment (as in Definition 4.56).

Chapter 5: Topology change with Morse functions

The observation that Cauchy sets in a Lorentzian length space need not be homeomorphic to each other leads us to Chapter 5, which can also be read independently of Chapter 4. Chapter 5 deals precisely with the phenomenon of topology change, in the concrete setting of *Morse spacetimes*, motivated by quantum gravity (see Section 1.2.3).

Morse spacetimes, introduced by Yodzis [Yod72; Yod73], are n-manifolds M equipped with a tensor of the form

$$g = \|df\|_h^2\, h - \zeta df \otimes df,$$

where f is a Morse function, h is a Riemannian metric and $\zeta > 1$ a constant. At the regular points (where $df \neq 0$), g is a Lorentzian metric, while at the critical points (where $df = 0$), $g = 0$. The topology of the level sets of f changes precisely at the critical values, hence their usefulness in this context.

Borde and Sorkin conjectured that Morse spacetimes are causally continuous precisely when f has no critical points of index 1 or $n-1$ (Conjecture 5.2). Their conjecture is intended to provide an admissibility criterion for topology changes. We prove the following special case.

Theorem 5.3 (simplified phrasing)**.** *Let (M,g) be a Morse spacetime of dimension n with a single critical point p_c of index $\lambda \neq 0,1,n-1,n$, which is contained in a coordinate neighborhood where*

$$f = \frac{1}{2}\sum_i a_i (x^i)^2, \qquad\qquad h = \sum_i (dx^i)^2,$$

for some real constants $a_i \neq 0$ satisfying

$$\frac{1}{\zeta} < \left|\frac{a_i}{a_j}\right| < \zeta \quad \text{and} \quad \frac{5}{8} \leq \left|\frac{a_i}{a_j}\right| \leq \frac{8}{5} \quad \text{for all } i,j.$$

Then (M,g) is causally continuous, as predicted by the Borde–Sorkin conjecture.

In order to contextualize our result, note that a Morse spacetime with various critical points can be decomposed into pieces with one critical point each [DGS00b]. The cases of index $0,1,n-1$ and n of the conjecture were already known to be true, as was Theorem 5.3 for $|a_i| = 1$ [Bor+99; DGS00b]. Moreover, we show in Proposition 5.15 that the most general case only differs from the case in Theorem 5.3 by two things: dropping the bounds on $\frac{a_i}{a_j}$ and adding a perturbation term to h (vanishing at p_c).

Finally, we argue (heuristically) via Example 5.17 that the bound $\frac{1}{\zeta} < \frac{a_i}{a_j} < \zeta$ is needed, and propose a revised version of the conjecture that incorporates this assumption. How to deal with a perturbation term of h is still completely open at this point. Indeed, we give some simple examples in Section 5.6 to illustrate that causal continuity is, in general, not stable under perturbations. While this does not necessarily mean that the (modified) Borde–Sorkin conjecture is false, it does imply that extending the proof of Theorem 5.3 is non-trivial.

Chapter 6: Global hyperbolicity and the null distance

Finally, in Chapter 6 we continue studying causality and time functions, but on the traditional smooth spacetimes. However, we view them as metric spaces via the null distance of Sormani and Vega [SV16], which assigns to a spacetime (M, g) with a choice of time function τ, a metric space structure $(M, \hat{d}_\tau)$. The null distance $\hat{d}_\tau(p, q)$ is given as the infimum of the *null lengths*

$$\hat{L}_\tau(\beta) := \sum_{i=1}^{k} \left|\tau(\beta(s_i)) - \tau(\beta(s_{i-1}))\right|$$

over all *piecewise causal* curves β from p to q (see Figure 1.8).

From this perspective, the causal structure is still visible, in the following sense.

Theorem 6.9. *Let (M, g) be a globally hyperbolic spacetime and τ a locally anti-Lipschitz time function such that all nonempty level sets are Cauchy. Then the null distance* encodes causality*, that is, for any $p, q \in M$,*

$$q \in J^+(p) \Longleftrightarrow \hat{d}_\tau(p, q) = \tau(q) - \tau(p). \tag{1.10}$$

The use of the condition (1.10) to encode causality was first proposed by Sormani and Vega [SV16]. The proof of Theorem 6.9 combines a local causality encodation result due to Sakovich and Sormani [SS23] with a novel local-to-global argument. We also give a self-contained proof of the local part for temporal functions. Regarding the assumptions, locally anti-Lipschitz τ is a standard assumption that guarantees definiteness of $\hat{d}_\tau$. Cauchyness of the level sets, on the other hand, is the crucial ingredient that allows passing from local to global. It can be slightly relaxed to future/past Cauchy, which allows us to prove Corollary 6.10 below. Global hyperbolicity of the underlying spacetime alone, however, is not enough to obtain Theorem 6.9, as shown by Example 6.34.

Corollary 6.10. *Let (M, g) be a spacetime that admits a regular cosmological time function τ. Then $\hat{d}_\tau$ encodes causality globally.*

The cosmological time function was introduced by Wald and Yip [WY81], and independently by Andersson, Galloway and Howard [AGH98]. It is defined as

$\tau(p) := \sup_x d_g(x,p)$ (where d_g is given by (1.7)), and has good properties only on certain spacetimes (where it is called regular). Its physical interpretation is that it measures proper time elapsed since the Big Bang. Sormani and Vega proposed to use the cosmological time function in order to resolve the ambiguity of having many time functions with respect to which one can define the null distance [SV16] (see also further discussion below).

If we are to use the null distance to study convergence of spacetimes, we need to be able to interpret the limit space, which will be a metric space equipped with a continuous "time" function, as a spacetime-like object as well. Results like Theorem 6.25 allow us to do this. Another perspective on this topic is given by the following theorem, which relates a causal property of the spacetime to a metric property of the null distance.

Theorem 6.4. *A spacetime (M, g) is globally hyperbolic if and only if there exists a time function τ such that $(M, \hat{d}_\tau)$ is a complete metric space.*

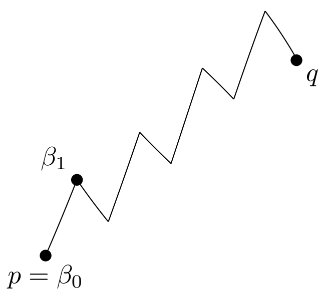


Figure 1.8: A piecewise causal curve β from p to q, with alternating future- and past-directed pieces.

We have thus found yet another characterization for global hyperbolicity (cf. Theorem 1.4). Note, however, that not every Cauchy time function induces a complete null distance (Example 6.39). A sufficient condition for completeness is for the time function to be *completely uniform*, a notion introduced by Bernard and Suhr [BS18; BS20]. The optimal condition for completeness still remains to be found.

More generally, the question arises of how much the null distance depends on the choice of time function; Corollary 6.11 tells us that any two null distances are always locally bi-Lipschitz equivalent, for *weak* temporal functions. The latter are a new regularity class in between temporal functions and time functions, which we define by requiring local bi-Lipschitz type bounds, but no differentiability.

Finally, note that the null distance has also been studied on Lorentzian length spaces by Kunzinger and Steinbauer [KS22]. This establishes a link between Chapters 4 and 6, since the existence of a time function is of course a prerequisite for defining the null distance. Furthermore, our results show that the null distance is especially well behaved for weak temporal functions. Unlike (smooth) temporal functions, their weak counterparts can be defined on Lorentzian length spaces, although a general existence result like Theorem 4.34 is still lacking. Also the validity of the theorems in Chapter 6 on Lorentzian length spaces is still open, as our proof techniques do rely on the differential structure.

Chapter 2

Causality theory of spacetimes with continuous Lorentzian metrics revisited

This chapter is based on the article [GH21] of the same title, published in *Classical and Quantum Gravity*.[1]

2.1 Introduction

The study of spacetimes with metrics of low regularity is a topic of rising importance in Lorentzian geometry. The main motivation stems from the strong cosmic censorship conjecture [DL17; Sbi18] and the occurrence of weak solutions to Einstein's equations coupled to certain matter models [BL14; GT87]. It has hence become an important research question to establish which properties of the usual, smooth spacetimes are more "robust" or "fundamental", in the sense that they continue to hold in lower regularity, and which, on the other hand, depend sensitively on the smoothness assumption. In trying to answer this question, the need arises to axiomatize the notion of spacetime, and in particular, to treat the causal structure in an order-theoretic way. This, in turn, connects well with ideas in quantum gravity, such as causal set theory [Bom+87]. To make matters more concrete, in the present paper we shall study spacetimes (M, g) where g is a continuous Lorentzian metric. However, since our approach is indeed of the order-theoretic type, it can easily be adapted to other settings.

Let us start by recalling the case of a classical spacetime (M, g) where g is smooth. The chronological and causal relations I^+ and J^+ can then be defined using the notions of timelike and non-spacelike curve respectively. The three following facts are well-known:

[1] I am very grateful to Annegret Burtscher for discussions and detailed comments on the draft. I also wish to thank Klaas Landsman, Ettore Minguzzi and two anonymous referees for further comments.

(i) The push-up lemma: if $p \in I^+(q)$ and $q \in J^+(r)$ then $p \in I^+(r)$.

(ii) The limit curve theorem: the uniform limit of a converging sequence of causal curves is a causal curve.

(iii) Openness of chronological pasts and futures: the sets $I^\pm(p)$ are open, for any $p \in M$.

With these three results at hand, one can develop a large portion of causality theory, such as the causal ladder and the characterization of time functions, without ever again mentioning the metric g or the manifold structure on M explicitly. This is confirmed by the "Lorentzian length spaces" approach of Kunzinger and Sämann [KS18] and follow-up work [ACS20; BGH21].

In order to do causality theory on a spacetime with a $\mathcal{C}^0$-metric, the first question is how the causal structure should even be defined. The obvious answer is to define I^+ and J^+ through timelike and non-spacelike curves, just as in the smooth case. However, there are two potential problems:

A. Points where the metric is not $\mathcal{C}^2$ do not admit normal neighborhoods.

B. The regularity class where we define timelike curves becomes important.

Chruściel and Grant [CG12] showed that because of A, the push-up lemma fails, while the limit curve theorems are unaffected (points 1 and 2 above). As a consequence, spacetimes with continuous metric exhibit so called "causal bubbles", open regions contained in J^+ but not in I^+. Regarding B, when the metric is at least $\mathcal{C}^2$, it was shown by Chruściel [Chr20] that one obtains the same chronological relation I^+ regardless of whether timelike curves are required to be Lipschitz or piecewise-differentiable. In the case of continuous metrics, however, it was shown by Grant et al. [Gra+20] that this choice makes an important difference. In particular, they showed that the chronological futures and pasts are open when using piecewise-differentiable curves, but not when using Lipschitz curves (point 3 above).

A radically different, and in fact earlier, approach is that of Sorkin and Woolgar [SW96]. They propose to keep the definition of I^+ by piecewise-differentiable timelike curves, and then introduce another relation K^+ as the smallest transitive, topologically closed relation containing I^+. The relation K^+ can then be used to replace J^+. Even for smooth metrics, the two relations K^+ and J^+ do not coincide. Nonetheless, it is possible to define the usual causal curves (and hence J^+) in terms of K^+, without referring to the metric directly. The K^+-relation has since found a variety of applications, most notably Minguzzi's works on stable causality [Min09] and time functions [Min10]. However, there is no push-up lemma for the K^+-relation.

Following a similar philosophy, one can define a relation d^+ as the largest relation such that the push-up lemma holds true. On spacetimes with $\mathcal{C}^2$-metrics, it was shown by Minguzzi [Min08b] that such a maximal relation exists, and various characterizations of it are given, using the name D^+. We will discuss which of these characterizations are valid for continuous metrics (not all of them, hence the change in notation). The D^+-relation was used by Minguzzi in order to characterize the causality condition known as weak distinction: it is shown that a spacetime is weakly distinguishing if and only if (I^+, D^+) is a causal structure in the sense of Kronheimer and Penrose [KP67]. This result also holds in the $\mathcal{C}^0$-setting for d^+, as we will see in detail. Further, we propose a definition of causal curve based on d^+, similar to Sorkin and Woolgar's K^+-causal curves. When the metric is smooth, causal curves defined through d^+ coincide with those defined by the metric g, but when the metric is merely continuous, they do not. As a consequence, we show that while our new causal relation satisfies the push-up lemma, the limit curve theorems cease to hold. We then argue that, essentially because we chose d^+ to be maximal, there in fact do not exist any causal relations that can satisfy both properties at the same time. That is, at least, if we want to keep the usual definition of (piecewise continuously differentiable) timelike curve, the only one that guarantees open futures. We also explore the possibility of alternative chronological relations, but we conclude that one runs into the same problems.

We conclude that spacetimes with continuous metrics are unavoidably pathological. This strengthens the view, already present in the literature, that one should focus on a special class of continuous metrics, the so-called causally plain ones. These are the ones where the usual push-up lemma holds, and include, for example, the class of locally Lipschitz metrics, but not the class of Hölder continuous metrics [CG12, Thm. 1.20]. In this sense, our paper can be seen to support a $\mathcal{C}^{0,1}_{\text{loc}}$-formulation of strong cosmic censorship, such as in Sbierski's recent work [Sbi22]. Moreover, the methods and results of this paper are also highly relevant to the study of axiomatic causality relations in other settings; for example, in the study of spacetimes with degenerate metrics [DGS00a], or with metrics that are not even continuous [GT87].

Outline. In Section 2.2 we provide more background, define the new causal relation d^+ and study its properties. In Section 2.3 we define causal curves in terms of d^+, and show that these are just the usual causal curves when the metric is smooth. In Section 2.4, we discuss other possible choices of causal structure. In Section 2.5, we summarize and discuss our results.

2.2 The d^+-relation

2.2.1 Basic notions in causality theory

Let M denote a Hausdorff, paracompact $\mathcal{C}^1$-manifold, and g a $\mathcal{C}^0$-Lorentzian metric. Assume that (M, g) admits a $\mathcal{C}^0$-vector field X such that $g(X, X) < 0$, called a time orientation. The pair (M, g) together with a choice of time orientation is called a $\mathcal{C}^0$*-spacetime*, following the nomenclature of Ling [Lin20]. Whenever we say that g is $\mathcal{C}^2$ (or smooth), or that (M, g) is a $\mathcal{C}^2$ (or smooth) spacetime, we mean that M admits a $\mathcal{C}^3$ (resp. smooth) subatlas such that g is $\mathcal{C}^2$ (resp. smooth) in this subatlas. By *relation* on M we will mean a subset of $M \times M$. The closure of a relation R, denoted by $\overline{R}$, is the topological closure in the product topology on $M \times M$. Likewise, we say that R is open if it is an open subset of $M \times M$.

There exist two different definitions for the notion of *timelike curve*, and thus two different notions of *chronological relation*. For $\mathcal{C}^2$-metrics the two notions are equivalent [Chr20, Cor. 2.4.11], but not for $\mathcal{C}^0$-metrics [Gra+20]. One is based on the class $\mathcal{L}$ of locally Lipschitz curves, and the other one on the class $\mathcal{C}^1_{\mathrm{pw}}$ of piecewise continuously differentiable curves:

$$\begin{aligned} I^+ := \{(p,q) \in M \times M \mid \, & \text{there exists an } \mathcal{L}\text{-curve } \gamma : [a,b] \to M \\ & \text{such that } \gamma(a) = p,\ \gamma(b) = q, \\ & \text{and } g(\dot\gamma,\dot\gamma) < 0,\ g(\dot\gamma, X) < 0 \text{ almost everywhere}\}, \\ \check{I}^+ := \{(p,q) \in M \times M \mid \, & \text{there exists an } \mathcal{C}^1_{\mathrm{pw}}\text{-curve } \gamma : [a,b] \to M \\ & \text{such that } \gamma(a) = p,\ \gamma(b) = q, \\ & \text{and } g(\dot\gamma,\dot\gamma) < 0,\ g(\dot\gamma, X) < 0 \text{ everywhere}\}. \end{aligned}$$

Recall that an $\mathcal{L}$-curve is differentiable almost everywhere by Rademacher's theorem. In the second case, when γ is $\mathcal{C}^1_{\mathrm{pw}}$, the condition $g(\dot\gamma,\dot\gamma) < 0$ is meant to hold for both one-sided derivatives (which may differ from each other at break points). It was established by Grant et al. that $\check{I}^+$ is open, but I^+ is not necessarily open [Gra+20]. The standard *g-causal relation* is defined as

$$\begin{aligned} J^+ := \{(p,q) \in M \times M \mid \, & \text{there exists an } \mathcal{L}\text{-curve } \gamma : [a,b] \to M \\ & \text{such that } \gamma(a) = p,\ \gamma(b) = q, \\ & \text{and } g(\dot\gamma,\dot\gamma) \leq 0,\ g(\dot\gamma, X) < 0 \text{ almost everywhere}\}, \end{aligned}$$

where we say that γ is a *g-causal curve*. We can analogously define the past relations I^-, $\check{I}^-$ and J^- by requiring $g(\dot\gamma, X) > 0$, but since they are simply given by reversing the factors, there is no need to treat them separately. Therefore we also shall not specify every time that our timelike and causal curves are always future-directed. We do, however, sometimes use the notations $q \in J^+(p)$, and $p \in J^-(q)$, both meaning the same, namely $(p,q) \in J^+$. As mentioned in the introduction, Sorkin and Woolgar suggested an order-theoretic alternative K^+ to J^+.

Definition 2.1 ([SW96, Def. 8])**.** The relation K^+ is the smallest closed, transitive relation containing $\check{I}^+$.

We finish this subsection with a short digression about limit curve theorems. In the literature, there exist multiple statements with this name; a detailed review can be found in [Min08a] for smooth spacetimes. In [CG12, Thm. 1.6] it is shown how the limit curve theorems for the smooth case also carry over to continuous spacetimes (see also [Säm16, Thm. 1.5]). Roughly speaking, a limit curve theorem is the combination of the following two statements:

(i) Under certain assumptions, a sequence of causal curves has a convergent (in some appropriate sense) subsequence.

(ii) The limit of said subsequence is itself a causal curve.

Here causal usually means g-causal, but we will also discuss alternative notions of causal curve. Regarding part 1, there exist many versions tailored to different applications. A common variation is to require the curves to be Lipschitz (as we did in our definition of J^+) and add some compactness assumptions in order to apply the Arzelà–Ascoli theorem. Part 2 is where the causal structure becomes important; we discuss it in the context of our new d^+-relation in Remark 2.12.

2.2.2 Definition of d^+

We first introduce some nomenclature, the underlying concepts being fairly standard. By (M, g) we continue to denote a $\mathcal{C}^0$-spacetime, although some of the ideas make sense even if M is just a set. The following definition gives a compatibility condition between the chronological and causal relations, in this case denoted abstractly by R and S respectively.

Definition 2.2. Let $R, S \subseteq M \times M$ be two relations. We say that S satisfies *push-up relative to* R (or that R *is an* S*-ideal*, cf. [Min08b; Min19a]) if the following two properties hold:

(i) $(x, y) \in S, (y, z) \in R \implies (x, z) \in R$,

(ii) $(w, x) \in R, (x, y) \in S \implies (w, y) \in R$.

Let $R, S, S' \subseteq M \times M$ be relations. Then it is easy to see that:

(a) If R is transitive, then R satisfies push-up relative to itself.

(b) If S satisfies push-up relative to R, and $S' \subseteq S$, then also S' satisfies push-up relative to R.

(c) If S and S' each satisfy push-up relative to R, then so does $S \cup S'$.

If (M, g) is $\mathcal{C}^2$, then J^+ satisfies push-up relative to I^+ (and equivalently, $\check{I}^+$). This fact is known as the push-up lemma [Chr20, Lem. 2.4.14]. Those $\mathcal{C}^0$-spacetimes where J^+ satisfies push-up relative $\check{I}^+$ are called *causally plain*. The term was coined by Chruściel and Grant [CG12, Def. 1.16], but beware that they used I^+ in place of $\check{I}^+$. In any case, not all $\mathcal{C}^0$-spacetimes are causally plain [CG12, Ex. 1.11]. The failure of the push-up lemma on arbitrary $\mathcal{C}^0$-spacetimes motivates our next, central definition.

Definition 2.3. The d^+*-relation* is the largest relation that satisfies push-up relative to $\check{I}^+$.

Proposition 2.4. *There exists a unique relation d^+ satisfying Definition 2.3. Moreover, d^+ is transitive and reflexive.*

Proof. We construct d^+ by setting $(x, y) \in d^+$ if:

(i) $(y, z) \in \check{I}^+ \implies (x, z) \in \check{I}^+$ (in other words, $\check{I}^+(y) \subseteq \check{I}^+(x)$),

(ii) $(w, x) \in \check{I}^+ \implies (w, y) \in \check{I}^+$ (in other words, $\check{I}^-(x) \subseteq \check{I}^-(y)$).

Clearly, this relation is maximal such that Definition 2.2 is satisfied. Moreover, for any $p \in M$, (p,p) satisfies (i) and (ii), meaning that d^+ is reflexive. To show transitivity, assume $(x,y) \in d^+$ and $(y,z) \in d^+$. Then $\check{I}^+(z) \subseteq \check{I}^+(y) \subseteq \check{I}^+(x)$ and $\check{I}^-(x) \subseteq \check{I}^-(y) \subseteq \check{I}^-(z)$, hence $(x,z) \in d^+$. □

For spacetimes with $\mathcal{C}^2$-metrics, the existence of such a maximal relation (i.e., of d^+ according to our notation) was already noted by Minguzzi [Min08b], who defined

$$D^+ := \left\{(p,q) \in M \times M \mid q \in \overline{\check{I}^+}(p),\ p \in \overline{\check{I}^-}(q)\right\}, \tag{2.1}$$

maximality with respect to push-up being a consequence (rather than the defining property) in this case [Min08b, Lem. 2.8]. In the $\mathcal{C}^2$- (but *not* in the $\mathcal{C}^0$-) case, (2.1) is equal to

$$\left\{(p,q) \in M \times M \mid q \in \overline{J^+}(p),\ p \in \overline{J^-}(q)\right\},$$

which is subsequently taken to be the definition in the review article [Min19a, Sec. 4.1]. Note also that in the $\mathcal{C}^2$-case, the distinction between I^+ and $\check{I}^+$ is irrelevant. To avoid possible confusion, we adopt the lower-case notation d^+, although we now prove that $d^+ = D^+$ also on $\mathcal{C}^0$-spacetimes.

Lemma 2.5. *For D^+ as given by (2.1), it holds that $d^+ = D^+$. In particular, $d^+ \subseteq \overline{\check{I}^+}$.*

Proof. Suppose $(p,q) \in d^+$. Let $\gamma : [0,1) \to M$ be any timelike curve with $\gamma(0) = q$. Then for all $t \in (0,1)$, $(q,\gamma(t)) \in \check{I}^+$. Because $(p,q) \in d^+$, the push-up property implies $(p,\gamma(t)) \in \check{I}^+$. Since γ is continuous, $(p,\gamma(t)) \to (p,q)$ as $t \to 0$, hence $q \in \overline{\check{I}^+}(p)$. Similarly, $p \in \overline{\check{I}^-(q)}$, so we conclude that $(p,q) \in D^+$ and hence $d^+ \subseteq D^+$.

On the other hand, suppose that $(p,q) \in D^+$, and let $r \in \check{I}^+(q)$. By assumption, $q \in \overline{\check{I}^+}(p)$, so we can approximate q by a sequence $(q_i)_i$ in $\check{I}^+(p)$. By openness of $\check{I}^+$, $r \in \check{I}^+(q_i)$ for some large enough i. Then, by transitivity, we conclude that $r \in \check{I}^+(p)$. Since r was arbitrary, we conclude that $\check{I}^+(q) \subseteq \check{I}^+(p)$. Similarly, $\check{I}^-(p) \subseteq \check{I}^-(q)$, so D^+ satisfies push-up with respect to $\check{I}^+$, and hence $D^+ \subseteq d^+$.

Finally, note that $D^+ \subseteq \overline{\check{I}^+}$ as a direct consequence of (2.1). □

We end this subsection by showing how the d^+-relation fits into the current literature. A $\mathcal{C}^0$-spacetime (M,g) is said to be *weakly distinguishing* whenever, for all $p,q \in M$, $\check{I}^+(p) = \check{I}^+(q)$ and $\check{I}^-(p) = \check{I}^-(q)$ together imply $p = q$ [Min19a, Def. 4.47]. Given two relations R,S on (M,g) (or any set, for that matter), we say that the pair (R,S) is a *causal structure* in the sense of Kronheimer and Penrose [KP67, Def. 1.2] if:

(i) S is transitive, reflexive and antisymmetric,

(ii) R is contained in S and irreflexive,

(iii) S satisfies push-up relative to R.

The following proposition was shown by Minguzzi for $\mathcal{C}^2$-spacetimes (phrased in terms of D^+ and with a slightly different proof, see [Min19a, Thm. 4.49] or [Min08b, Lem. 2.7]).

Proposition 2.6. *Let (M, g) be a $\mathcal{C}^0$-spacetime. Then $(\check{I}^+, d^+)$ is a causal structure in the sense of Kronheimer and Penrose if and only if (M, g) is weakly distinguishing.*

Proof. Point (iii) is satisfied by the very definition of d^+. To see point (ii), recall that if M is weakly distinguishing, then M is chronological, i.e., it does not contain closed timelike curves (and hence $\check{I}^+$ is irreflexive). Indeed, suppose $\gamma : [a, b] \to M$ is a closed timelike curve. Then, by transitivity, $\check{I}^\pm(\gamma(t)) = \check{I}^\pm(\gamma(s))$ for all $s, t \in [a, b]$. On the other hand, γ must be non-constant, in contradiction to weak distinction. That $\check{I}^+$ is contained in d^+ is clear because $\check{I}^+$, being transitive, must satisfy push-up with respect to itself . Regarding point (i), since d^+ is always transitive and reflexive by Lemma 2.4, it only remains to show that d^+ is antisymmetric.

By the proof of Lemma 2.4, $(p, q) \in d^+$ if and only if $\check{I}^+(q) \subseteq \check{I}^+(p)$ and $\check{I}^-(p) \subseteq \check{I}^-(q)$. Hence, $(p, q) \in d^+$ and $(q, p) \in d^+$ if and only if $\check{I}^+(p) = \check{I}^+(q)$ and $\check{I}^-(p) = \check{I}^-(q)$. But $p = q$ for all such pairs (p, q) if and only if (M, g) is weakly distinguishing. □

Another way of phrasing the last result in the usual language of causality theory is to say that "(M, g) is d^+-causal if and only if it is weakly distinguishing".

2.2.3 The local structure of d^+

Given a neighborhood $U \subseteq M$, we can define the localized relations $\check{I}^+_U$, J^+_U and d^+_U by applying the usual definitions to the spacetime $(U, g|_U)$. It is easy to see that $\check{I}^+_U \subseteq \check{I}^+$ and $J^+_U \subseteq J^+$.

Lemma 2.7. *Let $U \subseteq M$ be an open neighborhood. Then $d^+_U \subseteq d^+$.*

Proof. By definition, d^+_U satisfies push-up relative to $\check{I}^+_U$. We need to show that d^+_U also satisfies push-up relative to $\check{I}^+$, and then the claim follows from maximality of d^+. Suppose that $(x, y) \in d^+_U$ and $(y, z) \in \check{I}^+$. Then there exists a timelike curve $\gamma : [0, 1] \to M$ from y to z. Since, by assumption, $y \in U$, we must have that for $\epsilon > 0$ small enough, $\gamma|_{[0,\epsilon]} \in U$. Thus we have $(y, \gamma(\epsilon)) \in \check{I}^+_U$, which implies $(x, \gamma(\epsilon)) \in \check{I}^+_U \subseteq \check{I}^+$ by definition of d^+_U. Since also $(\gamma(\epsilon), z) \in \check{I}^+$, transitivity of $\check{I}^+$ implies $(x, z) \in \check{I}^+$. Part (ii) of Definition 2.2 can be shown analogously. □

We want to investigate whether, for U small enough, d^+_U is closed. The motivation lies in the limit curve theorems (see Remark 2.12 for the details). By Lemma 2.5, $d^+_U \subseteq \overline{\check{I}^+_U}$. Since also $\check{I}^+_U \subseteq d^+_U$, we conclude that d^+_U is closed if and only if $d^+_U = \overline{\check{I}^+_U}$. By Definition 2.3, $d^+_U = \overline{\check{I}^+_U}$ if and only if $\overline{\check{I}^+_U}$ satisfies push-up. Unfortunately, the next example [Gra+20, Ex. 3.1], shows that the latter is not necessarily the case.

Example 2.8. Let $M = \mathbb{R}^2$ with metric given by

$$g_\alpha := -\sin(2\theta(x))\,dt^2 - 2\cos(2\theta(x))\,dxdt + \sin(2\theta(x))\,dx^2$$

where

$$\theta(x) := \begin{cases} 0, & x < -1 \\ \arccos|x|^\alpha, & -1 \leq x \leq 0 \\ \frac{\pi}{2}, & x > 0, \end{cases}$$

and $0 < \alpha < 1$ arbitrary. The metric g_α is α-Hölder continuous, and in fact smooth outside of $\{x = -1\} \cup \{x = 0\}$. This example was introduced by Grant et al. [Gra+20, Ex. 3.1], who showed that $\check{I}^+ \subsetneq I^+$.

Let $p = (0,0)$ and $U \subseteq M$ any open neighborhood of p. Then the following hold:

(i) I^+_U is not open,

(ii) d^+_U is not closed.

Point (i) is shown in [Gra+20, Ex. 3.1] (they in fact show that I^+ is not open, but their argument is also valid on neighborhoods).

In order to show point (ii), first note that the past $\check{I}^-(p)$ (blue region in Figure 2.1) is contained in $\{x > 0\}$. This is so because a timelike $\mathcal{C}^1_{\mathrm{pw}}$-curve must have a timelike tangent vector everywhere, which implies having a positive x-component when in $\{x \geq 0\}$. When using *Leb*-curves, the past set $I^-(p)$ is in fact bigger [Gra+20, Ex. 3.1], but we will not discuss this further.

Consider the curve $\gamma : (-\epsilon, 0) \to M$, $s \mapsto (t(s), x(s))$ given by

$$t(s) := \frac{1}{1-\alpha} A^{1-\alpha} s, \qquad\qquad x(s) := -A|s|^{\frac{1}{1-\alpha}},$$

where $A > 0$ is arbitrary and $\epsilon > 0$ is small. It is shown in [Gra+20, Ex. 3.1] that γ is a timelike curve in the $\mathcal{C}^1_{\mathrm{pw}}$-sense (but its extension to the endpoint $s = 0$ is not). Since $\gamma(s) \to p$ as $s \to 0$, we conclude that $(\gamma(-\epsilon'), p) \in \overline{\check{I}^+}$ for all $\epsilon' < \epsilon$.

Next we show that $(\gamma(-\epsilon'), p) \notin d^+$. To see this, consider any point of the form $q = (x(-\epsilon'), t_q)$ with $t_q < t(-\epsilon')$ (see Figure 2.1). Then $(q, \gamma(-\epsilon')) \in \check{I}^+$, the connecting vertical segment being an example of timelike $\mathcal{C}^1_{\mathrm{pw}}$-curve between q and $\gamma(-\epsilon')$. If we assume $(\gamma(-\epsilon'), p) \in d^+$, then by push-up it follows that $(q, p) \in \check{I}^+$. However, $(q, p) \notin \check{I}^+$ because $x(-\epsilon') < 0$ and $\check{I}^-(p)$ is contained in $\{x > 0\}$. Hence $(\gamma(-\epsilon'), p) \notin d^+$.

Since we can pick $\epsilon' > 0$ arbitrarily small, and t_q smaller but arbitrarily close to $t(-\epsilon')$, the previous discussion applies to any neighborhood U of p. Thus $d^+_U \subsetneq \overline{\check{I}^+}_U$, no matter how we choose U.

Grant et al. showed that in the previous example also $\check{I}^+ \subsetneq I^+$, and that I^+ is not open (recall that $\check{I}^+$ is always open). If we were to define d^+ by requiring push-up with respect to I^+ instead of $\check{I}^+$, it may be that d^+_U is closed (for small enough U). We do not explore this possibility here, and simply note that this would be at

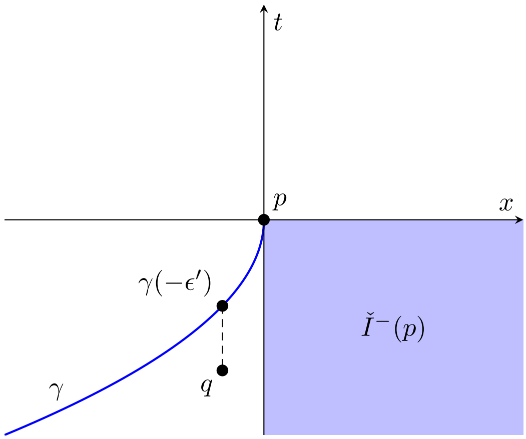


Figure 2.1: The spacetime of Example 2.8, with the curve γ that lies outside of $\check{I}^-(p)$, while nonetheless $(\gamma(-\epsilon'), p) \in \overline{\check{I}^+}$.

the cost of chronological futures not being open. Hence the conclusion is, either way, that one cannot have push-up, open futures and (local) closedness at the same time. That is, at least, if one wants the chronological relation to be given by the usual I^+ or $\check{I}^+$. We explore alternatives to I^+ and $\check{I}^+$ in Section 2.4, but the conclusion there is also that one of the three properties has to be sacrificed.

2.3 Causal curves in terms of the d^+-relation

In this subsection, we define a variation on the d^+-relation, which we call $\tilde{d}^+$, based on the notion of *d^+-causal curves.* The motivation for this is two-fold. Firstly, we will show that $\tilde{d}^+$, unlike d^+, has the desirable property that $\tilde{d}^+ = J^+$ on smooth spacetimes. Secondly, we are interested in studying the potential validity of limit curve theorems for d^+-causal curves. Throughout this section, F denotes an interval, meaning any connected subset of $\mathbb{R}$.

Definition 2.9. A continuous curve $\gamma : F \to M$ is called *d^+-causal* if for every $t \in F$ and every open neighborhood $U \subseteq M$ of $\gamma(t)$, there exists an open neighborhood $V \subseteq F$ of t such that

$$s_1 < s_2 \implies (\gamma(s_1), \gamma(s_2)) \in d_U^+ \quad \text{for all} \quad s_1, s_2 \in V.$$

Remark 2.10. Similarly to Definition 2.9, one can define J^+-causal curves, as was done already by Hawking and Ellis in 1973 [HE73, Chap. 6.2], and K^+-causal curves [SW96, Def. 17]. Any g-causal curve is automatically also J^+-causal. However, the converse is not true, since a J^+-causal curve may not even have a well defined tangent vector. Nonetheless, if two points $p, q \in M$ can be joined by a J^+-causal curve γ, then $(p, q) \in J^+$. In particular, there exists a g-causal curve σ, not necessarily equal to γ, which joins p and q.

Similarly to the previous remark, by transitivity and Lemma 2.7 it follows that if two points $p, q \in M$ can be joined by a d^+-causal curve, then $(p, q) \in d^+$. The next example motivates why Definition 2.9 has to be formulated in a local way, i.e. why we do not simply require $s_1 < s_2 \implies (\gamma(s_1), \gamma(s_2)) \in d^+$ for all $s_1, s_2 \in F$.

Example 2.11. Let $M = S^1 \times \mathbb{R}$ with metric $ds^2 = -dt^2 + dx^2$. This spacetime is totally vicious, so $\check{I}^+ = M \times M$ and hence also $d^+ = M \times M$. Therefore, any $\mathcal{C}^0$-curve $\gamma : F \to M$ satisfies $(\gamma(s_1), \gamma(s_2)) \in d^+$ for all $s_1, s_2 \in F$. However, M locally looks like Minkowski spacetime, where $d^+ = J^+$, hence not all curves on M are d^+-causal in the sense of Definition 2.9.

Having defined d^+-causal curves, we briefly return to Example 2.8 in order to better understand the relationship between closedness of d^+ and limit curve theorems.

Remark 2.12 (On limit curve theorems)**.** Suppose that $(\gamma_n)_n$ is a sequence of d^+-causal curves converging pointwise to a $\mathcal{C}^0$-curve $\gamma_\infty : F \to M$. Suppose that for every $t \in F$, there exists a neighborhood $U \subseteq M$ of $\gamma_\infty(t)$ such that d^+_U is closed. Then the curve γ_∞ is d^+-causal, because any pair of points on γ_∞ can be written as a limit of pairs of points on γ_n.

In Example 2.8, we showed that the point $p = (0, 0)$ does not admit any neighborhood U such that d^+_U is closed. Let γ be as in Example 2.8, and consider the sequence of curves given by $\gamma_n = \gamma|_{(-\epsilon, 1/n]}$. The sequence $(\gamma_n)_n$ converges pointwise (even uniformly, after appropriate reparametrization) to a curve $\gamma_\infty : (-\epsilon, 0] \to M$ which is just γ with p added as its endpoint. However, we showed in Example 2.8 that $(\gamma(t), p) \notin d^+$ for all $-\epsilon < t < 0$, hence γ_∞ is not d^+-causal.

Moving on, we use d^+-causal curves to define a new causal relation on M.

Definition 2.13. We define the $\tilde{d}^+$-relation as follows: $(p, q) \in \tilde{d}^+$ if there exists a d^+-causal curve from p to q.

By Lemma 2.7 and transitivity of d^+, $\tilde{d}^+ \subseteq d^+$. In particular, $\tilde{d}^+$ satisfies push-up relative to $\check{I}^+$. It is also clear that the concatenation of two d^+-causal curves is again d^+-causal, hence $\tilde{d}^+$ is transitive. Example 2.11 shows that $\tilde{d}^+$ can be strictly smaller than d^+. We finish this section with one of the main results of the paper, namely that $\tilde{d}^+ = J^+$ on smooth (and more generally, causally plain) spacetimes. Recall that a $\mathcal{C}^0$-spacetime is called causally plain if J^+ satisfies push-up relative to $\check{I}^+$.

Lemma 2.14. *Let (M, g) be a $\mathcal{C}^0$-spacetime, and $\gamma : F \to M$ a d^+-causal curve. Then γ is also J^+-causal.*

Proof. Let $t \in F$ be arbitrary. By [CG12, Proposition 1.10], there exists a neighborhood U of $p := \gamma(t)$, called a cylindrical neighborhood, such that $\overline{\check{I}^\pm_U(p)} \subseteq J^\pm_U(p)$ (here we mean the closure of the set $\check{I}^\pm_U(p) \subseteq U$). Let $V \in F$ be a neighborhood of t as in Definition 2.9, and $s \in V$. Suppose $s \geq t$, the other case being analogous. Because γ is d^+-causal, we have $(p, \gamma(s)) \in d^+_U$. Let $\sigma : [0, \epsilon) \to U$ be any timelike $\mathcal{C}^1_{\mathrm{pw}}$-curve with $\sigma(0) = \gamma(s)$. By push-up, $\mathrm{Im}(\sigma) \subseteq \check{I}^+(p)$. By continuity, $\sigma(u) \to \gamma(s)$ as $u \to 0$. Hence, by the previous and our choice of U,

$\gamma(s) \in \overline{\check{I}^+_U(p)} \subseteq J^+_U(p)$. In other words, $(\gamma(t), \gamma(s)) \in J^+_U$. Since t, s are arbitrary (as long as they are close enough), we conclude that γ is J^+-causal. □

Theorem 2.15. *Let (M, g) be a causally plain $\mathcal{C}^0$-spacetime. Then $\tilde{d}^+ = J^+$.*

Proof. If (M, g) is causally plain, then J^+ satisfies push-up, hence $J^+ \subseteq d^+$. In particular, on a subset $U \subset M$, we have $J^+_U \subseteq d^+_U$. Assume $(p, q) \in J^+$. Then there exists a g-causal curve $\gamma : [a, b] \to M$ from p to q. By continuity, for every $t \in [a, b]$ and every neighborhood U of $\gamma(t)$, there exists a neighborhood $V \subseteq [a, b]$ of t small enough such that $\gamma|_V$ is contained in U. If $s_1, s_2 \in V$ and $s_1 < s_2$, then $(\gamma(s_1), \gamma(s_2)) \in J^+_U \subseteq d^+_U$. Thus γ is a d^+-causal curve, and since γ is arbitrary, we conclude that $J^+ \subseteq \tilde{d}^+$. The other inclusion follows from Lemma 2.14, by noting that if two points p, q can be joined by a J^+-causal curve, then $(p, q) \in J^+$. □

2.4 Other causal structures

It is possible to repeat the procedure of Section 2.3 for Sorkin and Woolgar's K^+, and define a relation $\tilde{K}^+$ based on K^+-causal curves (the latter class of curves is also studied in [SW96, Section 3]). On a smooth spacetime, every point admits an arbitrarily small neighborhood U (a convex normal neighborhood) such that $J^+_U = \overline{\check{I}^+_U}$. Thus we conclude that on smooth spacetimes, $\tilde{K}^+ = J^+$. Unfortunately, Example 2.8 and Remark 2.12 tell us that $\tilde{K}^+$ cannot satisfy push-up with respect to $\check{I}^+$ on all $\mathcal{C}^0$-spacetimes. That is because, if it did, then $\tilde{K}^+ \subseteq d^+$. Recall that K^+ is closed and contains $\check{I}^+$. It follows that if a curve γ is the limit of a sequence of $\mathcal{C}^1_{\text{pw}}$-timelike curves, then γ must be d^+-causal. However, in Remark 2.12, we saw an example of such a γ where the endpoints are not d^+-related to each other, a contradiction.

A different approach is to consider J^+ to be the more fundamental relation, and then find an appropriate notion of chronological order, say $\mathcal{I}^+$. Ideally, we would like all of the following three properties to hold.

(i) $\mathcal{I}^+$ is open and contained in J^+.

(ii) J^+ satisfies push-up relative to $\mathcal{I}^+$.

(iii) For every point $p \in M$ and every neighborhood U of p, $\mathcal{I}^\pm(p) \bigcap U \neq \emptyset$.

Proposition 2.16. *Given J^+, if there exists a relation $\mathcal{I}^+$ satisfying (i), (ii) and (iii) above, then $\mathcal{I}^+ = \operatorname{Int} J^+$.*

Proof. By property (i), it is clear that $\mathcal{I}^+ \subseteq \operatorname{Int} J^+$. Suppose there exists $(p, q) \in \operatorname{Int} J^+ \setminus \mathcal{I}^+$. By propery (iii), with U a neighborhood of q, we can find $r \in U \cap \mathcal{I}^-(q)$. We may choose U small enough so that $\{p\} \times U \subseteq \operatorname{Int} J^+$, meaning in particular that $(p, r) \in J^+$. Then, by property (ii), we have that $(p, q) \in \mathcal{I}^+$, obtaining a contradiction. Therefore $\mathcal{I}^+ = \operatorname{Int} J^+$. □

Next, we present an example where J^+ does not satisfy push-up relative to $\operatorname{Int} J^+$ (point (ii)). In view of the previous proposition, we conclude that there cannot exist a relation $\mathcal{I}^+$ satisfying (i)-(iii) above.

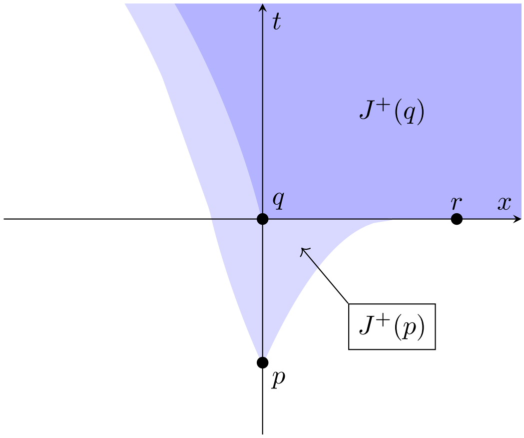


Figure 2.2: The points p, q, r in Example 2.17, which satisfy $(p, q) \in \operatorname{Int} J^+$ and $(q, r) \in J^+$ but $(p, r) \notin \operatorname{Int} J^+$.

Example 2.17. This example is adapted from [CG12, Ex. 1.11] and [Lin20, Sec. 4.1]. Let $M = (-2, 2) \times \mathbb{R}$ with metric given by

$$ds^2 = -dt^2 - 2(1 - |t|^{1/2})dtdx + |t|^{1/2}(2 - |t|^{1/2})dx^2.$$

This metric is smooth everywhere except on the x-axis. A null curve starting at the point $p = (-1, 0)$ can be parametrized as $x \mapsto \gamma(x) = (t(x), x)$, and then

$$\dot{t} = \begin{cases} |t|^{1/2} & \text{or} \\ |t|^{1/2} - 2. \end{cases}$$

By solving this equation, we obtain the boundary of $J^+(p)$ (light blue region in Figure 2.2). Consider the first case of the equation, which is when the null curve γ moves upwards and to the right. We are interested in finding the value x_1 such that $t(x) \to 0$ as $x \to x_1$. It can easily be computed by separation of variables:

$$x_1 = \int_0^{x_1} dx = \int_{-1}^{0} \frac{dt}{|t|^{1/2}} = 2.$$

Let $q = (0, 0)$, then $(p, q) \in \operatorname{Int} J^+$. Consider a third point $r = (0, 3)$, as in Figure 2.2. Now $(q, r) \in J^+$, because the curve $x \mapsto (0, x)$ is null. However, $(p, r) \notin \operatorname{Int} J^+$, since there are points of the form $(-\epsilon, 3)$ arbitrarily close to r which cannot lie in $J^+(p)$ because they lie below the x-axis and their x-coordinate is larger than 2. Hence J^+ does not satisfy push-up relative to $\operatorname{Int} J^+$ in this example.

Finally, we point out yet another option, which is to define chronological futures via the Lorentzian distance. We recall the basic properties of the Lorentzian distance, and refer to [BEE96, Chap. 4] for further details. The length of a causal curve $\gamma : [a, b] \to M$ is defined as

$$L_g(\gamma) := \int_a^b \sqrt{-g(\dot\gamma, \dot\gamma)}ds,$$

where it is enough for the integrand to be defined almost everywhere. The Lorentzian distance between two points $p, q \in M$ is then defined by

$$d(p, q) := \sup \{L_g(\gamma) \mid \gamma \text{ a } g\text{-causal curve with endpoints } p, q\} \tag{2.2}$$

if $(p, q) \in J^+$, and $d(p, q) := 0$ otherwise. Recall that when the spacetime metric g is smooth, the Lorentzian distance d satisfies $d(p, q) > 0 \iff (p, q) \in \check{I}^+$ [BEE96, Eqn. 4.2]. On $\mathcal{C}^0$-spacetimes, only the " $\Longleftarrow$ " implication continues to hold. To see how " $\Longrightarrow$ " can fail, take the points p, q, r in Example 2.17 (depicted in Figure 2.2). We can connect p and q by a vertical segment, which is timelike (hence causal) and has length equal to 1. We can connect q and r by a horizontal segment, which is also causal, and has length 0. Concatenating the two segments, we get a causal curve of length 1 from p to r, despite the fact that $(p, r) \notin \check{I}^+$. Further, we see that $r \in \partial J^+(p)$, and since $\{d > 0\} \subseteq J^+$ by definition, we conclude that $\{d > 0\}$ is not open in this example. Nonetheless, another property of the Lorentzian distance, the inverse triangle inequality

$$d(p, r) \geq d(p, q) + d(q, r) \ \text{ if } (p, r), (r, q) \in J^+,$$

does hold for all $\mathcal{C}^0$-metrics, since it is a direct consequence of (2.2) and the fact that the concatenation of two causal curves is causal. We deduce from the inverse triangle inequality that J^+ satisfies push-up relative to $\{d > 0\}$. Hence the combination of J^+ and $\{d > 0\}$ gives us push-up and limit curve theorems, but at the price of non-open futures.

2.5 Conclusions

Table 2.1 summarizes the properties of different choices of causal and chronological relation on $\mathcal{C}^0$-spacetimes. While each of the choices (rows) is distinct from the others for $\mathcal{C}^0$-metrics, they all coincide for smooth metrics. In particular, in the smooth case, the standard causal structure ticks all three boxes. For $\mathcal{C}^0$-metrics, on the other hand, no combination of chronological and causal order has all of the three properties that we considered. The newly introduced $(\check{I}^+, \tilde{d}^+)$ is the only causal structure that has both push-up and an open chronological relation. Moreover, it defines a causal structure in the sense of Kronheimer and Penrose (see Proposition 2.6).

It is fair to say that we have exhausted all reasonable possibilities. For if we want the chronological relation to be given by timelike $\mathcal{C}^1_{\mathrm{pw}}$-curves, then Example 2.8 and Remark 2.12 tell us that no causal relation can satisfy push-up and at the same time admit a limit curve theorem. We do not know if this changes when using timelike $\mathcal{L}$-curves equipped with their classical differential, but even if so,

Table 2.1: Comparison of different causal structures on $\mathcal{C}^0$-spacetimes by three important properties. The (non-)openness of $\check{I}^+$ and I^+ was established by Grant et al. [Gra+20], and the push-up and limit curve properties of J^+ by Chruściel and Grant [CG12]. The rest of entries in the table were, to the best of our knowledge, never explicitly stated before.

Chronological order	Causal order	Push-up	Open futures	Limit curves
$\check{I}^+$	J^+	✗	✓	✓
$\check{I}^+$	$\tilde{K}^+$	✗	✓	✓
$\check{I}^+$	$\tilde{d}^+$	✓	✓	✗
I^+	J^+	✗	✗	✓
I^+	$\tilde{K}^+$	?	✗	✓
I^+	$\tilde{d}^+$	✓	✗	?
$\{d > 0\}$	J^+	✓	✗	✓
$\operatorname{Int} J^+$	J^+	✗	✓	✓

we would lose the openness of chronological futures instead [Gra+20]. If, on the other hand, we choose for the causal relation to be given by g-causal curves, and try to define a compatible chronological relation, we run into the same problems by the discussion in Section 2.4.

Chapter 3

Causal bubbles in globally hyperbolic spacetimes

This chapter is based on the article [GHS22] of the same title, published in *General Relativity and Gravitation* together with Elefterios Soultanis.[1][2]

3.1 Introduction

Spacetimes where the Lorentzian metric is merely continuous often appear in mathematical General Relativity as weak solutions of Einstein's Equations [BL14; GT87; Sbi18]. Chruściel and Grant [CG12] were the first to study their causal structure, discovering a phenomenon called *causal bubbling* (see below). Later, Sämann [Säm16] studied globally hyperbolic spacetimes with continuous metrics in detail and showed that, just as in the smooth case, global hyperbolicity can be equivalently characterized by any of the following:

(i) Non-total imprisonment and compact causal diamonds,

(ii) Existence of a Cauchy hypersurface,

(iii) Existence of a Cauchy time function.

In the same paper (see Section 7) Sämann raises the question whether globally hyperbolic spacetimes can be causally bubbling. In this short note we give an example of a **globally hyperbolic spacetime with continuous metric** $-\mathrm{d}t^2 + \rho \mathrm{d}x^2$ **exhibiting causal bubbling**.

We define causal bubbling on a spacetime (M, g) with continuous metric g as follows. Set

$$\mathcal{B}^{\pm}(p) = J^{\pm}(p) \setminus \overline{I^{\pm}(p)}$$

[1] Radboud University. Present address: University of Warwick.

[2] We thank Annegret Burtscher, Eric Ling and Clemens Sämann for interesting discussions and comments.

The set $\mathcal{B}^{\pm}(p)$ is called a (future/past) *causal bubble* if it is non-empty.[3] Here $I^{\pm}(p)$ and $J^{\pm}(p)$ are the timelike and causal future/past cones defined using *Lipschitz* (or equivalently absolutely continuous) curves. We refer to [Gra+20] for a discussion on timelike and causal cones defined via smooth curves and their relationship with the notion used here. See also [GH21; Lin20] and [KS18, Section 5.1] for further analyses on causality with continuous metrics.

Our example stands out from previously known examples of causally bubbling spacetimes for two reasons.

(a) It is manifestly globally hyperbolic,

(b) The metric splits orthogonally into a timelike and a spacelike part.

The second statement (b) would follow automatically from (a) if the metric were at least $\mathcal{C}^2$ [BS05], but not if it is merely continuous. This is because the splitting in [BS05] is realized by flowing along the gradient vector field of a temporal function τ. The gradient vector field depends on both τ and g, and hence need not be sufficiently regular even if τ is smooth. In particular, while examples 3.1 and 3.2 in [Gra+20] (which exhibit *internal* bubbling) likely are globally hyperbolic, it is not clear whether the metric splits orthogonally as in (b). Whether Example 1.11 in [CG12] is globally hyperbolic is less clear, but it is known to be strongly causal (see [KS18, Section 5.1]).

3.2 The example

Consider the $(1 + 1)$-dimensional spacetime $\mathbb{R}^2$ equipped with the continuous Lorentzian metric

$$g := -\mathrm{d}t^2 + \rho(t,x)\mathrm{d}x^2, \quad \rho(t,x) := 1 + \sqrt{(t-|x|)_+}. \tag{3.1}$$

With the natural choice of time orientation, t is a time function. Since the light-cones of the metric (3.1) are narrower than those of the Minkowski metric, t is a Cauchy time function, and hence the spacetime is globally hyperbolic.

To see that the causal future of the origin contains a causal bubble, we begin by considering the ODE for the null curves $\gamma(s) = (\alpha(s), \beta(s))$:

$$0 = -\alpha'(s)^2 + \left(1 + \sqrt{(\alpha(s) - |\beta(s)|)_+}\right)\beta'(s)^2.$$

A null curve γ starting at the origin, parametrized as $(\alpha(s), s)$ thus satisfies $\alpha'(s)^2 = 1 + \sqrt{(\alpha(s) - s)_+}$ or, by denoting $y(s) = \alpha(s) - s \geq 0$,

$$y'(s) + 1 = \sqrt{1 + \sqrt{y(s)}}, \quad y(0) = 0. \tag{3.2}$$

[3] This definition corresponds to *external* bubbling in [Gra+20] and is equivalent to the failure of the *push up* property, see [Gra+20, Theorem 2.12].

In addition to the trivial solution $y \equiv 0$, the initial value problem (3.2) also admits another solution, expressed in implicit form as

$$s = \frac{4}{3}\left[\left(1+\sqrt{y(s)}\right)^{3/2} - 1\right] + 2\sqrt{y(s)}.$$

Indeed, differentiating both sides yields

$$\begin{aligned} 1 &= 2\left(1+\sqrt{y(s)}\right)^{1/2}\frac{\mathrm{d}}{\mathrm{d}s}\sqrt{y(s)} + 2\frac{\mathrm{d}}{\mathrm{d}s}\sqrt{y(s)} \\ &= y'(s)\frac{\sqrt{1+\sqrt{y(s)}}+1}{\sqrt{y(s)}} = \frac{y'(s)}{\sqrt{1+\sqrt{y(s)}}-1}. \end{aligned}$$

Denoting

$$f(y) := \frac{4}{3}\left[(1+\sqrt{y})^{3/2} - 1\right] + 2\sqrt{y},$$

we conclude that $\gamma = (s + f^{-1}(s), s)$ is a null curve as well as the straight line given by $\gamma = (s, s)$. In fact there is a 1-parameter family of null curves starting at 0, given by

$$\gamma_u(s) := \begin{cases} (s, s) & \text{for } 0 \le s < u, \\ (s + f^{-1}(s-u), s) & \text{for } s \ge u, \end{cases} \tag{3.3}$$

where $u \in [0, \infty]$. One can check that the curves γ_u are smooth for all parameter values except $s = u$, where they are only $\mathcal{C}^{1,1}$ regular. Note also that on smooth, two-dimensional spacetimes, every null curve is a null geodesic. Hence, for $u < \infty$, $\gamma_u|_{(u,\infty)}$ is a null geodesic, as it is contained in the region where the metric is smooth. At $s = u$, however, γ_u is not locally length maximizing.

Proposition 3.1. *The non-empty open set* $A := \{(t,x) : x > 0,\ 0 < t - x < f^{-1}(x)\} \subset \mathbb{R}^2$ *consists of points in* $J^+(0) \setminus \overline{I^+(0)}$.

Consequently $\mathbb{R}^2$ equipped with the globally hyperbolic metric (3.1) contains causal bubbles.

Proof. The inclusion $A \subset J^+(0)$ is clear. Let $p \in A$. We first show that $p \notin I^+(0)$. Since the set A is foliated by the null curves (3.3), there is a unique $u \in (0, \infty)$ such that γ_u passes through p. Because the metric is smooth outside of $t = \pm x$, γ_u is a null geodesic generating the boundary of $I^-(p)$, at least from (u, u) until p. This means that any past-directed timelike curve σ starting at p must intersect the diagonal at some point $(\bar{u}, \bar{u})$ with $\bar{u} \ge u > 0$ (see Figure 3.1). It follows that σ cannot reach 0; indeed, following the diagonal would introduce a null piece, while leaving the diagonal violates the causality of the curve (the metric (3.1) has narrower lightcones than those of the Minkowski metric). We conclude that $0 \notin I^-(p)$, (equivalently, $p \notin I^+(0)$) and hence $A \cap I^+(0) = \emptyset$. But since A is open, it cannot contain any boundary points of $I^+(0)$ either, hence also $A \cap \overline{I^+(0)} = \emptyset$, concluding the proof. $\square$

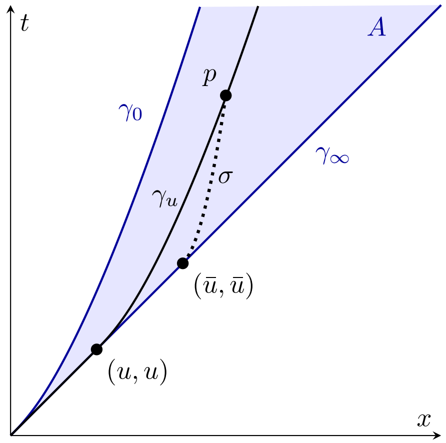


Figure 3.1: The causal bubble A as in Proposition 3.1.

3.3 Discussion

3.3.1 Strong energy condition

As in previously known examples [CG12; GH21; Gra+20; Lin20] causal bubbling arises from the branching of null geodesics. Chruściel and Grant noted [CG12, Rem. 1.19] that, in the Riemannian case, branching is associated with curvature being unbounded from below. Indeed, in our example (in $\{t \neq |x|\}$) the Ricci scalar, given by

$$R = \begin{cases} -\frac{\rho+(\rho-1)^2}{4\rho^2(\rho-1)^3} & \text{if } t > |x| \\ 0 & \text{if } t < |x|, \end{cases}$$

diverges to $-\infty$ as $t \searrow |x|$. Note however that in dimension $1+1$ the Ricci *tensor* is given by $\mathrm{Ric} = \frac{1}{2}Rg$ and thus

$$\mathrm{Ric}(v, v) \geq 0 \quad \text{for all causal vectors } v \text{ in } \{t \neq |x|\}. \tag{3.4}$$

In other words the strong energy condition is satisfied away from $\{t = |x|\}$. Physically speaking causal bubbling appears to be a consequence of the presence of infinite (but positive) effective energy density (see [KS20] for an in-depth discussion of energy conditions).

3.3.2 Synthetic curvature bounds

Recently, a *synthetic* notion of *timelike curvature dimension (TCD) bounds* on (non-smooth) *Lorentzian pre-length spaces* has been put forth using optimal transport [CM20; McC20], in analogy with the very successful metric theory [AGS14; LV09; Stu06]. The *entropic convexity* condition defining $\mathrm{TCD}(K, N)$-spaces asks that the *Rényi entropy*

$$\mathrm{Ent}(f\,\mathrm{vol}) := \int f \log f \mathrm{d}\,\mathrm{vol}$$

is (K, N)-convex along timelike geodesics in a space of probability measures. We refer to [KS18, Definition 2.8] and [Bra23, Definition 2.17] (see also [CM20; CM22]) for the definitions and properties of Lorentzian pre-length spaces and (K, N)-convexity, respectively. On smooth spacetimes, such convexity properties characterize the strong energy condition, cf. [McC20].

Despite satisfying (3.4) the Rényi entropy associated to the volume measure of the metric (3.1) is not (K, N)-convex along Lorentz-Wasserstein geodesics, i.e. **the metric** (3.1) **does not satisfy the entropic convexity condition for any** (K, N).[4] This follows from the fact that $t \mapsto -\log \rho(t, x)$ is not (K, N)-convex for any $x \in \mathbb{R}$, cf. [Sou22].

However, restricted to a suitable subset, our example demonstrates that the TCD-condition does not prevent causal bubbling. Indeed, the closed set $Y := \{t \geq |x|\} \subset \mathbb{R}^2$ with the restriction of the metric (3.1) satisfies the (weak) entropic $(0, 2)$-convexity condition but contains causal bubbles. Notice that (Y, g) is obtained as the uniform pointwise limit of the sequence of smooth metrics

$$g_j := -\mathrm{d}t^2 + \rho_j(t, x)\mathrm{d}x^2, \quad \rho_j(t, x) = \rho(t + 1/j, x) \tag{3.5}$$

on Y, which all satisfy the strong energy condition (3.4) everywhere on Y and are thus $\mathrm{TCD}(0, 2)$-spaces.[5]

In closing we point out that, in the metric setting, branching of geodesics is excluded by the *Riemannian curvature-dimension* (RCD) condition but not by the (weak) CD-condition, see [Den21] and [Mag22], respectively.

Remark 3.2. Note that all the conclusions in this note, including those about curvature, are valid in higher dimensions as seen by considering metrics $g = -\mathrm{d}t^2 + \rho(t, |\boldsymbol{x}|)\mathrm{d}\boldsymbol{x}^2$ in $\mathbb{R}^{n+1}$.

[4] While Lorentzian pre-length spaces are required (by definition) to have the push up property the entropic convexity condition makes sense regardless of the validity of push up.

[5] Note that, while the spaces (Y, g_j) are Lorentzian *pre*-length spaces, every point on $\{t = |x|\} \subset Y$ has empty chronological past and thus they fail to be Lorentzian length spaces, cf. [KS18, Definition 3.22].

Chapter 4

Time functions on Lorentzian length spaces

This chapter is based on the preprint [BGH21], written together with Annegret Burtscher and submitted for publication.[1]

4.1 Introduction

On a smooth spacetime (M, g) a continuous function $t\colon M \to \mathbb{R}$ is called a *time function* if it satisfies

$$p < q \implies t(p) < t(q) \quad \text{for all} \quad p, q \in M,$$

where $p < q$ means that there exists a causal curve from p to q, and that $p \neq q$. Time functions play a crucial role in Lorentzian causality theory and Einstein's general theory of relativity.

The study of time functions has a long history in general relativity. Their origin can be traced back to the works of Geroch and Hawking in the late 1960s. Geroch introduced *volume time functions* by normalizing the volume of a spacetime to one, and by defining the time of a point p as the volume of its chronological past $I^-(p)$. In his seminal work [Ger70] from 1970, Geroch used these volume functions to characterize global hyperbolicity by the existence of Cauchy surfaces and to obtain a topological splitting. Global hyperbolicity is the strongest and most important causality condition in general relativity. Cauchy surfaces represent the natural sets to pose initial conditions for the Einstein equations (for a self-contained exposition see [Rin09]). Moreover, global hyperbolicity and its different characterizations play a crucial role in the singularity theorems of Penrose and Hawking (see [HE73, Section 8.2]), the Lorentzian splitting theorems (see [Bar88;

[1] LGH would like to thank Didier Solis for communication related to their preprint [ACS20]. The authors are also grateful to Ettore Minguzzi for feedback on the introduction. AB's research is supported by the Dutch Research Council (NWO), Project number VI.Veni.192.208.

Esc88; Gal89] and follow-up work) and the formulation of Quantum Field Theory on curved backgrounds [BGP07, Chapter 4].

Building upon Geroch's idea, Hawking [Haw68] showed that volume functions can be "smeared out" to obtain time functions at a significantly lower step on the causal ladder, namely stable causality (see also the work of Minguzzi [Min09; Min10] for the same result via the equivalent notion of K-causality and [Min18, Figure 20] for a depiction of the complete causal ladder). This result contributed to Hawking's program to find the minimal causality conditions that one should impose on a spacetime in order to consider it as physically reasonable. As an in-between result between stable causality and global hyperbolicity, Hawking and Sachs [HS74] showed in 1974 that Geroch's volume functions are continuous themselves (that is, without the use of an averaging procedure) precisely when the spacetime is causally continuous. Their proof, however, contained a loophole that was later filled by Dieckmann [Die87; Die88b].

After these foundational works, the question remained whether time functions and Cauchy surfaces can be chosen smooth, rather than just continuous. Despite several attempts by Seifert [Sei77], Sachs and Wu [SW77] and Dieckmann [Die87; Die88a], this problem remained open for decades. Only in the early 2000s it was firmly established by Bernal and Sánchez [BS03; BS05; BS06a] that a spacetime that admits a continuous time function also admits a smooth one, and that Geroch's topological splitting of globally hyperbolic spacetimes can be promoted to a smooth, orthogonal splitting. While their work differs significantly from previous approaches, it was recently established by Chruściel, Grant and Minguzzi [CGM16] that also a family of Geroch's time functions are continuously differentiable for globally hyperbolic metrics, and that Hawking's time functions can be smoothed out.

A radically different approach to show smoothness of time functions is that of Fathi and Siconolfi [Fat15; FS12], which uses weak KAM theory. It has the added advantage that it is formulated for smooth manifolds equipped with a continuous field of closed convex tangent cones, of which Lorentzian manifolds are just one class of examples. Note that also in the work of Chruściel, Grant and Minguzzi [CGM16], while staying in the more traditional framework of Lorentzian manifolds, the metric tensor can be of regularity as low as $C^{2,1}$. Similarly, global hyperbolicity was characterized by the existence of Cauchy hypersurfaces and of Cauchy time functions on manifolds with continuous Lorentzian metrics by Sämann [Säm16] building on work by Chruściel and Grant [CG12]. One should be aware that such low-regularity spacetimes already exhibit considerable pathological behavior including "causal bubbles" (failure of the push-up property) and not necessarily open chronological futures/pasts [CG12; GH21; Gra+20; Lin20; SW96]. Note that also the cone structure approach of Fathi and Siconolfi has been developed further in very recent works of Bernard and Suhr [BS18; BS20] (extending Conley theory to this setting) and Minguzzi [Min19a] (using traditional arguments). Besides the elimination of the need of a Lorentzian metric in the theory of closed (causal) cones, also the smooth manifold structure should be removable to some degree. Indeed, it was already pointed out by Minguzzi [Min13] that sufficient conditions for the existence of time functions can be obtained for more general topological spaces through Nachbin's theory of closed ordered spaces [Nac65].

It remained open, however, if and to what extent general (and in particular, all finer) results about the existence and properties of time functions known for smooth spacetimes rely on the causal and topological structure alone. Our present work answers this question fully in the abstract framework of Lorentzian (pre-)length spaces by characterizing the existence of several types of time functions with different steps on the causal ladder.

The framework of Lorentzian (pre-)length spaces is widely applicable as a broad range of singular and regular "spacetimes" appearing in the literature are encompassed, including the above-mentioned closed cone fields and causally plain continuous spacetimes. Introduced by Kunzinger and Sämann [KS18] in 2018, the notion of Lorentzian length spaces makes explicit what is already evident in the early works of Weyl, Penrose and what has also been proposed by various other authors (see, for instance, [Bom+87; BS90; BS06b; Bus67; EPS72; KP67; Wey88; Woo73] and [Min18, Section 4.2.4]), namely that one should treat the causal structure as the most fundamental geometric object in the general theory of relativity. At the same time, an attempt has been made to translate metric geometric techniques à la Gromov that have revolutionized Riemannian geometry to Lorentzian geometry. In this spirit, Lorentzian (pre-)length spaces are defined as essentially metric spaces equipped with a chronological and a causal order satisfying basic order-theoretic properties such as open chronological futures/pasts and the push-up property (see Section 4.2 for precise definitions). This general theory of "spacetime-like" spaces encompasses basically all previously mentioned low-regularity settings. Despite its youth the Lorentzian (pre-)length framework has already celebrated important successes in the context of causality theory [ACS20; KS18], inextendibility results [Ale+19; GKS19], synthetic curvature bounds [CM20; McC20; MS22c] (related to energy conditions in general relativity) and stability [AB22; KS22] with respect to the null distance.

Nonetheless, the question of when a time function on Lorentzian (pre-)length space exists has been neglected until now, the only exception being the recent work of Kunzinger and Steinbauer [KS22] where it is shown that the existence of certain time functions implies strong causality (the converse, however, is false even in the manifold setting). In the present work, we fill this gap by establishing several sharp existence results. The statements and a discussion of our main results follows.

Main results

In this paper we establish three major results relating the existence (and properties) of time functions to three different steps on the causal ladder, starting from the optimal condition for existence (K-causality) and building it up to the top one (global hyperbolicity). All our results in the main body of the paper are obtained for Lorentzian pre-length spaces obeying milder axioms than those required of a Lorentzian length space (all definitions are presented in Section 4.2 in a self-contained way). Establishing our results in the context of Lorentzian pre-length spaces is important because they are more widely applicable and often sufficient (see also [CM20; KS22]). In this introduction, however, we state a simplified version of our theorems in the setting of Lorentzian length spaces (at the end we briefly comment on the pre-length case). Generally it is useful to recall that on a smooth spacetime, the local causal and topological structure is very rigid (all neighborhoods look the same). Our proofs, on the other hand, rely almost exclusively on global arguments. This way, we can reduce the local assumptions to the bare minimum, doing away completely with the manifold structure.

Our first result characterizes the mere existence of time functions on Lorentzian (pre-)length spaces by K-causality, generalizing a result of Hawking [Haw68] and Minguzzi [Min10] for smooth spacetimes (note that K-causality is equivalent to stable causality in the smooth setting [Min09]).

Theorem 4.1. *Suppose X is a second countable, locally compact Lorentzian length space. Then X is K-causal if and only if X admits a time function.*

Here it is crucial that the K-relation is closed and transitive (by Definition 4.33). Then, as already pointed out by Minguzzi [Min13], the general theory of topological ordered spaces yields time functions on K-causal spaces. To prove the converse statement, namely that existence of a time function implies K-causality, we closely follow the approach of Minguzzi [Min10], which requires the use of limit curve theorems.

Our second result is concerned with an explicit construction of time functions on Lorentzian (pre-)length spaces that can be equipped with Borel probability measures. Here, we are influenced by Geroch's notion of volume functions [Ger70] as well as Hawking's averaging procedure [Haw68]. Since the boundaries of light cones, however, are no longer hypersurfaces with measure zero as in the smooth manifold setting, the definition of our *averaged volume functions* as well as the proof of their causal properties and continuity are significantly more involved. The result we obtain is essentially a generalization of a theorem by Hawking and Sachs [HS74] (and Dieckmann's rigorous follow-up work [Die88b]), and thus the corresponding step on the causal ladder is that of causal continuity.

Theorem 4.2. *Let X be a second countable, locally compact Lorentzian length space. Then X is causally continuous if and only if the averaged volume functions on X are time functions.*

While the assumption on second countability of the underlying metric space is crucial in order to have a suitable measure at hand, local compactness can be removed by using a weaker (but in the smooth case equivalent) notion of causal continuity.

Our third result characterizes globally hyperbolic Lorentzian (pre-)length spaces by the existence of Cauchy time functions, whose level sets are Cauchy sets that are intersected by every inextendible causal curve exactly once. The smooth spacetime analogue is the seminal 1970 result of Geroch [Ger70].

Theorem 4.3. *Let X be a second countable Lorentzian length space with a proper metric structure. Then the following are equivalent:*

(i) X is globally hyperbolic,

(ii) X is non-totally imprisoning and the set of causal curves between any two points is compact,

(iii) X admits a Cauchy set,

(iv) X admits a Cauchy time function.

To establish Theorem 4.3 we utilize our averaged volume functions introduced already for Theorem 4.2, as well as the behavior of inextendible causal curves. The use of *averaged* volume functions poses an additional difficulty in our proofs, compared to the smooth case. The other main challenge is the fact that our Cauchy sets are not hypersurfaces, and in fact very little can be deduced about their topology (hence the name Cauchy *set* instead of surface, see also our discussion below).

The above theorems follow immediately from their sharper versions, Theorems 4.34, 4.52 and 4.59, which are obtained for Lorentzian pre-length spaces. Some essential conditions (which are part of the axioms of Lorentzian length spaces), however, still need to be assumed. In Theorem 4.34 (generalizing Theorem 4.1), for instance, we need to additionally impose the existence of causal curves and their limit curves. The chronological relation, however, is not needed in the proof at all. The proof of Theorem 4.52 (corresponding to Theorem 4.2), on the other hand, does not require causal curves, but does make use of both the causal and chronological relation. Furthermore, the two relations need to satisfy a compatibility condition that we call "approximating", which simply means that the causal futures and pasts are contained in the closure of the chronological ones. Finally, Theorem 4.59 (the pre-length version of Theorem 4.3) builds upon Theorem 4.52, hence the "approximating" condition is also needed here. Moreover, causal curves are used already in the definition of Cauchy set and Cauchy time function, and the limit curve theorems will be important again.

Discussion and outlook

Since its inception, the framework of Lorentzian length spaces has seen a rapid expansion [ACS20; Ale+19; CM20; GKS19; KS18; KS22]. Notably, Cavalletti and Mondino [CM20] recently introduced a notion of Ricci curvature bounds for Lorentzian pre-length spaces, based on optimal transport theory, that mimics the strong energy condition of general relativity, and implies a version of Hawking's singularity theorem. Our work completes another important milestone in establishing the potential of this non-smooth causal theory by fully characterizing the existence of time functions in terms of the causal ladder. As an immediate application, we have now unambiguously established when one can make use of time functions to define the null distance of Sormani and Vega [SV16]. Very recently, this notion (as well as convergence) has been investigated by Kunzinger and Steinbauer [KS22] in the context of Lorentzian pre-length spaces (certain limits of examples in [AB22, Section 5] are also of this type). With our new characterizations, in particular, of global hyperbolicity via Cauchy time functions, more refined convergence/stability results may be obtained. Moreover, it should be straightforward to carry over the cosmological time function of Andersson, Galloway and Howard [AGH98] to the Lorentzian length space framework using the time separation function in place of the Lorentzian distance.

While in broad terms we show that the classical results about time functions admit direct generalizations for Lorentzian length spaces, we also find interesting and not so subtle differences. In particular, although global hyperbolicity is also characterized by the existence of Cauchy sets, these Cauchy sets need not be homeomorphic to each other. This is in stark contrast to the case of spacetimes, where Geroch's celebrated splitting theorem [Ger70] (later refined by Bernal and Sánchez [BS03]) shows that all Cauchy surfaces on a given spacetime must have the same topology. In fact, Geroch already showed in [Ger67] that transitions between compact spatial topologies not only contradict global hyperbolicity, but in fact even violate the most basic of all assumptions, chronology. While time travel is a no-go in any physically sound theory, Sorkin [Sor97] argues that topology change is a necessary feature of *any* convincing candidate theory of quantum gravity. Going beyond the setting of smooth Lorentzian manifolds is thus a necessity in order to admit topology change without violating chronology, a common approach being the use of degenerate Lorentzian metrics [Bor+99; Hor91]. Current proposals for quantum gravity also predict that physical spacetimes are represented by non-manifold-like structures at small scales, such as causal sets [Sur19], causal dynamical triangulations [Lol20], causal fermion systems [Fin18], or spin foams [NP07]. An interesting and important next step will be to see if and how the framework of Lorentzian length spaces fits into these quantum gravitational theories.

Outline

The paper is structured as follows. In Section 4.2, we give a self-contained account of the relevant aspects of the theory of Lorentzian (pre-)length spaces, drawing from the existing literature but also introducing new material. We then prove Theorems 4.1, 4.2 and 4.3 (more precisely, the corresponding sharper Lorentzian pre-length space versions) in Sections 4.3, 4.4 and 4.5, respectively.

4.2 Lorentzian pre-length spaces

In this section we recall, and partly refine, the definition of Lorentzian (pre-)length spaces, their causality conditions and the limit curve theorems. We use the notation and results of [ACS20; KS18]. A reader familiar with the smooth case will find that most classical concepts are defined in the same way for Lorentzian pre-length spaces (with the difference that some important properties do not follow automatically but have to be imposed separately, such as causal curves themselves).

In Section 4.2.1 we recall the definition of Lorentzian pre-length space, and of causal curve. This is standard material, except that we use a more precise nomenclature for inextendible causal curves (Definition 4.11). In Section 4.2.2 we revisit the limit curve theorems of Kunzinger and Sämann [KS18] in the slightly weaker framework of "local weak causal closedness" following a suggestion of Aké et al. [ACS20]. We also introduce the new notion of *Lorentzian pre-length spaces with limit curves*, which encompasses all necessary assumptions needed for the application of the limit curve theorems. In Section 4.2.3 we introduce the notion of *approximating Lorentzian pre-length space*, which can be seen as a much weaker version of Kunzinger and Sämann's "localizability". Most importantly, we show that our newly introduced properties are, in particular, satisfied by all Lorentzian length spaces. Finally, in Section 4.2.4, we define time functions and introduce some elements of causality theory for Lorentzian pre-length spaces. Notably, we give some new characterizations of non-total imprisonment, both in terms of causal curves and of time functions. Additional causality conditions, including K-causality and global hyperbolicity, are introduced in later sections when needed.

4.2.1 Basic definitions and properties

Definition 4.4 ([KS18, Definition 2.1])**.** A *causal set* is a set X equipped with a preorder $\leq$ (called *causal* relation) and a transitive relation $\ll$ (called *chronological* or *timelike* relation) contained in $\leq$.

The following notation for the timelike/causal future or past of a point is standard

$$\begin{aligned} I^+(p) &:= \{x \in X \mid p \ll x\}\,, & J^+(p) &:= \{x \in X \mid p \leq x\}\,, \\ I^-(p) &:= \{x \in X \mid x \ll p\}\,, & J^-(p) &:= \{x \in X \mid x \leq p\}\,, \\ I(p,q) &:= I^+(p) \cap I^-(q), & J(p,q) &:= J^+(p) \cap J^-(q), \end{aligned}$$

and we write $p < q$ if $p \leq q$ and $p \neq q$.

Definition 4.5 ([KS18, Definition 2.8])**.** A *Lorentzian pre-length space* is a causal set $(X, \ll, \leq)$ equipped with a metric d and a lower semicontinuous function $\tau\colon X \times X \to [0, \infty]$ satisfying, for all $x, y, z \in X$

(i) $\tau(x, y) > 0$ if $x \ll y$,

(ii) $\tau(x, y) = 0$ if $x \not\leq y$,

(iii) $\tau(x, z) \geq \tau(x, y) + \tau(y, z)$ if $x \leq y \leq z$.

Occasionally we denote a Lorentzian pre-length space simply by X. It follows from Definition 4.5 that the sets $I^\pm(p)$ are open for all $p \in X$, a fact that we will also refer to as the *openness of* $\ll$. The crucial push-up property extends from the smooth situation.

Lemma 4.6 (Push-up [KS18, Lemma 2.10])**.** *Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space and $x, y, z \in X$ with $x \ll y \leq z$ or $x \leq y \ll z$. Then $x \ll z$.*

The function τ in Definition 4.5 is often called *time separation function* or *Lorentzian distance function*. We will not use this terminology. In fact, we never need the function τ by itself but just the openness of $\ll$ and the push-up property (we could also trivially set $\tau(p, q) = \infty$ if $p \ll q$ and $= 0$ otherwise).

In smooth Lorentzian geometry the causal character of curves determines the timelike and causal future and past, that is $I^{\pm}$ and $J^{\pm}$, respectively. In Lorentzian pre-length spaces it is the other way round.

Definition 4.7 ([KS18, Definition 2.18]). Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space and I be any (open, half-open, or closed) interval in $\mathbb{R}$. A non-constant locally Lipschitz path $\gamma\colon I \to X$ is called a

(i) *future-directed causal curve* if $\gamma(s_1) \leq \gamma(s_2)$ for all $s_1 < s_2 \in I$.

(ii) *past-directed causal curve* if $\gamma(s_2) \leq \gamma(s_1)$ for all $s_1 < s_2 \in I$.

Future- and past-directed *timelike* curves are defined analogously by replacing $\leq$ with $\ll$.

Remark 4.8. By a result in metric geometry (see, for instance, [BBI01, Proposition 2.5.9]), we can parametrize any causal curve by d-arclength (the reference is for closed intervals only, but the proof is in fact valid for any interval). Recall that $\gamma\colon I \to X$ is parametrized by d-arclength iff

$$L^d(\gamma|_{[a,b]}) = b - a \text{ for all } [a, b] \subseteq I.$$

Since a curve that is parametrized by d-arclength is automatically 1-Lipschitz continuous, causal curves remain causal when parametrizing them by d-arclength. Hence, without loss of generality, we can assume that causal curves are parametrized in a way so that they are not locally constant, i.e., not constant on any open subinterval of $\mathbb{R}$.

Definition 4.9 ([KS18, Definition 3.1]). A Lorentzian pre-length space X is called *causally path-connected* if for every $p < q$ there exists a future-directed causal curve connecting p and q, and for every $p \ll q$ a future-directed timelike curve connecting p and q.

Definition 4.10 ([ACS20, Definition 2.19]). For a subset U of a Lorentzian pre-length space X we define the relation $\leq_U$ by

$$p \leq_U q :\Longleftrightarrow \text{there is a future-directed causal curve from } p \text{ to } q \text{ in } U.$$

A neighborhood U is called *weakly causally closed* if $\leq_U$ is closed, and the Lorentzian pre-length space X is called *locally weakly causally closed* if every point $p \in X$ is contained in a weakly causally closed neighborhood U.

Definition 4.9 is satisfied on *any* smooth Lorentzian manifold, and thus acts as a replacement for regularity on a Lorentzian pre-length space. In contrast, the "local causal closedness" condition of Kunzinger and Sämann [KS18, Definition 3.4] is stronger than Definition 4.10 because it requires that $\leq$ restricted to $U \times U$ is closed. For instance, in the smooth case the latter notion would only be satisfied on strongly causal spacetimes (see [ACS20, p. 6] for a detailed discussion).

Finally, we refine the concept of an inextendible curve[2].

[2] Doubly-inextendible causal curves are often simply called "inextendible". On the other hand, Kunzinger and Sämann [KS18] call a curve inextendible if it is either future- or past-inextendible (or both), and have no need for the concept of double-inextendibility. We will be more precise when needed.

Definition 4.11. Let X be a Lorentzian pre-length space and $\gamma\colon (a,b) \to X$ be a future-directed causal curve. If there exists a causal curve $\bar{\gamma}\colon (a,b] \to X$ such that $\bar{\gamma}|_{(a,b)} = \gamma$, we say that γ is *future-extendible.* If there exists a causal curve $\tilde{\gamma}\colon [a,b) \to X$ such that $\tilde{\gamma}|_{(a,b)} = \gamma$, we say that γ is *past-extendible.* We say that γ is *future-(past-)inextendible* if it is not future-(past-)extendible, and *doubly-inextendible* if it is neither future- nor past-extendible.

The analogous definition for past-directed causal curves is obtained by interchanging future and past in Definition 4.11. The definition applies accordingly to half-open intervals. If a path is defined on all of $\mathbb{R}$, we mean extendibility to $\pm\infty$. Alternatively we can parametrize it by arclength and apply the following lemma (which, of course, admits also a past version).

Lemma 4.12. *Let X be a locally weakly causally closed, causally path-connected Lorentzian pre-length space, let $-\infty < a < b \leq \infty$ and let $\gamma\colon [a,b) \to X$ be a future-directed causal curve parametrized with respect to d-arclength. If (X,d) is a proper metric space or the curve γ is contained in a compact set, then γ is future-inextendible if and only if $b = \infty$. In this case $L^d(\gamma) = \infty$. Moreover, γ is future-inextendible if and only if* $\lim_{t\nearrow b}\gamma(t)$ *does not exist.*

Proof. The proof is the same as [KS18, Lemma 3.12]. Note that there, the assumption of local "strong" causal closedness is only applied to points which lie on γ. Hence that proof also works with our notion of weakly causally closed neighborhood. ☐

In the remaining subsection we recall the definition of Lorentzian length space (including necessary preliminary notions) as introduced in [KS18]. While we will not directly work with Lorentzian length spaces in the main body of this paper, our results about pre-length spaces immediately also lead to useful Corollaries in this setting (see Introduction). A Lorentzian length space is, in essence, just the Lorentzian analogue of length metric spaces generalizing Riemannian manifolds where the time separation function τ is used in place of a distance function to measure lengths and the admissible class of curves respects causality. More precisely, if $\gamma\colon [a,b] \to X$ is a future-directed causal curve, then its *τ-length* $L_\tau(\gamma)$ is defined by (see [KS18, Definition 2.24])

$$L_\tau(\gamma) := \inf\left\{ \sum_{i=0}^{N-1} \tau(\gamma(t_i),\gamma(t_{i+1})) \;\middle|\; a = t_0 < t_1 < \ldots < t_N = b, N \in \mathbb{N} \right\}.$$

In addition, the notion of localizability is needed. In the main sections of this paper, only assumptions (i) and (ii) are needed, thus we restate them separately below (see Definitions 4.17 and 4.20, Lemma 4.22 and Proposition 4.23).

Definition 4.13 ([KS18, Definition 3.16] and [ACS20, Definition 2.22][3])**.** We call a Lorentzian pre-length space $(X,d,\ll,\leq,\tau)$ *localizable* if for every point $p \in X$, there exists a neighborhood U_p of p such that

[3]Similarly to the difference between weak and "strong" local causal closedness, there is a difference between the notions of localizability in [KS18, Definition 3.16] and [ACS20, Definition 2.22] regarding the meaning of $\ll_{U_p}, \leq_{U_p}$.

(i) There exists a constant $C > 0$ such that for all causal curves contained in U_p we have $L^d(\gamma) \leq C$.

(ii) For every $q \in U_p$ we have $I^{\pm}(q) \cap U_p \neq \emptyset$.

(iii) There exists a continuous function $\omega_p\colon X \times X \to [0, \infty)$ such that $(U_p, d|_{U_p \times U_p}, \ll_{U_p}, \leq_{U_p}, \omega_p)$ is a Lorentzian pre-length space. Moreover, for all $x, y \in U_p$ with $x < y$, it holds that

$$\omega_p(x, y) = \max\{L_\tau(\gamma) \mid \gamma\colon [a, b] \to U_p \text{ fd. causal from } x \text{ to } y\ \},$$

so in particular there exists a maximizing causal curve between x and y.

By further assuming that also τ is given by length-maximization (but without necessarily requiring the existence of global maximizers), one obtains a Lorentzian length space.

Definition 4.14 ([KS18, Definition 3.22])**.** A causally path-connected, locally (weakly) causally closed and localizable Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called a *Lorentzian length space* if for all $p, q \in X$

$$\tau(p, q) = \sup\{L_\tau(\gamma) \mid \gamma \text{ future-directed causal from } p \text{ to } q\}.$$

4.2.2 Limit curve theorems

We revisit the limit curves theorems of Kunzinger and Sämann [KS18, Section 3.2] and relax their assumption of local causal closedness to local weak causal closedness (see Definition 4.10). That this extension is possible was already pointed out by Aké et al. [ACS20, p. 8]. The limit curve theorems are crucial for Sections 4.3 and 4.5.

We start with [KS18, Lemma 3.6] where, instead of pointwise convergence, we need to assume locally uniform convergence.

Lemma 4.15. *Let X be a causally path-connected locally weakly causally closed Lorentzian pre-length space and let $(\gamma_n)_n$ be a sequence of future-directed causal curves $\gamma_n\colon I \to X$ converging locally uniformly to a non-constant Lipschitz curve $\gamma\colon I \to X$. Then γ is future-directed causal.*

Proof. For every $s \in I$ there exists a weakly causally closed neighborhood $U_{\gamma(s)}$. By continuity, we can pick $s_1 < s < s_2$ such that $\gamma([s_1, s_2]) \subseteq U_{\gamma(s)}$ (if s is a boundary point of I, then $s_1 = s$ or $s_2 = s$ is chosen). Assume additionally that we choose s_1, s_2 close enough such that $(\gamma_n)_n$ converges uniformly on $[s_1, s_2]$. This implies that there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$, $\gamma_n([s_1, s_2]) \subseteq U_{\gamma(s)}$. Now it follows from the definition of weakly causally closed neighborhood that γ restricted to $[s_1, s_2]$ is future-directed causal. Since s was arbitrary, we can decompose γ as a concatenation of future-directed causal curves, hence by transitivity of $\leq$, γ is future-directed causal on I. $\square$

Theorem 4.16 (Limit curve theorem)**.** *Let X be a causally path-connected locally weakly causally closed Lorentzian pre-length space. Let $(\gamma_n)_n$ be a sequence of future-directed causal curves $\gamma_n\colon [a,b] \to X$ that are uniformly Lipschitz continuous, i.e., there is an $L > 0$ such that $\mathrm{Lip}(\gamma_n) \leq L$ for all $n \in \mathbb{N}$. Suppose that there exists a compact set that contains every γ_n or that d is proper and that the curves $(\gamma_n)_n$ accumulate at some point, i.e., there is a $t_0 \in [a,b]$ such that $\gamma_n(t_0) \to x_0 \in X$. Then there exists a subsequence $(\gamma_{n_k})_k$ of $(\gamma_n)_n$ and a Lipschitz continuous curve $\gamma\colon [a,b] \to X$ such that $\gamma_{n_k} \to \gamma$ uniformly. If γ is non-constant, then γ is a future-directed causal curve.*

Proof. The proof of [KS18, Theorem 3.7] goes through. The assumption of local "strong" causal closedness is only used to invoke [KS18, Lemma 3.6], but since the convergence is uniform, we can replace it by our Lemma 4.15. $\square$

This first limit curve theorem is already very useful. In order to formulate our second limit curve theorem, we need a better control over the d-length of causal curves.

Definition 4.17 ([KS18, Definition 3.13])**.** A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *d-compatible* if for every $p \in X$ there exists a neighborhood U of p and a constant $C > 0$ such that $L^d(\gamma) \leq C$ for all causal curves γ contained in U.

To ease the nomenclature, we group some of our previous assumptions into the following definition.

Definition 4.18. A *Lorentzian pre-length space with limit curves* is a causally path-connected, locally weakly causally closed, and d-compatible Lorentzian pre-length space.

It is an easy consequence that Lorentzian length spaces are particular cases of Lorentzian pre-length spaces with limit curves (see also Proposition 4.23 below).

Theorem 4.19 (Limit curve theorem for inextendible curves)**.** *Let X be a Lorentzian pre-length space with limit curves. Let $(\gamma_n)_n$ be a sequence of future-directed causal curves $\gamma_n\colon [0, L_n] \to X$ which are parametrized with respect to d-arclength and satisfy $L_n := L^d(\gamma_n) \to \infty$. If there exists a compact set that contains every curve $\gamma_n([0, L_n])$ or if d is proper and $\gamma_n(0) \to x$ for some $x \in X$, then there exists a subsequence $(\gamma_{n_k})_k$ of $(\gamma_n)_n$ and a future-directed causal curve $\gamma\colon [0,\infty) \to X$ such that $\gamma_{n_k} \to \gamma$ locally uniformly. Moreover, γ is future-inextendible.*

Proof. The proof of [KS18, Theorem 3.14] goes through. The assumption of local "strong" causal closedness is only used to invoke [KS18, Lemmas 3.6 and 3.12, Theorem 3.7], so we can replace them by our Lemma 4.15, Lemma 4.12 and Theorem 4.16 respectively. $\square$

While the limit curve theorems are stated for future-directed curves, they of course also hold for past-directed ones.

4.2.3 Approximating Lorentzian pre-length spaces

In this subsection we introduce our new “approximating” condition relating the causal structure and the topology on X. It is satisfied on all spacetimes regardless of their place in the causal ladder, and will be crucial in Section 4.4. In Proposition 4.23 we show that all Lorentzian length spaces automatically fulfill the “approximating” condition, and also our earlier Definition 4.18.

Definition 4.20. A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *approximating* if for all points $p \in X$ it holds that $J^{\pm}(p) \subseteq \overline{I^{\pm}(p)}$.

It is called *future-(past-)approximating* if the approximating property holds for $+(-)$.

The approximating property can equivalently be characterized via sequences as follows.

Lemma 4.21. *Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. Then X is (future-/past-)approximating if and only if for every point $p \in X$ there exists a sequence $(p_n^{\pm})_n$ in $I^{\pm}(p)$ such that $p_n^{\pm} \to p$ as $n \to \infty$.*

We say that the sequence $(p_n^+)_n$ approximates p from the future, and that the sequence $(p_n^-)_n$ approximates p from the past.

Proof. That such sequences exist on approximating spaces is obvious, because $p \in J^{\pm}(p) \subseteq \overline{I^{\pm}(p)}$. To show the converse, suppose $q \in J^+(p)$ and (q_n^+) is a sequence in $I^+(q)$ approximating q from the future. By the push-up Lemma 4.6, $q_n^+ \in I^+(p)$ for all $n \in \mathbb{N}$, hence $q \in \overline{I^+(p)}$. □

Assuming that X is causally path-connected, we get even more characterizations, which in the smooth case are in fact the most widely used ones.

Lemma 4.22. *Let $(X, d, \ll, \leq, \tau)$ be a causally path-connected Lorentzian pre-length space. Then the following are equivalent:*

(i) X is future-(past-)approximating,

(ii) $I^+(p) \neq \emptyset$ ($I^-(p) \neq \emptyset$) for all $p \in X$,

(iii) for every point $p \in X$ is there exists a future-(past-)directed timelike curve $\gamma \colon [a, b) \to X$ with $\gamma(a) = p$.

Note that if X is approximating, we can always join the future- and past-directed curves from point (iii) to find a timelike curve $\gamma \colon (a, b) \to X$ through p.

Proof. (i) $\implies$ (ii) Let $p \in X$ be any point. Then $p \in J^+(p)$, so if X is future-approximating, we get that $\emptyset \neq J^+(p) \subseteq \overline{I^+(p)}$. This implies that $I^+(p) \neq \emptyset$.

(ii) $\implies$ (iii) By assumption, there exists points $q \in I^+(p)$. By causal path-connectedness, there exists a future-directed timelike curve γ from p to q, which by Definition 4.7 must be non-constant, even if $p = q$. We can then remove the appropriate endpoint of γ to get the desired curve.

(iii) $\implies$ (i) Let $p \in X$ and $\gamma\colon [a,b) \to X$ be a future-directed timelike curve with $\gamma(a) = p$. By continuity of γ, we have $p = \lim_{s \to a} \gamma(s)$, so $p \in \overline{I^+(p)}$.

The past statements are proved analogously. $\square$

We can now easily see how Lorentzian length spaces are particular cases of the more general pre-length spaces that we will be working with in the rest of the paper.

Proposition 4.23. *If X is a localizable, causally path-connected Lorentzian pre-length space, then X is approximating and d-compatible. If X is a Lorentzian length space, then X is an approximating Lorentzian pre-length space with limit curves.*

Proof. If X is localizable, then by property (i) in Definition 4.13, X is d-compatible. Furthermore, by property (ii), every point $q \in X$ has $I^\pm(q) \neq \emptyset$, and then by Lemma 4.22, X is approximating. The second statement follows trivially from the definitions. $\square$

Remark 4.24. In connection with the null distance on Lorentzian pre-length spaces, Kunzinger and Steinbauer [KS22, Definition 3.4] introduced the notion of *sufficiently causally connectedness (scc)*. A Lorentzian pre-length space is scc if it is path-connected (in the sense of metric spaces), causally path-connected (Definition 4.9) and every point $p \in X$ lies on some timelike curve γ. While the last condition is reminiscent of property (iii) in our Lemma 4.22, it is in fact weaker, since scc puts no restriction on whether p should be an endpoint of γ. On the other hand, we do not need to assume path-connectedness.

Having established the existence of causal (even timelike) curves through every point in Lemma 4.22, the question remains whether one can find a (doubly-)inextendible causal curve through every point (see Definition 4.11). The following proposition and corollary answer this question in the affirmative, which will be crucial in Section 4.5 when studying Cauchy sets. We need to assume that (X,d) is proper in order to invoke the limit curve theorem. The use of the latter is also the reason why we only prove the existence of intextendible causal (and not timelike) curves.

Proposition 4.25 (Existence of maximal extensions of causal curves)**.** *Let X be an approximating Lorentzian pre-length space with limit curves. Suppose, in addition, that (X,d) is proper. Then, for every future-(past-)directed causal curve $\gamma\colon [a,b) \to X$ with $b < \infty$, there exists $c \in (b,\infty]$ and a future-(past-)inextendible causal curve $\lambda : [a,c) \to X$ such that $\lambda|_{[a,b)} = \gamma$.*

Proof. Consider, without loss of generality, the case that γ is future-directed. If γ is already inextendible, there is nothing to prove. Hence we consider the case of γ being extendible. Then γ has an endpoint, which we will, by abuse of notation, denote as $\gamma(b)$. Since X is approximating, by Lemma 4.22 there is a future-directed timelike curve starting at $\gamma(b)$. Concatenating it with γ, we get a proper extension $\tilde\gamma\colon [a,c) \to X$ of γ, where $c > b$. If we can choose $\tilde\gamma$ to be inextendible, we are done. Hence, suppose for the sake of contradiction that all extensions of γ are themselves extendible. There are two possible cases:

1. There exists a constant $C > 0$ such that the d-arclength of all extensions of γ is bounded by C. Suppose that we have chosen C as small as possible. Then there exists a sequence $(\gamma_n)_n$ of extensions such that $L^d(\gamma_n) \to C$. Since all the γ_n are extendible (hence we can add their future-endpoints) and agree at the point $\gamma(b)$, by Theorem 4.16 a subsequence converges to a limit curve $\gamma_\infty \colon [a, c] \to X$ of arclength $L^d(\gamma_\infty) = C$. But then, by the above, γ_∞ admits a future extension, which is then also an extension of γ and has arclength greater than C, a contradiction.

2. There exists a sequence $(\gamma_n)_n$ of extensions of γ such that $L^d(\gamma_n) \to \infty$. In this case we can apply Theorem 4.19 to find an inextendible limit curve γ_∞ of a subsequence. This γ_∞ is then the desired inextendible extension of γ. □

Combining Lemma 4.22 and Proposition 4.25 gives us an important conclusion.

Corollary 4.26. *Let X satisfy the assumptions of Proposition 4.25. Then, for every point in $p \in X$, there exists a doubly-inextendible causal curve passing through p.* □

4.2.4 Causality conditions and time functions

The conditions in the previous subsections relating the topology and the causal structure (such as approximating) are satisfied automatically when the topology is that of a manifold, and the causal structure is induced by a Lorentzian metric. They can thus be thought of as making our Lorentzian pre-length spaces more "manifold-like", while still being much more general. In this subsection, on the other hand, we are going to discuss causality conditions, i.e., steps on the causal ladder, which are not satisfied by all smooth spacetimes and hence also not by all Lorentzian pre-length spaces. They should be thought of as criteria for physical reasonability.

In this section, we consider the notions of causality and non-total imprisonment, and the definition of time functions (for an in-depth treatment of the causal ladder for Lorentzian length spaces, see [ACS20]). Most of the material is standard, but Theorem 4.30 and Proposition 4.31 are new. The goal of this paper is to characterize the existence of (certain kinds of) time functions by suitable causality conditions, which will be introduced in the main sections. The causality conditions in this section are weaker, but also play an important role.

A smooth spacetime is called causal if it contains no closed causal curves. The following equivalent definition is better suited for Lorentzian pre-length spaces.

Definition 4.27 ([KS18, Definition 2.35]). A Lorentzian pre-length space is called *causal* if for any two points $p, q \in X$, $p < q$ implies $q \not< p$.

Time functions too, can be defined either via causal curves (time functions are then required to be strictly increasing on future-directed causal curves), or in the following, more order-theoretic manner.

Definition 4.28. A function $f\colon X \to \mathbb{R}$ on a Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called a *generalized time function* if for all $p, q \in X$,

$$p < q \implies f(p) < f(q).$$

It is called a *time function* if it is also continuous.

Clearly, the existence of a (generalized) time function requires that the underlying space is at least causal.

In the smooth case, non-total imprisonment is equivalent to the following definition (see, for instance, [Min18, Theorem 4.39]).

Definition 4.29 ([KS18, Definition 2.35])**.** A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *non-totally imprisoning* if for every compact set $K \subseteq X$ there exists a constant $C > 0$ such that for every causal curve γ with image in K, $L^d(\gamma) \leq C$.

As a corollary to the limit curve theorems, we obtain the following alternative characterizations.

Theorem 4.30. *Let X be a Lorentzian pre-length space with limit curves. Then the following are equivalent.*

(i) X is non-totally imprisoning.

(ii) No compact set in X contains a future-inextendible causal curve.

(iii) No compact set in X contains a past-inextendible causal curve.

(iv) No compact set in X contains a doubly-inextendible causal curve.

Proof. The equivalence between (i), (ii) and (iii) is shown in [KS18, Corollary 3.15]. As a consequence of Lemma 4.12, any doubly-inextendible curve has infinite arclength. Thus (i) implies (iv).

It remains to be shown that (iv) implies (i). Suppose X is totally imprisoning. Then there exists a compact set K and a sequence of future-directed causal curves $\gamma_n\colon [0, L_n] \to X$, parametrized by arclength and contained in K, such that $L_n = L^d(\gamma_n) \to \infty$. Consider the sequence of future-directed causal curves $\bar{\gamma}_n\colon [0, L_n/2] \to X$ given by

$$\bar{\gamma}_n(s) := \gamma_n\left(\frac{L_n}{2} + s\right).$$

Then also $L^d(\bar{\gamma}_n) = L_n/2 \to \infty$ and we can apply Theorem 4.19 to find a converging subsequence $(\bar{\gamma}_{n_k})_k$ and a future-inextendible causal limit curve $\bar{\gamma}\colon [0, \infty) \to X$. Next consider the sequence of past-directed causal curves $\tilde{\gamma}_k\colon [-L_n/2, 0] \to X$ given by

$$\tilde{\gamma}_k(s) := \gamma_{n_k}\left(\frac{L_n}{2} + s\right).$$

Again we can apply Theorem 4.19 to find a converging subsequence $(\tilde{\gamma}_{k_m})_m$ and a past-inextendible limit curve $\tilde{\gamma}\colon (-\infty, 0] \to X$. Note that

$$\tilde{\gamma}(0) = \lim_{m\to\infty} \gamma_{n_{k_m}}(L_n/2) = \bar{\gamma}(0),$$

where the limit in the middle exists by compactness of K, a fact that we had already used implicitly when applying the Limit Curve Theorem 4.19. Thus the curve $\tilde{\gamma}$ joined with $\bar{\gamma}$ is a doubly-inextendible causal curve $(-\infty,\infty) \to X$ contained in K. This contradicts our assumption (iv). □

The following result was shown by Kunzinger and Sämann for Lorentzian length spaces, but only the assumption of causal path-connectedness is used in the proof.

Proposition 4.31 ([KS18, Theorem 3.26]). *Suppose $(X, d, \ll, \leq, \tau)$ is a causally path-connected Lorentzian pre-length space. If X is non-totally imprisoning, then X is causal.*

For spacetimes, it is well-known that the existence of a time function implies strong causality, and that strong causality implies non-total imprisonment. This result has been shown by Kunzinger and Steinbauer for Lorentzian length spaces [KS22, Theorem 3.13], under the additional assumption that the time function must be topologically locally anti-Lipschitz. We instead give a direct proof of the fact that for Lorentzian pre-length spaces with limit curves, the existence of any kind of time function implies non-total imprisonment.

Lemma 4.32. *Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space with limit curves. If X admits a time function, then X is non-totally imprisoning.*

Proof. Suppose X is totally imprisoning. Then by Theorem 4.30 there exists a compact set $K \subseteq X$ and a future-inextendible future-directed causal curve $\gamma\colon [0,\infty) \to K$. Note the following two facts:

(i) By Lemma 4.12, because γ is inextendible, $\lim_{s\to\infty} \gamma(s)$ does not exist.

(ii) By compactness of K, for every sequence $(s_i)_i$ in $[0,\infty)$, there exists a subsequence of $(\gamma(s_i))_i$ that converges in K.

Thus we can find two sequences $(r_i)_i$ and $(s_i)_i$ in $[0,\infty)$ such that $p := \lim_{i\to\infty} \gamma(r_i) \neq \lim_{i\to\infty} \gamma(s_i) =: q$ (in particular, both limits exist).

For these p, q, pick $\delta > 0$ small enough so that $B_\delta(p)$ is contained in a weakly causally closed neighborhood and $q \notin B_\delta(p)$. Assume w.l.o.g. that $r_1 < s_1 < r_2 < s_2 \dots$ and that for all $i \in \mathbb{N}$ we have $\gamma(r_i) \in B_{\delta/2}(p)$ and $\gamma(s_i) \notin B_{\delta/2}(p)$. Now define a third sequence $(a_i)_i$ with $r_i < a_i < s_i$ and such that a_i is the value at which $\gamma|_{[r_i,s_i]}$ first intersects $\partial B_{\delta/2}(p)$. We then have $r_i < a_i < s_i < r_{i+1}$. By compactness of $\partial B_{\delta/2}(p)$, there exists a subsequence of $(\gamma(a_i))_i$ that converges to a point $q' \neq p$. Since $r_i < a_i$, we have $\gamma(r_i) \leq \gamma(a_i)$. Because $B_\delta(p)$ is contained in a weakly causally closed neighborhood, we have $p < q'$. By assumption, X admits a time function $t\colon X \to \mathbb{R}$, for which it holds that $t(p) < t(q')$. On the other hand, since $a_i < r_{i+1}$, we have $\gamma(a_i) \leq \gamma(r_{i+1})$. Thus $t(\gamma(a_i)) \leq t(\gamma(r_{i+1}))$ for all $i \in \mathbb{N}$

and by continuity of t and γ also $t(q') \leq t(p)$. Combining this with the previous inequality, we obtain

$$t(p) < t(q') \leq t(p),$$

which is a contradiction. $\square$

4.3 Time functions and K-causality

The notion of K-causality was first introduced by Sorkin and Woolgar [SW96] to study spacetimes with continuous Lorentzian metrics. Among the multiple applications of this concept, we emphasize the work of Minguzzi [Min10], who showed that for smooth spacetimes, K-causality is equivalent to stable causality. Since Hawking [Haw68] had shown earlier that stable causality is equivalent to the existence of a time function, so is K-causality. Minguzzi in [Min10] also gave a direct proof of the equivalence between K-causality and the existence of time functions, which is more mathematically rigorous and less dependent on the Lorentzian manifold structure. Since K-causality is a purely order-theoretical notion, it can be used verbatim[4] for Lorentzian pre-length spaces.

Definition 4.33 ([SW96, Definitions 8 and 9])**.** Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. The K^+*-relation* on X is defined as the (unique) smallest transitive relation that contains $\leq$ and is (topologically) closed.

A Lorentzian pre-length space X is called K*-causal*[5] if the K^+-relation is antisymmetric.

In this section we establish the equivalence of the existence of time functions and K-causality on certain Lorentzian pre-length spaces (see Theorem 4.34 below and Theorem 4.1 formulated for Lorentzian length spaces). This result generalizes the analogous theorem known for smooth spacetimes by Minguzzi [Min10, Theorem 7] and the proof is obtained along the same lines.

Theorem 4.34. *Suppose τ is a second countable, locally compact Lorentzian pre-length space with limit curves. Then X is K-causal if and only if X admits a time function.*

Before proving the theorem, we show that time functions can exist more generally also on Lorentzian pre-length spaces that are not K-causal if they do not satisfy the limit curve property (Definition 4.18).

[4] In [SW96] the K^+-relation is only required to contain I^+. Our definition with J^+ can be traced back to [ACS20; Min10]. On spacetimes and approximating Lorentzian pre-length spaces, we have $J^+ \subseteq \overline{I^+}$, hence there it makes no difference.

[5] In [ACS20; KS18] a Lorentzian pre-length space with this property is called *stably causal* because in the smooth case K-causality and stable causality are equivalent [Min09]. The definition of stable causality [Haw68, p. 433], however, requires knowledge about causal properties of "nearby" Lorentzian metrics. Since (causal) stability of Lorentzian pre-length spaces has not yet been investigated in this sense, we prefer to use the standard term K-causal.

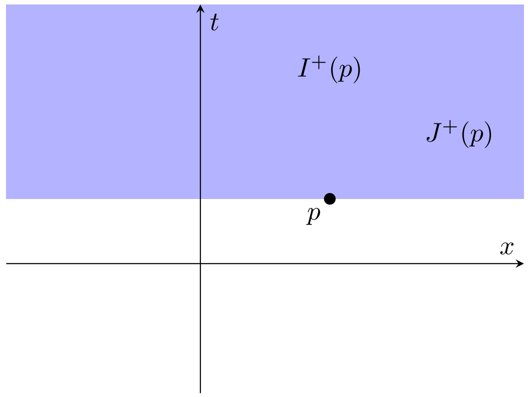


Figure 4.1: The sets $I^+(p)$ (blue, without boundary) and $J^+(p) = I^+(p) \cup \{p\}$ for a point p in Example 4.35.

Example 4.35 (There exist Lorentzian pre-length spaces that admit a time function but are neither strongly causal nor K-causal)**.** Let (X, d) be the Euclidean plane with coordinates (t, x). For any pair of points $p_i = (t_i, x_i)$, $i = 1, 2$, let

$$\begin{aligned} p_1 \ll p_2 &:\iff t_1 < t_2, \\ p_1 \leq p_2 &:\iff t_1 < t_2 \text{ or } p_1 = p_2, \end{aligned}$$

as depicted in Figure 4.1, and

$$\tau(p_1, p_2) = t_2 - t_1.$$

This equips (X, d) with the structure of a Lorentzian pre-length space. Clearly, the t-coordinate is a time function.

However, this space is not K-causal, as any closed relation containing $\leq$ (or $\ll$) must contain the relation $\leq_R$ given by

$$p_1 \leq_R p_2 \iff t_1 \leq t_2.$$

But $\leq_R$ is not antisymmetric, so the K-relation will not be antisymmetric either, and our space is therefore not K-causal.

Note that while K-causality implies strong causality on smooth spacetimes and locally compact Lorentzian length spaces [ACS20, Proposition 3.16] this need not be the case for Lorentzian pre-length spaces. To see that X is also not strongly causal, recall that a Lorentzian pre-length space X is called *strongly causal* if the Alexandrov topology, generated by

$$I(p, q) = \{r \in X \mid p \ll r \ll q\}, \qquad p, q \in X,$$

agrees with the metric topology [KS18, Definitions 2.4 and 2.35].

The Alexandrov topology of the above example is generated by the open horizontal stripes, hence is strictly coarser than the Euclidean topology and X therefore not strongly causal.

4.3.1 Proof of Theorem 4.34

We follow Minguzzi's proof for smooth spacetimes [Min10]. A key element is the following theorem from utility theory, a branch of mathematical economics with resemblances to causality theory.

Theorem 4.36 (Levin's Theorem [Lev83])**.** *Let X be a second countable, locally compact Hausdorff topological space and R be a closed preorder on X. Then there exists a continuous function $f : X \to \mathbb{R}$ such that*

$$(x, y) \in R \implies f(x) \leq f(y),$$

with equality if and only if $x = y$.

That K-causality implies the existence of a time function is a direct consequence of Levin's Theorem, and is in fact true in an even more general setting than ours, as already pointed out by Minguzzi [Min13].

It remains to show the converse. The next lemma gives us a more explicit characterization of the K-relation. Its proof is not significantly different to its smooth counterpart [Min10, Lemma 3] but included for the sake of clarity.

Lemma 4.37. *Let $(X, d, \ll, \leq, \tau)$ be a non-totally imprisoning, locally compact Lorentzian pre-length space with limit curves. If $(p, q) \in K^+ \subseteq X \times X$, then either $p \leq q$ or for every relatively compact open set B containing p, there exists $r \in \partial B$ such that $p < r$ and $(r, q) \in K^+$.*

Proof. For the purposes of this proof, it will be more convenient to denote relations as subsets of $X \times X$; in particular, $J^+ := \{(p, q) \in X \times X \mid p \leq q\}$. Consider the relation

$$\begin{aligned} R^+ := \{(p, q) \in K^+ \mid (p, q) \in J^+ &\text{ or for every relatively compact} \\ &\text{open set } B \text{ containing } p \text{ there is an } r \in \partial B \\ &\text{such that } p < r \text{ and } (r, q) \in K^+\}. \end{aligned}$$

Clearly, $J^+ \subseteq R^+ \subseteq K^+$. We will show that R^+ is closed and transitive, which then implies $R^+ = K^+$, and in turn proves the Lemma. Transitivity can be proven exactly in the same way as in the smooth case, so we refer to [Min10, Lemma 3].

To show closedness or R^+, consider $(p_n, q_n) \to (p, q)$ with $(p_n, q_n) \in R^+$ for all $n \in \mathbb{N}$. If $p = q$ then $(p, q) \in J^+ \subseteq R^+$. We can therefore assume that $p \neq q$ and it remains to be shown that $(p, q) \in R^+$ as well. Let B be an open relatively compact neighborhood of p. For sufficiently large n, $p_n \neq q_n$ and $p_n \in B$. By passing to a subsequence if necessary, we can assume that either $(p_n, q_n) \in J^+$ (case 1) or $(p_n, q_n) \in R^+ \setminus J^+$ (case 2) for all $n \in \mathbb{N}$:

1. Suppose $(p_n, q_n) \in J^+$ for all $n \in \mathbb{N}$. By assumption X is causally path-connected (and $p_n \neq q_n$), thus there exist future-directed causal curves $\gamma_n \colon [0, 1] \to X$ from p_n to q_n. By passing to a subsequence if necessary, we can assume that γ_n either lies entirely in $\overline{B}$ for all n, or leaves $\overline{B}$ for all n.

If all the γ_n lie inside $\overline{B}$, then by non-total imprisonment their lengths (and by linear reparametrization also their Lipschitz constants) are bounded above by a positive constant C independent of n. Thus we are in conditions to apply Theorem 4.16, which shows the existence of a limit causal curve connecting p and q, and hence $(p,q) \in J^+ \subseteq R^+$.

If, on the other hand, none of the γ_n lie entirely inside $\overline{B}$, then there is a first parameter value s_n at which γ_n leaves B (and $\gamma_n(s_n) \in \partial B$ by connectedness), and define a new sequence of curves $\tilde{\gamma}_n := \gamma_n|_{[0,s_n]}$. Linear reparametrization so that $\tilde{\gamma}_n \colon [0,1] \to X$ together with Theorem 4.16 shows the existence of a limit causal curve that connects p with a point $r = \lim_{n\to\infty} \gamma_n(s_n) \in \partial B$ (again w.l.o.g. by passing to a subsequence). Because $\gamma_n(s_n) \leq q$ for all n it follows that $(r,q) \in \overline{J^+} \subseteq K^+$ by closedness of the K-relation. Since also $p < r$, we again conclude that $(p,q) \in R^+$.

2. Suppose $(p_n, q_n) \in R^+ \setminus J^+$ for all n. Then there exist points $r_n \in \partial B$ such that $p_n < r_n$ and $(r_n, q_n) \in K^+$. By passing to a subsequence if necessary, we can assume that $r_n \to r \in \partial B$. Arguing as in case 1 (for the sequence (p_n, r_n)), either $p < r$, or there exists $r' \in \partial B$ such that $p < r'$ and $(r', r) \in \overline{J^+} \subseteq K^+$. Combining this with the fact that $(r,q) \in K^+$ by closedness of the K-relation, it follows that $(p,q) \in R^+$.

In both cases we have thus shown that $(p,q) \in R^+$, which concludes the proof. □

The next and final lemma is key, and tells us that time functions are K-utilities, in the language of economics. In other words, a time function with respect to the causal relation $\leq$ is automatically also a time function with respect to the K^+-relation.

Lemma 4.38. *Let $(X, d, \ll, \leq, \tau)$ be a locally compact Lorentzian pre-length space with limit curves and $t \colon X \to \mathbb{R}$ be a time function. If $(p,q) \in K^+$, then either $p = q$ or $t(p) < t(q)$.*

Proof. The proof is the same as the one of [Min10, Lemma 4], replacing [Min10, Lemma 2] and [Min10, Lemma 3] by our Lemmas 4.32 and 4.37, respectively. □

Proof of Theorem 4.34. That K-causal spaces admit time functions is a direct consequence of Levin's Theorem 4.36. Conversely, if X admits a a time function, then the K^+-relation must be antisymmetric, as otherwise it would contradict Lemma 4.38. □

Remark 4.39. It is worth pointing out that, throughout this section, we have not made use of the chronological relation $\ll$, nor of timelike curves. Hence, Theorem 4.34 is still valid if in Definition 4.9 we only require the existence of causal (and not of timelike) curves, or even if $\ll$ is empty.

4.4 Volume time functions

In this section we introduce and explicitly construct special types of functions, called *averaged volume functions*, on Lorentzian pre-length spaces that are equipped with probability measures. While the existence of a suitable measure solely depends on the topology (in fact, the metric structure) it is the causal structure that determines whether these functions are time functions. More precisely, we will see that the averaged volume functions are time functions if and only if the underlying Lorentzian pre-length space is causally continuous in Theorem 4.52 (see Theorem 4.2 for the Lorentzian length space version thereof).

Our results generalize a classical theorem of Dieckmann [Die87; Die88b] (also stated earlier by Hawking and Sachs [HS74], but with an incomplete proof). Volume functions had already been introduced earlier by Geroch [Ger70, Sec. 5] to study global hyperbolicity; we will replicate those results in Section 4.5. In this section, we follow the approach of Dieckmann, but also make use of an averaging procedure similar to that used by Hawking [Haw68] to study stable causality and time functions. Besides that, our methods in this section are based on order- and measure-theoretical arguments, and we do not need to assume the existence of causal curves.

4.4.1 Averaged volume functions

To construct averaged volume functions on a Lorentzian pre-length space X, we equip X with a Borel measure μ satisfying

(i) $\mu(X) = 1$, i.e., μ is a probability measure, and

(ii) $\operatorname{supp}(\mu) = X$, i.e., μ has full support.

When X is finite the construction of μ is trivial. Otherwise such a measure exists precisely when (X, d) is a separable metric space (which is automatically satisfied for all compact spaces).

Proposition 4.40. *A metric space (X, d) admits a Borel probability measure μ with $\operatorname{supp}(\mu) = X$ if and only if it is separable (equivalently, second-countable).*

Proof. Suppose (X, d) is separable. Then it contains a countable dense subset $D = \{p_n \mid n \in \mathbb{N}\}$. Denote by δ_n the Dirac delta measure centered at p_n, and define

$$\mu := \sum_{n\in\mathbb{N}} 2^{-n}\delta_n.$$

The measure μ has the desired properties since (i) $\mu(X) = \sum_n 2^{-n} = 1$ and (ii) for all open sets A, $A \cap D \neq \emptyset$ by denseness and hence $\mu(A) > 0$. The proof of the converse can be found in [MS48, p. 134].

Finally, secound countability implies separability, and on metric spaces the two notions in fact are equivalent. $\square$

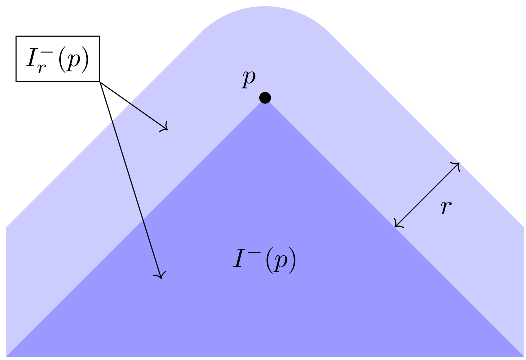


Figure 4.2: The sets $I^-(p)$ (dark blue) and $I_r^-(p)$ (light and dark blue) for some point p in Minkowski spacetime, with d the Euclidean distance.

Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space equipped with a Borel probability measure μ of full support. For $r \in (0,1)$ and $p \in X$, let

$$\begin{aligned} I_r^\pm(p) &:= \left\{x \in X \mid d\left(x, I^\pm(p)\right) < r\right\}, \\ V_r^\pm(p) &:= \mu\left(I_r^\pm(p)\right). \end{aligned}$$

We call $I_r^\pm(p)$ the r-thickening of $I^\pm(p)$, as depicted in Figure 4.2, and $V_r^\pm(p)$ its volume.

Definition 4.41. Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space equipped with a Borel probability measure μ of full support. The *future* $(+)$ and *past* $(-)$ *averaged volume functions* of μ are defined by

$$t^\pm(p) := \mp \int_0^1 V_r^\pm(p)\, dr, \qquad p \in X.$$

Note that the integral exists for all points $p \in X$ because the function $r \mapsto V_r^\pm(p)$ is increasing and bounded.

Remark 4.42 (Comparison to previous definitions of volume functions)**.** The classical definition of a volume function by Geroch [Ger70, Section 5] is simpler and reads $t_{\mathrm{cl}}^\pm(p) = \mp\mu\left(I^\pm(p)\right)$. However, it was discovered by Dieckmann [Die88b, Def. 1.2] that in order to show continuity, one has to require that μ also satisfies the property (iii) $\mu(\partial I^\pm(p)) = 0$. On a smooth spacetime (M, g) one can always construct such an admissible measure μ from the volume form using a partition of unity and utilize that $\partial I^\pm(p)$ is a hypersurface having zero Lebesgue measure in charts [Die88b, Prop. 1.1]. Since we do not have a manifold structure and the Lebesgue measure at our disposal, we instead integrate over r to "average out" discontinuities, hence the addition of "averaged" in the naming of volume functions in Definition 4.41. This averaging procedure is inspired by the work of Hawking [Haw68] on stable causality and time functions.

Remark 4.43 (μ-dependence)**.** It is clear that the above constructions of $I_r^\pm$, $V_r^\pm$ and $t^\pm$ depend crucially on the choice of d and μ. We will, however, see that the *existence* of (generalized) time functions is at this point independent of the particular choice of μ and also of d (as long as the metric is second countable). More precisely, whether the averaged volume functions $t^\pm$ are indeed (continuous) time functions depends only on the causal structure of X.

We end this subsection by proving that $t^\pm$ are *isotone* or *causal* functions (see [Min18, Def. 1.17]), a property that is weaker than being a time function.

Lemma 4.44. *Let X be a Lorentzian pre-length space as in Definition 4.41 and let $p, q \in X$. If $p \leq q$, then $V_r^-(p) \leq V_r^-(q)$ and $V_r^+(p) \geq V_r^+(q)$. In particular,*

$$p \leq q \Longrightarrow t^\pm(p) \leq t^\pm(q).$$

Proof. By the push-up Lemma 4.6, $p \leq q$ implies $I^-(p) \subseteq I^-(q)$, and hence $I_r^-(p) \subseteq I_r^-(q)$ for all $r \in (0,1)$. The first conclusion thus follows from the monotonicity of μ, and the second one from the monoticity of the integral. □

4.4.2 Averaged volume functions as generalized time functions

Finally, in order to show that $t^\pm$ are generalized time functions we need to apply the standard distinguishing causality condition, or at least our own weaker version thereof.

Definition 4.45 ([KP67, p. 486])**.** Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. We say that

(i) X is *past-distinguishing* if

$$I^-(p) = I^-(q) \Longrightarrow p = q, \qquad p, q \in X,$$

(ii) X is *future-distinguishing* if

$$I^+(p) = I^+(q) \Longrightarrow p = q, \qquad p, q \in X.$$

We call X *distinguishing* if it is both past- and future-distinguishing.

Definition 4.46. We say that X is *causally* (*past*- or *future*-) *distinguishing* if the conditions of Definition 4.45 are only required to hold for all $p, q \in X$ with $p \leq q$.

Furthermore, we assume that the Lorentzian pre-length spaces are approximating (see Section 4.2.3). This avoids the pathological situation where the future or past of a point could be "far away" from the point itself, or empty.

Having equipped our spaces with sufficient causal and topological conditions, we are in a position to establish the well-known classical result about generalized time functions [Die88b, Prop. 2.2] also for separable Lorentzian pre-length spaces (for which averaged volume functions $t^\pm$ from Definition 4.41 are well-defined by Proposition 4.40).

Proposition 4.47. *Let $(X, d, \ll, \leq, \tau)$ be a past-(future-)approximating Lorentzian pre-length space equipped with a Borel probability measure of full support. Then X is causally past-(future-)distinguishing if and only if t^- (t^+) is a generalized time function i.e., for all $p, q \in X$*

$$p < q \Longrightarrow t^{\mp}(p) < t^{\mp}(q).$$

Proof. We prove the past version. If X is causally past-distinguishing, then for all $p, q \in X$ with $p < q$,

$$I^-(p) \subsetneq I^-(q).$$

Clearly, $q \notin I^-(p)$, and we show that also $q \notin \partial I^-(p)$: For any $x \in I^-(q)$ the future $I^+(x)$ is open and thus contains an open set V around q. If $q \in \partial I^-(p)$ then there is a point $y \in V \cap I^-(p)$, thus $x \ll y \ll p$ and by transitivity $x \in I^-(p)$. Therefore $I^-(q) \subseteq I^-(p)$, a contradiction.

Since any metric space is normal, the disjoint closed sets $\{q\}$ and $\overline{I^-(p)}$ can be separated by disjoint open sets. In particular, there exists an open neighborhood U around q such that $r_0 := d(U, I^-(p)) > 0$. The intersection $U \cap I^-(q)$ is open, and by the past-approximating property of X at q (see Lemma 4.21) also non-empty. Thus

$$\begin{aligned} t^-(q) - t^-(p) &= \int_0^1 \mu\left(I_r^-(q) \setminus I_r^-(p)\right) dr \\ &\geq \int_0^{r_0} \mu\left(U \cap I^-(q)\right) dr > 0. \end{aligned}$$

To prove the converse, assume that for all $p < q$ we have $t^-(q) - t^-(p) > 0$. In order to show that X is causally past-distinguishing furthermore assume that $I^-(p) = I^-(q)$. But then, clearly $I_r^-(p) = I_r^-(q)$ for all $r \in [0, 1]$, and therefore $t^-(p) = t^-(q)$, a contradiction. Hence $p = q$. □

Example 4.35 is a Lorentzian pre-length space that is causally distinguishing but not distinguishing (any two points on the same level set of t have the same past, but are not causally related to one another). One can show that under certain conditions the two notions agree, in particular, for smooth spacetimes.

Proposition 4.48. *Let $(X, d, \ll, \leq, \tau)$ be a locally compact, causally path-connected, locally weakly causally closed, (past-/future-)approximating Lorentzian pre-length space. If X is causally (past-/future-)distinguishing, then it is also (past-/future-)distinguishing.*

Proof. Suppose X is causally past-distinguishing but not past-distinguishing (the future case is analogous). Then there exist two distinct points $p, q \in X$ such that $I^-(p) = I^-(q)$. By Lemma 4.21 there is a sequence $(p_n^-)_n$ that approximates p from the past. Because $p_n^- \in I^-(p) = I^-(q)$ and X is causally path-connected, there exists a sequence of causal curves $(\gamma_n)_n$ connecting p_n^- and q. Let δ be small enough so that $B_\delta(p)$ is compact and does not contain q. Consider the sequence of points $(r_n)_n$ where γ_n first intersects $\partial B_{\delta/2}(p)$. By compactness, we may assume it converges to a point $r \in \partial B_{\delta/2}(p)$. Without loss of generality suppose that $B_\delta(p)$ is contained in a weakly causally closed neighborhood, and so furthermore $p \leq r$ and $I^-(p) \subseteq I^-(r)$ by the push-up Lemma 4.6. Moreover, by openness of $\ll$ we have $I^-(r) \subseteq \bigcup_n I^-(r_n)$, and since $r_n \ll q$, by transitivity we have $I^-(r_n) \subseteq I^-(q)$. Hence $I^-(p) \subseteq I^-(r) \subseteq I^-(q)$, which together with our initial assumption $I^-(p) = I^-(q)$ implies that $I^-(p) = I^-(r)$. Since X is causally past-distinguishing we thus know that $p = r$, which contradicts the fact that $d(p,r) = \frac{\delta}{2} > 0$. Thus X must indeed also be past-distinguishing. □

4.4.3 Continuity of averaged volume functions

We conclude this section by generalizing the equivalence between causal continuity and the continuity of volume functions (shown in the smooth case by Dieckmann [Die88b, Proposition 2.5]) to establish when $t^\pm$ of Definition 4.41 are time functions.

Definition 4.49. Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. If for all $p, q \in X$,

(i) $I^+(p) \subseteq I^+(q)$ implies $I^-(q) \subseteq I^-(p)$, we say that X is *past reflecting*.

(ii) $I^-(p) \subseteq I^-(q)$ implies $I^+(q) \subseteq I^+(p)$, we say that X is *future reflecting*.

If X is both past and future reflecting, we say that it is *reflecting*.

We define causal continuity as in the smooth case. Aké et al. showed that this notion of causal continuity implies strong causality [ACS20, Proposition 3.15], as it does in the smooth case. However, it turns out that the optimal condition for our Theorem 4.52 is a weaker version thereof. This weaker version is not (trivially) sufficient for the results of Aké et al. to still hold. In any case, both definitions are equivalent when the conditions of Proposition 4.48 are met (in particular, in the smooth case).

Definition 4.50 ([ACS20, Definition 3.9])**.** A Lorentzian pre-length space is called *causally continuous* if it is reflecting and distinguishing.

Definition 4.51. A Lorentzian pre-length space is called *weakly causally continuous* if it is reflecting and causally distinguishing.

Using this terminology we can state the main result of this section (see Theorem 4.2 for the corresponding Lorentzian length space version).

Theorem 4.52. *Let X be an approximating Lorentzian pre-length space equipped with a Borel probability measure of full support. Then X is weakly causally continuous if and only if the averaged volume functions $t^\pm$ are time functions.*

Note that also in this section we combine a topological condition (second countable) and a causal condition ((weakly) causally continuous) to characterize the existence of volume time functions.

Proof. By Proposition 4.47 the averaged volume functions are generalized time functions if and only if X is causally distinguishing. In Lemma 4.55 below we show that their continuity is characterized by X being reflecting. $\square$

First we show that property (iii) discussed in Remark 4.42 holds for almost all thickenings of $I^{\pm}(p)$.

Lemma 4.53. *Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space equipped with a Borel probability measure μ of full support. For $p \in X$ consider the functions $r \mapsto V_r^{\pm}(p) = \mu(I_r^{\pm}(p))$ as defined in Section 4.4.1 and*

$$r \mapsto \overline{V_r^{\pm}}(p) := \mu\left(\overline{I_r^{\pm}(p)}\right).$$

Then the sets

$$S^{\pm}(p) := \left\{r \in (0,1) \mid \overline{V_r^{\pm}}(p) \neq V_r^{\pm}(p)\right\}$$

are countable in $(0,1)$ for all $p \in X$.

Proof. Let $p \in X$. By definition and additivity of μ we have

$$\begin{aligned}\overline{V_r^{\pm}}(p) - V_r^{\pm}(p) &= \mu(\overline{I_r^{\pm}(p)}) - \mu(I_r^{\pm}(p)) \\ &= \mu\left(\overline{I_r^{\pm}(p)} \setminus I_r^{\pm}(p)\right) = \mu(\partial I_r^{\pm}(p)).\end{aligned}$$

Hence we can rewrite $S^{\pm}(p)$ as

$$S^{\pm}(p) = \left\{r \in (0,1) \mid \mu\left(\partial I_r^{\pm}(p)\right) > 0\right\}.$$

To see that $S^{\pm}(p)$ is countable, consider the sets

$$S_n^{\pm}(p) := \left\{r \in (0,1) \;\middle|\; \mu\left(\partial I_r^{\pm}(p)\right) > \frac{1}{n}\right\},$$

for $n \in \mathbb{N}$. We show that $|S_n^{\pm}(p)| \leq n$. Otherwise there would exist at least $n+1$ distinct $r_i \in S_n^{\pm}(p)$. Since all $\partial I_{r_i}^{\pm}(p)$ are disjoint, this would imply that

$$\frac{n+1}{n} < \sum_{i=1}^{n+1} \mu(\partial I_{r_i}^{\pm}(p)) \leq \mu(X) = 1,$$

a contradiction. Finally, because $S^{\pm}(p) = \bigcup_{n\in\mathbb{N}} S_n^{\pm}(p)$, we deduce that the sets $S^{\pm}(p)$ are countable for any $p \in X$. $\square$

In addition to Lemma 4.53, the continuity of the averaged volume functions rests on the following general result which is based on Lebesgue's dominated convergence theorem (in [Jos05] formulated for $X = \mathbb{R}^d$ but true for all sequential spaces).

Theorem 4.54 ([Jos05, Theorem 16.10])**.** *Let (X, d) be a metric space, $U \subseteq X$ and $y_0 \in U$. Consider a function $f\colon \mathbb{R}^n \times U \to \mathbb{R} \cup \{\pm\infty\}$. Assume that*

(i) *for every fixed* $y \in U$ *the function* $x \mapsto f(x, y)$ *is integrable,*

(ii) *for almost all* $x \in \mathbb{R}^n$ *the function* $y \mapsto f(x, y)$ *is continuous at* y_0*,*

(iii) *there exists an integrable function* $F\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *with the property that for every* $y \in U$*,*

$$|f(x, y)| \leq F(x)$$

holds almost everywhere on $\mathbb{R}^n$*.*

Then the function

$$g(y) := \int_{\mathbb{R}^n} f(x, y)\, dx$$

is continuous at the point y_0*.*

With this, we can prove the last lemma of this section (compare with [Die88b, Proposition 1.6]). Together with Proposition 4.47, it constitutes the proof of Theorem 4.52.

Lemma 4.55. *Let* $(X, d, \ll, \leq, \tau)$ *be an approximating Lorentzian pre-length space equipped with a Borel probability measure* μ *of full support. For any point* $q \in X$*, the following are equivalent:*

(i) t^- *(*t^+*) is continuous at* q*.*

(ii) $I^+(q) \subseteq I^+(p) \implies I(\tilde{p}, p) \cap I^-(q) \neq \emptyset$ *for all* $\tilde{p} \ll p$
($I^-(q) \subseteq I^-(p) \implies I(p, \tilde{p}) \cap I^-(q) \neq \emptyset$ *for all* $p \ll \tilde{p}$*).*

(iii) X *is past (future) reflecting at* q*.*

(iv) $\bigcap_{q \ll x} I^-(x) \subseteq \overline{I^-(q)}$
($\bigcap_{x \ll q} I^+(x) \subseteq \overline{I^+(q)}$*).*

Proof. We show the past versions. Fix $q \in X$.

(i) $\implies$ (ii) Assume (i) holds but not (ii). Then there exist points p and $\tilde{p}$ such that $\tilde{p} \ll p$ and $I^+(q) \subseteq I^+(p)$ but $I(\tilde{p}, p) \cap I^-(q) = \emptyset$. Note that $I(\tilde{p}, p)$ is open and non-empty, because we can approximate p from the past, and by openness of $\ll$, any past-approximating sequence must enter $I^+(\tilde{p})$. Let $r_0 > 0$ be small enough, then also the set

$$B := \left\{x \in I(\tilde{p}, p) \mid d(x, I^-(q)) > r_0\right\} = I(\tilde{p}, p) \setminus \overline{I^-_{r_0}(q)}.$$

is open and non-empty: If not, then $I(\tilde{p}, p) \subseteq \overline{I^-_{r_0}(q)}$. But since $I(\tilde{p}, p)$ is open, this would mean that $I(\tilde{p}, p) \cap I^-(q) \neq \emptyset$, contradicting our earlier assumption. Clearly, $B \cap I^-_r(q) = \emptyset$ for all $r < r_0$. On the other hand, since $q_n^+ \in I^+(q) \subseteq I^+(p)$, transitivity of $\ll$ implies that $I(\tilde{p}, p) \subseteq I^-(p) \subseteq I^-(q_n^+)$. For a sequence (q_n^+) approximating q from the future we have

$$t^-(q_n^+) - t^-(q) = \int_0^1 \mu\left(I^-_r(q_n^+) \setminus I^-_r(q)\right) dr.$$

Since $B \subseteq I_r^-(q_n^+) \setminus I_r^-(q)$ for $r < r_0$, it follows that

$$t^-(q_n^+) - t^-(q) \geq \int_0^{r_0} \mu(B)\, dr > 0,$$

showing discontinuity of t^- at q, a contradiction to (i).

(ii) $\Longrightarrow$ (iii) Assume (ii) holds but not (iii). If X is not past reflecting at q, then there exists a $p \in X$ with $I^+(q) \subseteq I^+(p)$ but $I^-(p) \nsubseteq I^-(q)$. Thus there is a $\tilde{p} \in I^-(p) \setminus I^-(q)$ and by (ii) there exists a point $\hat{p} \in I(\tilde{p}, p) \cap I^-(q)$. But then $\tilde{p} \ll \hat{p} \ll q$, and by transitivity $\tilde{p} \ll q$, a contradiction.

(iii) $\Longrightarrow$ (iv) If $\bigcap_{q \ll x} I^-(x) = \emptyset$ the conclusion is trivial. Suppose $p \in \bigcap_{q \ll x} I^-(x)$. Then $x \in I^+(p)$ for all $x \gg q$, i.e., $I^+(q) \subseteq I^+(p)$. By (iii) X is past reflecting at q, hence $I^-(p) \subseteq I^-(q)$. Since X is past-approximating, $p \in J^-(p) \subseteq \overline{I^-(p)} \subseteq \overline{I^-(q)}$. Since p was an arbitrary point in the intersection, (iv) follows.

(iv) $\Longrightarrow$ (i) Fix $r \in (0,1)$ and let $\epsilon > 0$. Because X is past-approximating, we can find a sequence $(q_i^-)_i$ that approximates q from the past. By openness of $\ll$, it then follows that

$$\bigcup_{i=1}^{\infty} I_r^-(q_i^-) = I_r^-(q),$$

and by standard measure theory [KP08, Theorem 1.2.5] there exists $i_0 \in \mathbb{N}$, such that

$$\mu\left(I_r^-(q)\right) - \mu\left(I_r^-(q_{i_0}^-)\right) < \epsilon.$$

By Lemma 4.44, we deduce

$$V_r^-(q) - V_r^-(p) < \epsilon \quad \text{for all} \quad p \in I^+(q_{i_0}^-). \tag{4.1}$$

Next, consider $(q_j^+)_j$ a sequence approximating q from the future. Assumption (iv) implies that

$$\bigcap_{j=1}^{\infty} I_r^-(q_j^+) \subseteq \overline{I_r^-(q)},$$

and hence

$$\mu\left(\bigcap_{j=1}^{\infty} I_r^-(q_j^+)\right) \leq \mu\left(\overline{I_r^-(q)}\right).$$

By [KP08, Theorem 1.2.5], for r given, there exists $j_0 \in \mathbb{N}$ such that

$$\mu\left(I_r^-(q_{j_0}^+)\right) - \mu\left(\overline{I_r^-(q)}\right) < \epsilon.$$

Then, using Lemma 4.44 we deduce

$$V_r^-(p) - \overline{V_r^-}(q) < \epsilon \quad \text{for all} \quad p \in I^-(q_{j_0}^+). \tag{4.2}$$

By Lemma 4.53, $\overline{V_r^-}(q) = V_r^-(q)$ for all but countably many r. Hence for almost all $r \in (0,1)$, we can combine (4.1) and (4.2) to write

$$|V_r^-(q) - V_r^-(p)| < \epsilon \quad \text{for all} \quad p \in I(q_{i_0}^-, q_{j_0}^+).$$

We conclude that almost all functions $V_r^-: X \to \mathbb{R}$, $r \in (0,1)$, are continuous at q. Thus it follows from Theorem 4.54 that

$$t^-(p) = \int_0^1 V_r^-(p)\,dr$$

is continuous at q. □

4.5 Global hyperbolicity and Cauchy time functions

The highest step on the causal ladder, namely global hyperbolicity, is fundamental for a number of deep and important results in general relativity, such as the study of the Cauchy problem of the Einstein equations, the singularity theorems, and Lorentzian splitting theorems. This is due to the fact global hyperbolicity is equivalent to the existence of a Cauchy time function, whose level sets in turn are Cauchy surfaces, i.e., domains suitable for specifying initial data for hyperbolic PDEs, and for imposing focusing conditions for geodesics. This characterization of global hyperbolicity was first obtained by Geroch [Ger70] in 1970 and makes use of volume time functions. In the same vein, in this section we characterize global hyperbolicity for Lorentzian pre-length spaces in four different ways by also utilizing our constructions from Section 4.4.

4.5.1 Definitions and main result

The causality conditions we use for Lorentzian pre-length space are defined analogously to the smooth case as follows (cf. [Min18]).

Definition 4.56 ([KS18, Definition 2.35])**.** A Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called *globally hyperbolic* if it is non-totally imprisoning and the causal diamonds $J(p,q)$ are compact for all $p, q \in X$.

Definition 4.57. A time function $t\colon X \to \mathbb{R}$ on a Lorentzian pre-length space $(X, d, \ll, \leq, \tau)$ is called a *Cauchy time function* if for every doubly-inextendible causal curve γ we have $\mathrm{Im}(t \circ \gamma) = \mathbb{R}$.

For smooth and continuous Lorentzian metrics, global hyperbolicity is also characterized by the existence of a Cauchy surface, which is then a topological (even smooth) hypersurface. We extend the definition verbatim, but adopt the name *Cauchy set* to emphasize that we are not in the manifold setting.

Definition 4.58. Let $(X, d, \ll, \leq, \tau)$ be a Lorentzian pre-length space. A *Cauchy set* is a subset $\mathcal{S} \subseteq X$ such that every doubly-inextendible causal curve intersects $\mathcal{S}$ exactly once.

Geroch [Ger70, Section 4] makes use of Leray's notion of global hyperbolicity which is formulated in terms of the topology on the collection of certain curves. We will show that Definition 4.56 is equivalent to this notion also for Lorentzian pre-length spaces. To this end, for any two points $p, q \in X$, we consider the set $\mathcal{C}(p,q)$ the equivalence class of future-directed causal curves from p to q with continuous, strictly monotonically increasing reparametrizations, equipped with the Hausdorff distance between the images of the curves as subsets in X, i.e.,

$$d_H(\gamma_1,\gamma_2) = \max\{\sup_{x\in\gamma_1} d(x,\gamma_2), \sup_{y\in\gamma_2} d(\gamma_1,y)\}$$

(we write γ_i also for the image $\mathrm{Im}(\gamma_i)$, since the parametrization does not matter).

In this section we establish the third, and last, main result of this paper (which via 4.23 immediately implies the Lorentzian length version stated in Theorem 4.3 of the Introduction).

Theorem 4.59. *Let $(X, d, \ll, \leq, \tau)$ be an approximating Lorentzian pre-length space with limit curves. Suppose, in addition, that (X, d) is second countable and proper. Then the following are equivalent:*

(i) X is globally hyperbolic,

(ii) X is non-totally imprisoning and $\mathcal{C}(p,q)$ is compact, for any $p, q \in X$,

(iii) X admits a Cauchy time function,

(iv) X admits a Cauchy set.

Remark 4.60 (Smooth spacetimes)**.** Manifolds are second-countable by definition. Moreover, any differentiable manifold admits a complete Riemannian metric [NO61], and hence, by the Hopf–Rinow Theorem, a proper distance. For spacetimes with continuous metrics, however, Theorem 4.59 cannot be applied unrestrictedly, since only causally plain C^0-spacetimes satisfy the axioms of Lorentzian pre-length spaces [KS18, Example 5.2] (but a characterization of global hyperbolicity on *all* C^0-spacetimes has been obtained in [Säm16]). On the other hand, Theorem 4.59 is of course valid well beyond the manifold setting. For instance, by the Hopf–Rinow–Cohn-Vossen Theorem it is sufficient that (X, d) is a complete, locally compact length metric space for it to be proper.

Remark 4.61 (Topology change)**.** On a spacetime (M, g), if $\mathcal{S}$ is a smooth Cauchy surface, then any other Cauchy surface is diffeomorphic to $\mathcal{S}$ [BS03] (similarly, if one works with continuous Cauchy surfaces, then they are homeomorphic). The diffeomorphism can be constructed by following the flow of the time-orientation vector field. Further, M is foliated by Cauchy surfaces.

For Lorentzian pre-length spaces X, it is still true that in the setting of Theorem 4.59 the level sets of Cauchy time functions yield a decomposition of X as a disjoint union of Cauchy sets. Different Cauchy sets, however, need not be homeomorphic, nor even homotopy equivalent as the next examples show.

Example 4.62 (Degenerate subsets of Minkowski space)**.** Let (M, η) be $(n+1)$-dimensional Minkowski spacetime, with coordinates $(t, x) \in \mathbb{R} \times \mathbb{R}^n$. Define

$$X := \{(t, x) \mid t \leq 0, x = 0 \text{ or } t > 0, |x| = t/2\}$$

(here $|x|$ denotes the Euclidean norm of $x \in \mathbb{R}^n$), equipped with the causal and chronological relations induced by η and the Euclidean distance. Then X is a globally hyperbolic Lorentzian pre-length space and satisfies the assumptions of Theorem 4.59. The function $f(t,x) = t$ is a Cauchy time function, with some of its level sets being points (if $t \leq 0$) and some being $(n-1)$-spheres (if $t > 0$).

Example 4.63 (Degenerate generalized cones)**.** If (X_1, d_1), (X_2, d_2) are separable, proper geodesic length spaces one can construct generalized cones Y_1 and Y_2 in the sense of [Ale+19] over them and glue them together at the tip. The resulting space can be equipped with the structure of a Lorentzian length space in the usual way, which is then globally hyperbolic by [Ale+19, Prop. 4.10], and has Cauchy sets homeomorphic to X_1, X_2 and $\{\text{tip}\}$.

4.5.2 Properties of Cauchy sets

In what follows we prove several results involving Cauchy sets that are crucial in the proof of Theorem 4.59 for the implication (iv) $\Longrightarrow$ (ii), but are also of independent interest. We make use of several results about inextendible causal curves, in particular, about the existence of maximal extension of causal curves obtained in Section 4.2.1.

Proposition 4.64 (Basic properties of Cauchy sets)**.** *Let $(X, d, \ll, \leq, \tau)$ be an approximating Lorentzian pre-length space with limit curves such that (X, d) is proper. If $\mathcal{S} \subseteq X$ is a Cauchy set, then the following properties hold:*

(i) $\mathcal{S}$ is acausal, *i.e., distinct points on $\mathcal{S}$ are not causally related,*

(ii) $X = J^-(\mathcal{S}) \cup J^+(\mathcal{S})$,

(iii) $X = I^-(\mathcal{S}) \sqcup \mathcal{S} \sqcup I^+(\mathcal{S})$, where $\sqcup$ denotes the disjoint union,

(iv) $J^\pm(\mathcal{S}) = \overline{I^\pm(\mathcal{S})} = I^\pm(\mathcal{S}) \sqcup \mathcal{S}$, and $\mathcal{S}$ and $J^\pm(\mathcal{S})$ are closed.

Proof. (i) Suppose $p < q$ and $p, q \in \mathcal{S}$. Then by causal path-connectedness, there exists a causal curve $\gamma\colon [a,b] \to X$ from p to q. By Proposition 4.25 there exists a maximal doubly-inextendible extension λ of γ, which therefore intersects $\mathcal{S}$ more than once, a contradiction to $\mathcal{S}$ being a Cauchy set.

(ii) Let $p \in X$. By Corollary 4.26, there exists a doubly-inextendible causal curve γ through p. Since $\mathcal{S}$ is a Cauchy set, it intersects γ exactly once, say at $\gamma(0)$. Hence $p \leq \gamma(0)$ or $\gamma(0) \leq p$.

(iii) Let $p \in X$. By Lemma 4.22, we can assume that the curve γ in (ii) is timelike in a neighborhood of p. Then, if $p \leq \gamma(0)$, we have either $p = \gamma(0) \in \mathcal{S}$ or $p \ll \gamma(-\epsilon) < \gamma(0)$ for some $\epsilon > 0$, which then by the push-up Lemma 4.6 implies $p \ll \gamma(0)$. Applying the same argument in the case $\gamma(0) \leq p$, we conclude that $p \in I^-(\mathcal{S}) \cup \mathcal{S} \cup I^+(\mathcal{S})$.

It remains to be shown that the union is disjoint. Since $\mathcal{S}$ is acausal by (i) and the push-up property it is clear that $\mathcal{S}$ and $I^\pm(\mathcal{S})$ are disjoint. Moreover, $I^+(\mathcal{S})$ and $I^-(\mathcal{S})$ are disjoint since otherwise, there exists a causal curve that intersects

$\mathcal{S}$ at two different points or is closed timelike (if the points coincide), and hence contradict $\mathcal{S}$ being Cauchy.

(iv) The sets $I^\pm(\mathcal{S}) = \bigcup_{p\in\mathcal{S}} I^\pm(p)$ are unions of open sets and hence open. By (iii) is $\mathcal{S} = X \setminus (I^-(\mathcal{S}) \sqcup I^+(\mathcal{S}))$ the complement of open sets and thus closed. By the same argument is $X \setminus I^\mp(\mathcal{S}) = I^\pm(\mathcal{S}) \sqcup \mathcal{S}$ closed, and hence

$$\overline{I^\pm(\mathcal{S})} \subseteq \overline{I^\pm(\mathcal{S}) \sqcup \mathcal{S}} = I^\pm(\mathcal{S}) \sqcup \mathcal{S} \subseteq J^\pm(\mathcal{S}) \subseteq \overline{I^\pm(\mathcal{S})},$$

where the last inclusion is due to the assumption that X is approximating. Thus $J^\pm(\mathcal{S}) = \overline{I^\pm(\mathcal{S})}$, and hence closed. □

Closedness of Cauchy sets, in particular, will be key in establishing (iv) $\Longrightarrow$ (ii) of Theorem 4.59. In the remaining lemmas of this subsection we prove this implication in steps.

Lemma 4.65. *Let $(X, d, \ll, \leq, \tau)$ be an approximating Lorentzian pre-length space with limit curves. Suppose, in addition, that (X, d) is proper. If X contains a Cauchy set $\mathcal{S}$, then X is non-totally imprisoning.*

Proof. Suppose X is totally imprisoning. Then by Theorem 4.30, there is a compact set $K \subseteq X$ and a doubly-inextendible future-directed causal curve $\gamma\colon \mathbb{R} \to K$. By Definition 4.58 we know that, without loss of generality, $\gamma \cap \mathcal{S} = \{\gamma(0)\}$.

First we show that $\tilde\gamma := \gamma|_{[1,\infty)}$ must come arbitrarily close to $\mathcal{S}$, that is, there exist parameter values s_n such that $d(\tilde\gamma(s_n), \mathcal{S}) < 1/n$, for all $n \in \mathbb{N}$. Suppose this is not the case: then there exists a δ such that $d(\tilde\gamma(s), \mathcal{S}) > \delta$ for all $s \in [1, \infty)$, and hence $\tilde\gamma$ is contained in the compact set $K \setminus \mathcal{S}_\delta$, where $\mathcal{S}_\delta := \{x \in X \mid d(x, \mathcal{S}) < \delta\}$. But then, since $\tilde\gamma$ is future-inextendible, by the proof of Theorem 4.30, there also exists a doubly-inextendible causal curve contained in $K \setminus \mathcal{S}_\delta$, in contradiction to $\mathcal{S}$ being a Cauchy set. Thus we conclude that there must exist a sequence $(s_n)_n$ such that $d(\tilde\gamma(s_n), \mathcal{S}) \to 0$. Moreover, since $\mathcal{S}$ is closed by Proposition 4.64 (iv), $K \cap \mathcal{S}$ is compact, and hence $(\tilde\gamma(s_n))_n$ converges, up to a subsequence, to some point $p \in \mathcal{S}$. This implies, since $\gamma \cap \mathcal{S} = \{\gamma(0)\}$, that $s_n \to \infty$.

Having established the existence of a sequence $s_n \to \infty$ in $\mathbb{R}$ such that $\gamma(s_n) \to p \in \mathcal{S}$, we may consider the sequence of past-directed causal curves $\lambda_n : [0, s_n] \to X$ given by

$$\lambda_n(s) = \gamma(s_n - s).$$

By Lemma 4.12, γ has infinite arclength, and thus $L^d(\lambda_n) \to \infty$. Since $\lambda_n(0) \to p$, we can apply Theorem 4.19 to find a past-directed limit curve $\lambda : [0, \infty) \to X$ with $\lambda(0) = p$, hence $\lambda([0, \infty)) \subseteq J^-(p) \subseteq J^-(\mathcal{S})$. On the other hand, we have that $\gamma([0, \infty)) \subseteq J^+(\gamma(0)) \subseteq J^+(\mathcal{S})$, therefore $\lambda_n([0, s_n]) \subseteq J^+(\mathcal{S})$. By Proposition 4.64(iv), $J^+(\mathcal{S})$ is closed, so as a limit curve, λ is contained in it. But then $\lambda([0, \infty)) \subseteq J^-(\mathcal{S}) \cap J^+(\mathcal{S})$. This is a contradiction since, by Proposition 4.64 (iv), $J^-(\mathcal{S}) \cap J^+(\mathcal{S}) = \mathcal{S}$, and by Proposition 4.64 (i), Cauchy surfaces are acausal. □

Finally, the last two lemmas prove compactness of $\mathcal{C}(p, q)$. They are adapted from Geroch's original proof for smooth spacetimes [Ger70]. Similarly to $\mathcal{C}(p, q)$, we denote by $\mathcal{C}(p, \mathcal{S})$ the space of causal curves from a point $p \in X$ to a Cauchy set $\mathcal{S}$, equipped with the Hausdorff distance.

Lemma 4.66. *Let $(X, d, \ll, \leq, \tau)$ be an approximating Lorentzian pre-length space with limit curves. Suppose, in addition, that (X, d) is proper. Let $\mathcal{S}$ be a Cauchy surface in X. Then for all $p \in X$, the space $\mathcal{C}(p, \mathcal{S})$ is compact.*

Proof. Suppose $\mathcal{C}(p, \mathcal{S}) \neq \emptyset$, otherwise the statement is trivial. In particular, this means that $p \in J^-(\mathcal{S})$. The set $\mathcal{C}(p, \mathcal{S})$ together with the Hausdorff distance d_H is a metric space, and hence it remains to prove sequential compactness. Let $(\gamma_n)_n$ be a sequence in $\mathcal{C}(p, \mathcal{S})$. We distinguish between two cases:

1. There exists a constant $C > 0$ such that $L^d(\gamma_n) < C$ for all $n \in \mathbb{N}$. Then by Theorem 4.16, a subsequence of $(\gamma_n)_n$ converges uniformly (and thus also with respect to d_H) to a future-directed causal limit curve $\gamma\colon [0, 1] \to X$. Since $\gamma_n(0) = p$, also $\gamma(0) = p$. Moreover, $\gamma_n(1) \in \mathcal{S}$ for all $n \in \mathbb{N}$, and $\mathcal{S}$ is closed by Proposition 4.64 (iv), thus also $\gamma(1) \in \mathcal{S}$. Hence $\gamma \in \mathcal{C}(p, \mathcal{S})$.

2. On the other hand, assume that $L^d(\gamma_n) \to \infty$. By Theorem 4.19 a subsequence of $(\gamma_n)_n$ converges locally uniformly to a future-inextendible causal limit curve $\gamma\colon [0, \infty) \to X$. Since $\gamma_n \subseteq J^-(\mathcal{S})$, and by Proposition 4.64(iv), $J^-(\mathcal{S})$ is closed, we have $\gamma \subseteq J^-(\mathcal{S})$. By Proposition 4.25, there exists a doubly-inextendible extension $\tilde{\gamma}\colon \mathbb{R} \to X$ of γ, which by transitivity of $\leq$ is also contained in $J^-(\mathcal{S})$. Then $\tilde{\gamma}$ must intersect $\mathcal{S}$, say at $\tilde{\gamma}(s_0)$. This implies that for all $s > s_0$, $\gamma(s) \in J^-(\mathcal{S}) \cap J^+(\mathcal{S}) = \mathcal{S}$. Moreover, $\gamma|_{(s_0,\infty)}$ cannot be constant, because then by Lemma 4.12, γ would be future-extendible. Since by Proposition 4.64 (i) $\mathcal{S}$ is acausal, we arrive at a contradiction. □

Lemma 4.67. *Let $(X, d, \ll, \leq, \tau)$ be a approximating Lorentzian pre-length space with limit curves. Suppose, in addition, that (X, d) is proper. If X contains a Cauchy set $\mathcal{S}$, then $\mathcal{C}(p, q)$ is compact for all $p, q \in X$.*

Proof. By Lemma 4.65, if X contains a Cauchy set $\mathcal{S}$, then X is non-totally imprisoning and hence causal. Thus we may assume without loss of generality that $p < q$ so that $\mathcal{C}(p, q)$ is nontrivial. We prove sequential compactness, with $(\gamma_n)_n$ always denoting a sequence in $\mathcal{C}(p, q)$. In view of Proposition 4.64 (i) and (iii), we can distinguish three cases:

1. $p \in \mathcal{S}$ and $q \in I^+(\mathcal{S})$ (or analogously, $q \in \mathcal{S}$ and $p \in I^-(\mathcal{S})$): Then $(\gamma_n)_n$ can be seen as a sequence in $\mathcal{C}(\mathcal{S}, q)$. By Lemma 4.66, there exists a limit curve γ in $\mathcal{C}(\mathcal{S}, q)$. But since γ_n starts at p for all n, also γ must start at p, hence it is an element of $\mathcal{C}(p, q)$.

2. $p, q \in I^+(\mathcal{S})$ (or analogously, $p, q \in I^-(\mathcal{S})$): There exists a future-directed timelike curve λ from $\mathcal{S}$ to p. Construct a new sequence $(\tilde{\gamma}_n)_n$ by concatenating λ with γ_n. Then $(\tilde{\gamma}_n)_n$ is a sequence in $\mathcal{C}(\mathcal{S}, q)$ and thus has a causal limit curve $\tilde{\gamma}$ by Lemma 4.66. Because of how the sequence was constructed, $(\tilde{\gamma}_n)_n$ must be the concatenation of λ with a causal curve $\gamma \in \mathcal{C}(p, q)$ which is the Hausdorff limit of $(\gamma_n)_n$.

3. $p \in I^-(\mathcal{S})$ and $q \in I^+(\mathcal{S})$: By Proposition 4.25 we can extend each γ_n, and the maximal extension must intersect $\mathcal{S}$ exactly once, say at $\gamma(0)$. Because

of Proposition 4.64 (iv), $\gamma(0)$ cannot lie to the past of p nor to the future of q, hence it must in fact lie on γ_n. Consider the sequence $(\bar{\gamma}_n)_n$ where $\bar{\gamma}_n$ is the restriction of γ_n from p to $\gamma_n(0)$. By Lemma 4.66 a subsequence $(\bar{\gamma}_{n_k})_k$ converges to a limit curve $\bar{\gamma}$ in $\mathcal{C}(p, \mathcal{S})$. Similarly, consider the sequence $(\tilde{\gamma}_{n_k})_k$ of the restrictions of the original curves γ_{n_k} from $\gamma_{n_k}(0)$ to q. By Lemma 4.66, we may assume it converges to a limit curve $\tilde{\gamma}$ in $\mathcal{C}(\mathcal{S}, q)$. Since by construction the endpoints of $\bar{\gamma}$ and $\tilde{\gamma}$ agree on $\mathcal{S}$, we can join them to obtain a limit causal curve of (a subsequence of) the original sequence $(\gamma_n)_n$ from p to q. □

4.5.3 Proof of Theorem 4.59

We prove Theorem 4.59 in several steps. The implications (ii) $\Longrightarrow$ (i) and (iii) $\Longrightarrow$ (iv) are straightforward. The most involved step (iv) $\Longrightarrow$ (ii) follows from our results in Section 4.5.2 about properties of Cauchy sets. Finally, for the implication (i) $\Longrightarrow$ (iii) we show that the averaged volume functions of Section 4.4 have additional properties on globally hyperbolic Lorentzian pre-length spaces.

Proof of Theorem 4.59. (ii) $\Longrightarrow$ (i) By (ii), X is already non-totally imprisoning, and thus it remains to be shown that the causal diamonds $J(p, q)$ are compact for all $p \leq q$. If $p = q$, then $J(p, q) = \{p\}$ because non-total imprisonment implies causality. Suppose that $p < q$. Let $(x_n)_n$ be any sequence in $J(p, q)$. By causal path-connectedness of X, every x_n lies on a causal curve $\gamma_n \colon [0, 1] \to X$ from p to q. By (ii), the space of curves $\mathcal{C}(p, q)$ is compact, and hence a subsequence $(\gamma_{n_k})_k$ of $(\gamma_n)_n$ converges to a causal curve $\gamma \in \mathcal{C}(p, q)$ in the Hausdorff sense. In particular, for the corresponding points, $d(x_{n_k}, \gamma) \to 0$ as $k \to \infty$. Since γ itself is compact, a subsequence of $(x_{n_k})_k$ must converge to a point on γ. Hence $J(p, q)$ is compact.

(iii) $\Longrightarrow$ (iv) Suppose t is a Cauchy time function. Then the level sets of t are Cauchy sets: For $s \in \mathbb{R}$ consider the preimage $\mathcal{S} = t^{-1}(\{s\})$. Since t is a time function, any future-directed causal curve γ intersects $\mathcal{S}$ at most once. If γ is furthermore doubly-inextendible, and since t is Cauchy, $\mathrm{Im}(t \circ \gamma) = \mathbb{R}$ and thus $\mathcal{S} \cap \gamma \neq \emptyset$. Hence doubly-inextendible causal curves intersect $\mathcal{S}$ at exactly one point.

(iv) $\Longrightarrow$ (ii) Suppose X admits a Cauchy set. By Lemma 4.65, X is non-totally imprisoning. By Lemma 4.67, the set of future-directed causal curves $\mathcal{C}(p, q)$ between p and q is compact for any $p, q \in X$.

(i) $\Longrightarrow$ (iii) Let

$$t := \ln\left(-\frac{t^-}{t^+}\right).$$

By (i), X is globally is hyperbolic, and thus it follows immediately from Lemma 4.69 (see below) that t is a Cauchy time function. □

We finish the remaining parts of the proof of (i) $\Longrightarrow$ (iii). By definition, $I^\pm(p) \subseteq J^\pm(p)$ for all points p in a Lorentzian pre-length space X. If X is approximating then furthermore $J^\pm(p) \subseteq \overline{I^\pm(p)}$, and thus $\overline{I^\pm(p)} = \overline{J^\pm(p)}$. If X is globally hyperbolic, we can say even more, which will allow us to apply our results of Section 4.4.

Proposition 4.68. *Let $(X, d, \ll, \leq, \tau)$ be a causally path-connected, approximating Lorentzian pre-length space. If X is globally hyperbolic, then X is* causally simple, *meaning that $J^{\pm}(p)$ is closed and thus $J^{\pm}(p) = \overline{I^{\pm}(p)}$ for all $p \in X$. Moreover, if X is causally simple, then X is causally continuous.*

Proof. These statements were shown by Aké et al. for Lorentzian length spaces [ACS20, Propositions 3.13 & 3.14]. The same proof goes through for our assumptions, because the assumption of localizability is only needed to invoke [ACS20, Sequence Lemma 2.25] which in our case is replaced by Lemma 4.21. □

We also need the following result.

Lemma 4.69. *Let $(X, d, \ll, \leq, \tau)$ be an approximating Lorentzian pre-length space with limit curves. Suppose, in addition, that (X, d) is a proper metric space equipped with a Borel probability measure of full support and corresponding averaged volume functions $t^+ : X \to [-\infty, 0]$ and $t^- : X \to [0, \infty]$ (see Definition 4.41). If X is globally hyperbolic, then $t^{\pm}$ are time functions. Moreover, for every doubly-inextendible future-directed causal curve $\gamma : (a, b) \to X$ it holds that*

$$\lim_{s \to b} t^+(\gamma(s)) = \lim_{s \to a} t^-(\gamma(s)) = 0.$$

Proof. We follow [Die87, Satz II.20]. Suppose X is globally hyperbolic, approximating Lorentzian pre-length space with limit curves. Then by Proposition 4.68 X is causally continuous, and thus by Theorem 4.52 $t^{\pm}$ are time functions.

To show the second part of the statement, assume for contradiction that $\gamma : (a, b] \to X$ is a future-directed past-inextendible causal curve with

$$\lim_{s \to a} t^-(\gamma(s)) > 0. \tag{4.3}$$

On the other hand, by standard measure theory [KP08, Thm. 1.2.5] and Lebesgue's dominated convergence theorem,

$$\begin{aligned}
\lim_{s \to a} t^-(\gamma(s)) &= \lim_{s \to a} \int_0^1 \mu\left(I_r^-(\gamma(s))\right) dr \\
&= \int_0^1 \lim_{s \to a} \mu\left(I_r^-(\gamma(s))\right) dr \\
&= \int_0^1 \mu\left(\bigcap_{s > a} I_r^-(\gamma(s))\right) dr.
\end{aligned}$$

Assumption (4.3) implies that there exists an $r \in (0, 1)$ and a point

$$p \in \bigcap_{s > a} I_r^-(\gamma(s)).$$

This means that for every $s > a$, there exists a $q \in I^-(\gamma(s))$ such that $d(q, p) < r$. In particular, we can find a sequence $q_n \in I^-(\gamma(a + 1/n))$ such that $d(q_n, p) <$

r for all $n \in \mathbb{N}$. Because d is proper, the closed ball of radius r around p is compact, and hence q_n converges to a limit point q (up to a subsequence). Now for a given $s > a$, we choose n_0 such that $s > a + 1/n_0$. Then, by the push-up Lemma 4.6, $q_n \in I^-(\gamma(s))$ for all $n \geq n_0$. Because of this, $q \in \overline{I^-(\gamma(s))} = J^-(\gamma(s))$, where the last equality follows by global hyperbolicity and Proposition 4.68. Since $s > a$ was arbitrary, we have that $\gamma \subseteq J^+(q)$. In particular, $q \leq \gamma(b)$, and by global hyperbolicity, $J(q, \gamma(b))$ is compact. But then the curve γ is imprisoned in $J(q, \gamma(b))$, in contradiction to global hyperbolicity. $\square$

This concludes our characterization of global hyperbolicity with the existence of Cauchy time functions (and Cauchy sets) in the setting of Lorentzian pre-length spaces.

Chapter 5

Topology change with Morse functions: progress on the Borde–Sorkin conjecture

This chapter is based on the preprint [GH22] of the same title, which has been accepted for publication in *Advances in Theoretical and Mathematical Physics.*[1]

5.1 Introduction

In General Relativity, by solving the initial value problem for Einstein's Equations, one finds the time evolution of the spacetime metric. In this picture, the topology of the constant time slices always remains the same. Precisely speaking, the maximal globally hyperbolic development of some initial data V is, on the level of topology (Geroch [Ger70]) and differentiable structure (Bernal and Sánchez [BS03; BS05]), simply $V \times \mathbb{R}$. The question remains whether this rigid product structure is desirable, or whether we should allow the topology to change over time as well.

There are several instances where topology change is desirable. The dynamical creation of a wormhole, for example, is necessarily a topology changing process, as it involves attaching a handle to space. Already in 1957, Wheeler argued that quantum fluctuations of spacetime should modify the topology [Whe57]. Moreover, in certain approaches to Quantum Gravity, instead of considering the deterministic evolution of a spatial slice under the Einstein Equations, the idea is to find the transition probability between two spatial slices V_1, V_2. This is done by computing a path-integral over all cobordisms between V_1 and V_2; that is, manifolds $\mathcal{M}$ with boundary $\partial\mathcal{M} = V_1 \sqcup V_2$. These cobordisms also have to be equipped with a Lorentzian (or, in Euclidean Quantum Gravity, Riemannian) metric, and possibly

[1] I would like to thank Elefterios Soultanis and Maximilian Ruep for very interesting discussions, and Annegret Burtscher for comments on the draft. I am also grateful to Bernardo Araneda and Simon Pepin Lehalleur for pointing me to reference [KS21].

satisfy some additional conditions. It is then natural to think that the transition probability between V_1 and V_2 might be non-zero also when V_1 and V_2 are not homeomorphic, as long as appropriate cobordisms exist. We refer to [AL98; Dow02; Sor97] for further discussion on the role of topology change in Quantum Gravity.

Which properties should a Lorentzian cobordism satisfy, in order to consider it physically reasonable? In this paper, we will focus on the case of compact cobordisms (i.e. spatially closed universes). Geroch [Ger67] showed that any non-trivial (meaning with $V_1 \neq V_2$) compact Lorentzian cobordism must contain closed timelike curves. Because of this, the only way to have topology change without time travel is by allowing the spacetime metric to degenerate at certain singular points [Kun67; Yod72]. Notice that the case of non-compact time slices is less restricted, with examples of topology change without closed timelike curves and without singular points obtained by multiple authors (see Sánchez [Sán23] for the most recent ones and for the overview of previous work on p. 16).

One interpretation of a degenerate metric is to consider the singular points as naked singularities, and not as points in the spacetime manifold. In this paper, however, we do the opposite: we consider the singular points as points in the spacetime, where nothing special happens, except that, in some sense, the topology change happens there. Our point of view implies that the spacetime metric is not Lorentzian everywhere, but this is not so bad, since the metric is not a physical observable in itself. Indeed, we will show that the causal and length structures can be satisfactorily generalized to include the singular points (some work on the curvature has also been done [LS97]). Still, allowing degenerate metrics does introduce many new questions and problems (irrespective of our point of view on the singular points). Already in the 1980s, Anderson and DeWitt showed that on their famous "trousers spacetime", quantum fields create infinite bursts of energy in the presence of singular points [AD86]. This result was later refined and confirmed in Manogue et al. [MCD88] and Buck et al. [Buc+17]. The aim of subsequent work was to impose additional conditions that avoid such pathologies.

A concrete and very useful construction of degenerate Lorentzian metrics on cobordisms was given by Yodzis [Yod72; Yod73] using Morse functions. This idea was further developed by Sorkin [Sor89] and collaborators [Bor+99; DS98; DG98; DGS00a; DGS00b; LS97], under the name of *Morse geometries*. We continue this approach in the present paper.

The construction of a Morse geometry is as follows. Let $\mathcal{M}$ be a compact cobordism of dimension n, h a Riemannian metric, $\zeta > 1$ a constant, and f a Morse function. Recall that a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ is called a Morse function if all its critical points (where $df = 0$) are non-degenerate (not to be confused with the (non-)degeneracy of the spacetime metric). This is equivalent to saying that around each critical point, there exist coordinates x^i such that

$$f = \frac{1}{2} \sum_i a_i (x^i)^2, \tag{5.1}$$

where $a_i \neq 0$ are constants. It follows, in particular, that the critical points are isolated. The *index* of a critical point is defined as the number of negative a_i (see

[Mil63] for more details). Louko and Sorkin define the *Morse metric* corresponding to $(\mathcal{M}, h, f, \zeta)$ by

$$g = \|df\|_h^2 \, h - \zeta df \otimes df, \tag{5.2}$$

which, in coordinates, gives

$$g_{\mu\nu} := (h^{\alpha\beta} \partial_\alpha f \partial_\beta f) h_{\mu\nu} - \zeta \partial_\mu f \partial_\nu f. \tag{5.3}$$

Let $M = \mathcal{M} \setminus (\partial\mathcal{M} \cup \{p_i\}_i)$, where $\{p_i\}_i$ is the set of critical points of f. By abuse of notation, we call the restriction of g to M also g. Since df vanishes only at the critical points, g is Lorentzian on M, and the pair (M, g) forms a spacetime in the usual sense. It is clear from (5.2) that f is a time function on (M, g), when choosing the time orientation to be given by the gradient vector field $\nabla^\alpha f := h^{\alpha\beta} \partial_\beta f$. Following the nomenclature of Borde et al. [Bor+99], we call (M, g) a *Morse spacetime* and $(\mathcal{M}, h, f, \zeta)$ a Morse geometry[2]. It is known that for any pair of connected 3-manifolds, there exists a Morse geometry interpolating between them [DS98; DGS00b].

According to the following two conjectures, the infinite bursts of energy found by Anderson and DeWitt are only present on certain topology-changing spacetimes, but not on others.

Conjecture 5.1 (Sorkin)**.** *A quantum field propagating on a Morse geometry $(\mathcal{M}, h, f, \zeta)$ has an unphysical singular behaviour if and only if the Morse spacetime (M, g) is causally discontinuous.*

Conjecture 5.2 (Borde–Sorkin)**.** *The Morse spacetime (M, g) induced by a Morse geometry $(\mathcal{M}, h, f, \zeta)$ is causally continuous if and only if all critical points of f have index different from 1 and $n-1$.*

Recall that causal continuity roughly means that the past and future $I^\pm(p)$ varies continuously with the point p (see Appendix 5.6 for details). Causal continuity was introduced in Hawking and Sachs [HS74] as a minimal requirement for a spacetime to be physically reasonable, for reasons unrelated to quantum theory. Thus Conjecture 5.2 is also interesting beyond the obvious link to Conjecture 5.1. Conjecture 5.1 is mentioned as early as 1990 in Sorkin [Sor89], while the earliest reference for Conjecture 5.2 is an indirect source (Dowker and Garcia [DG98] from 1998). Both of them remain open to this day. Conjecture 5.2 has seen important progress through the works of Borde, Dowker, Garcia, Sorkin and Surya [Bor+99; DGS00a; DGS00b]. In this paper, we contribute to this effort by showing the following special case:

Theorem 5.3. *Let $(\mathcal{M}, h, f, \zeta)$ be a Morse geometry of dimension n with a single critical point $p_c \in \mathcal{M}$. Suppose that p_c has index $\lambda \neq 0, 1, n-1, n$, and is contained in a coordinate neighborhood where*

$$f = \frac{1}{2} \sum_i a_i (x^i)^2, \qquad h = \sum_i (dx^i)^2, \tag{5.4}$$

[2] In [DGS00b], the inverted nomenclature is used.

for some real constants $a_i \neq 0$ satisfying

$$\frac{1}{\zeta} < \left|\frac{a_i}{a_j}\right| < \zeta \quad \textit{and} \quad \frac{5}{8} \leq \left|\frac{a_i}{a_j}\right| \leq \frac{8}{5} \quad \textit{for all } i, j. \tag{5.5}$$

Then the corresponding Morse spacetime (M, g) is causally continuous.

In Section 5.3 (Proposition 5.15) we will show that one can *always* find coordinates where (5.4) holds, up to adding a perturbation to h which vanishes at p_c. Moreover, we conjecture that the first part of (5.5) is sharp, in the sense that its violation leads to causal discontinuity (see Example 5.17 and Conjecture 5.18).

Combining Theorem 5.3 with previous results by other authors (fleshed out below), we can summarize the current status of Conjecture 5.2 in the next theorem.

Theorem 5.4. *Let $(\mathcal{M}, h, f, \zeta)$ be a Morse geometry of dimension $n \geq 2$, and (M, g) the corresponding Morse spacetime. Assume f has a single critical point for each critical value.*

(i) If f has at least one critical point of index $\lambda = 1, n-1$, then (M, g) is causally discontinuous.

(ii) If each critical point of f has index $\lambda = 0, n$, or has any index $\lambda \neq 1, n-1$ and is contained in a neighborhood as in Theorem 5.3, then (M, g) is causally continuous.

The case $\lambda = 0, n$ in part (ii) was solved in Borde et al. [Bor+99], along with the special case of Theorem 5.3 corresponding to $|a_i| = 1$ for all i. Part (i) of Theorem 5.4 was shown in Dowker et al. [DGS00b]. Also in [DGS00b], it was shown that the case of multiple critical points (as long as there is only one per critical value) reduces to the case of a single critical point: the Morse spacetime is causally continuous if and only if every critical point has a causally continuous neighborhood.

The proof of Theorem 5.3 is contained in Section 5.2. In Section 5.3, we discuss the necessity of our assumptions, and possible generalizations of our proof. Based on this discussion, we propose a modified version of the Borde–Sorkin conjecture in Section 5.4, where we also give concluding remarks. Appendix 5.5 contains results of [Bor+99] that we need in our proofs, and Appendix 5.6 gives some background on causal continuity.

5.2 Proof of Theorem 5.3

Before starting, let us give a brief outline of the proof. Recall from the introduction that the case of $|a_i| = 1$ for all i has already been solved in Borde et al. [Bor+99], a result that we build upon. While in the case of $|a_i| = 1$ there are a lot of symmetries, which allow for good coordinate choices (see Appendix 5.5), this is no longer true in the general case. Our strategy is to extend the causal structure from (M, g) to $\mathcal{M}$, in a way that preserves its most important properties: openness of the chronological relation I^+, the push-up principle $J^+(I^+(q)) = I^+(J^+(q)) = I^+(q)$, and the properties of limits of causal curves. Once these properties are proven, causal continuity follows almost immediately, as it would in Minkowski spacetime.

The most difficult to establish, out of the three properties, is the openness of the chronological relation. We do this in Subsection 5.2.2. The argument is based on reduction to the $|a_i| = 1$ case. Once openness of the chronological relation is established, the rest of the proof can be performed without the need to make any coordinate choices whatsoever, and without further use of the assumptions (5.4) and (5.5). This second part of the proof is contained in Subsection 5.2.3. It requires heavy use of the limit curve theorems of Minguzzi [Min08a].

5.2.1 Notation and first steps

Throughout this section, we assume that $(\mathcal{M}, h, f, \zeta)$ is a Morse geometry of dimension $n \geq 4$, with a single critical point p_c of index $\lambda \neq 0, 1, n-1, n$ lying in the interior of $\mathcal{M}$. As in the introduction, we write $M := \mathcal{M} \setminus (\partial\mathcal{M} \cup \{p_c\})$, and g denotes the metric (5.2), which is Lorentzian on M and degenerate-Lorentzian on $\mathcal{M}$. We do not use Einstein's summation convention: all sums are written out, but without making explicit the summation limits. Hence $\sum_i$ means $\sum_{i=1}^n$, and similarly $\max_i$ means $\max_{i=1,\ldots,n}$. For convenience, we refer to the hypothesis of Theorem 5.3 as Condition 1.

Condition 1. *There exists an open set $\mathcal{U} \subseteq \mathcal{M}$ with $p_c \in \mathcal{U}$, an open ball $\mathcal{B} \in \mathbb{R}^n$ around the origin, and a coordinate chart $\phi : \mathcal{U} \to \mathcal{B}$ of $\mathcal{M}$ such that $\phi(p_c) = 0$ and*

$$f \circ \phi^{-1} = \frac{1}{2} \sum_i a_i (x^i)^2, \qquad h \circ \phi^{-1} = \sum_i (dx^i)^2, \tag{5.6}$$

for some real constants $a_i \neq 0$. Moreover, setting

$$\zeta_c := \max_{i,j} \left| \frac{a_i}{a_j} \right|,$$

we have

$$\zeta_c \leq \frac{8}{5} \qquad \text{and} \qquad \zeta_c < \zeta. \tag{5.7}$$

The value of ζ_c does not depend on the choice of coordinates, as long they satisfy (5.6) (we give a detailed argument for this in Section 5.3.1). We will usually suppress the coordinate map ϕ from the notation, and whenever we write x^i, it will refer to the coordinates as given by Condition 1. In these coordinates, the metric (5.2) takes the form

$$g = \sum_{i,j} \left(a_i x^i dx^j \right)^2 - \zeta \left(\sum_k a_k x^k dx^k \right)^2. \tag{5.8}$$

An important tool in our proof will be to reduce some computations to the case of *isotropic neighborhoods* as studied in Borde et al. [Bor+99] (see also Appendix

5.5). These are metrics where Condition 1 is satisfied, but with the stronger requirement that $|a_i| = 1$ for all i. The following lemma gives us such an isotropic neighborhood metric g_{iso} with lightcones narrower than those of g.

Lemma 5.5. *Suppose that Condition 1 is satisfied, and consider on $\mathcal{U}$ the linear change of coordinates*

$$x^i \mapsto y^i := \sqrt{|a_i|}\zeta_c^{\frac{1}{4}} x^i.$$

Then the tensor given in these new coordinates by

$$g_{\text{iso}} := \sum_{i,j} \left(y^i dy^j\right)^2 - \frac{\zeta}{\zeta_c}\left(\sum_k \operatorname{sign}(a_k) y^k dy^k\right)^2 \tag{5.9}$$

is a Lorentzian metric on $\mathcal{U} \setminus \{p_c\}$, with lightcones narrower that those of g.

Proof. By (5.7), $\zeta/\zeta_c > 1$, and hence g_{iso} is a *neighborhood metric* in the sense of Borde et al. (see Appendix 5.5). In particular, g_{iso} is Lorentzian everywhere except at the origin. In the y-coordinates, the metric (5.8) takes the form

$$g = \frac{1}{\zeta_c}\sum_{i,j}\left|\frac{a_i}{a_j}\right|\left(y^i dy^j\right)^2 - \frac{\zeta}{\zeta_c}\left(\sum_k \operatorname{sign}(a_k) y^k dy^k\right)^2.$$

For a vector $V \in T\mathcal{U}$ (with components V^i in the y-coordinates), this means

$$\begin{aligned} g(V,V) &= \frac{1}{\zeta_c}\sum_{i,j}\left|\frac{a_i}{a_j}\right|\left(y^i V^j\right)^2 - \frac{\zeta}{\zeta_c}\left(\sum_k \operatorname{sign}(a_k) y^k V^k\right)^2 \\ &\le \sum_{i,j}\left(y^i V^j\right)^2 - \frac{\zeta}{\zeta_c}\left(\sum_k \operatorname{sign}(a_k) y^k V^k\right)^2 \\ &= g_{\text{iso}}(V,V) \end{aligned}$$

Hence if $g_{\text{iso}}(V,V) \le 0$, then also $g(V,V) \le 0$. In other words, g_{iso} has narrower lightcones than g. $\square$

Another crucial element in the proof will be the extension of the causal relation from M to $\mathcal{M}$. Let $\gamma\colon I \to \mathcal{M}$ be a locally Lipschitz curve. By continuity, we can write $\gamma^{-1}(M) = \gamma^{-1}(\mathcal{M}) \setminus \gamma^{-1}(p_c)$ as a union of intervals $\bigcup_i I_i$. If $\gamma\colon I_i \to M$ is future-directed (f.d.) causal for every i, then we say that $\gamma\colon I \to \mathcal{M}$ is future-directed causal, and analogously for timelike and/or past-directed curves. This gives rise to a notion of futures and pasts $I^{\pm}_{\mathcal{M}}(p), J^{\pm}_{\mathcal{M}}(p)$ in $\mathcal{M}$. Additionally, for $p \neq p_c$, we denote by $I^{\pm}_M(p), J^{\pm}_M(p)$ the usual past and future sets in the spacetime (M,g).

The following lemma tells us that no causal curve can be imprisoned in a neighborhood of p_c (see [BEE96, pp. 61-62] for the definition of non-imprisonment on non-degenerate spacetimes).

Lemma 5.6. *Let $\gamma\colon (a,b) \to M$ be a causal curve which is future inextendible in M. Then either $\lim_{s\to b}\gamma(s) = p_c$ or γ runs into $\partial\mathcal{M}$.*

Proof. Because f is a time function on (M,g), (M,g) is strongly causal. Then, by [BEE96, Prop. 3.13], given any compact set $K \subseteq M$, there exists $\delta > 0$ such that $\gamma(s) \notin K$ for all $s \in (b-\delta, b)$ (in other words, γ must leave K and never enter it again). Let $U \subseteq \mathcal{M}$ be any open set (not necessarily connected) containing p_c and $\partial\mathcal{M}$. Then we can choose $K = \mathcal{M}\setminus U$, and hence there exists $\delta > 0$ such that $\gamma(s) \in U$ for all $s \in (b-\delta, b)$. Since U was arbitrary, we are done. □

5.2.2 Openness of chronological pasts and futures

In this subsection, we characterize the past $I^-_{\mathcal{M}}(p_c)$ of the critical point p_c. Every statement has a time-reversed analogue for the future $I^+_{\mathcal{M}}(p_c)$ (which we do not write out explicitly). The following condition is very important. It states that if from a point $q \in M$ we can reach p_c via timelike curves, then we can also reach a whole neighborhood of p_c. This is a well-known fact for spacetimes without singular points.

Condition 2 (Openness of $I^+_{\mathcal{M}}$)**.** *For every $q \in I^\pm_{\mathcal{M}}(p_c)$ there exists a neighborhood U of p_c such that $U \setminus \{p_c\} \subseteq I^\mp_M(q)$.*

An important consequence of Condition 2 is that the chronological relation is not altered by removing p_c.

Lemma 5.7 ($I^+_M = I^+_{\mathcal{M}} \cap M$)**.** *Suppose Condition 2 is satisfied. Then, for every $p \in M$ it holds that $I^+_M(p) = I^+_{\mathcal{M}}(p) \cap M$.*

Proof. The inclusion $I^+_M(p) \subset I^+_{\mathcal{M}}(p) \cap M$ is trivial. It remains to show the other direction. Let $q \in I^+_M(p)$, and let $\gamma : [a,b] \to \mathcal{M}$ be a timelike curve from p to q. If γ avoids p_c, there is nothing to prove. Hence suppose that $\gamma(c) = p_c$ for some c. Then $p_c \in I^+_{\mathcal{M}}(p) \cap I^-_{\mathcal{M}}(q)$, so by Condition 2 we can find neighborhoods U, V of p_c such that $U \setminus \{p_c\} \subseteq I^+_M(p)$ and $V \setminus \{p_c\} \subseteq I^-_M(q)$. But then we can find a point $z \in U \cap V \setminus \{p_c\}$, and a timelike curve $\sigma : [a,b] \to M$ from p to q passing through z. □

In Appendix 5.5 (Lemma 5.19), we show that Condition 2 holds for the isotropic metric g_{iso}, which is simpler than g, and has narrower lightcones (Lemma 5.5). Making use of this fact, we show through the following lemma that Condition 2 also holds for our metric of interest g.

Lemma 5.8. *Condition 1 implies Condition 2.*

The rest of this subsection is dedicated to proving Lemma 5.8. We start by discussing coordinate choices. Assume w.l.o.g. that we have ordered our coordinates x^i, where f,h take the form (5.6), so that $a_i < 0$ for $i = 1, ..., \lambda$ and $a_j > 0$ for $j = \lambda+1, ..., n$. We then define the following "radial" coordinates

$$r^2 := \frac{1}{2}\sum_{i=1}^{\lambda} -a_i (x^i)^2, \qquad \rho^2 := \frac{1}{2}\sum_{j=\lambda+1}^{n} a_j (x^j)^2.$$

By following the flow of the gradient vector ∇r (by which we mean the gradient taken with respect to h, so that $h(\nabla r, \cdot) = dr(\cdot)$) we get a diffeomorphism from $\mathbb{R}^{n-\lambda} \setminus \{0\}$ to $\mathbb{R} \times S^{\lambda-1}$. This gives us a coordinate system $(r, \theta_1, ..., \theta_{\lambda-1})$ on $\mathbb{R}^\lambda$, where we view $\mathbb{R}^\lambda$ as the subspace spanned by the x^i coordinates with $i = 1, ..., \lambda$. Essentially, all we have done is changing to polar coordinates, but it is important that we have done so in a way that the angular directions are g-orthogonal to the r-direction. We can do the same construction with ρ, obtaining coordinates $(\rho, \phi_1, ..., \phi_{n-\lambda-1})$ on $\mathbb{R}^{n-\lambda}$. Furthermore, we have

$$f = \rho^2 - r^2, \qquad \alpha := (\rho r)^{\frac{1}{p}}, \tag{5.10}$$

where $p > 0$ is a constant, f is just our Morse function, and α is chosen so that $h(\nabla f, \nabla\alpha) = 0$. Using $(f, \alpha, \theta_1, ..., \theta_\lambda, \phi_1, ..., \phi_{n-\lambda-1})$ as our coordinates, the Euclidean metric h takes the form

$$h = \frac{df^2}{\|\nabla f\|^2} + \frac{d\alpha^2}{\|\nabla\alpha\|^2} + h_\Theta + h_\Phi,$$

and thus the Morse metric g takes the form

$$g = -(\zeta - 1)df^2 + \frac{\|\nabla f\|^2}{\|\nabla\alpha\|^2}d\alpha^2 + \|\nabla f\|^2(h_\Theta + h_\Phi). \tag{5.11}$$

Here we have used that, by definition, $\|\nabla f\| = \|df\|$. Having chosen our coordinates, we now state a lemma that constitutes the most important step in the proof of Lemma 5.8.

Lemma 5.9. *Suppose Condition 1 is satisfied. Let $q \in I^-_{\mathcal{M}}(p_c)$, and let $\gamma\colon [0,1] \to \mathcal{M}$ be any f.d. timelike curve from $\gamma(0) = q$ to $\gamma(1) = p_c$, which we express in coordinates as*

$$\gamma(s) = (f(s), \alpha(s), \Theta(s), \Phi(s)). \tag{5.12}$$

Then, for every $0 < \varepsilon < \alpha(0)$, the curve $\sigma\colon [0, s_\varepsilon] \to \mathcal{M}$ given by

$$\sigma(s) := (f(s), \alpha(s) - \varepsilon, \Theta(s), \Phi(s)) \tag{5.13}$$

is f.d. timelike. Here $s_\varepsilon := \min\{s \in (0,1) \mid \alpha(s) = \varepsilon\}$.

Proof. The statement is trivially true if $\alpha(0) = 0$ (since then there exist no suitable ε), and otherwise s_ε is well-defined (the minimum exists) by continuity of $\alpha(s)$. Moreover, our choice of ϵ and s_ε ensures that $\alpha(s) - \varepsilon \geq 0$ for all $s \in [0, s_\varepsilon]$, so that the curve σ is also well-defined.

Note that shifting α by a constant ε while leaving f, Θ, Φ fixed (as is done in (5.13)), is equivalent to shifting both ρ^2 and r^2 by a quantity $\epsilon(s)$. We are going to show that $g(\dot\sigma(s), \dot\sigma(s)) \leq g(\dot\gamma(s), \dot\gamma(s)) < 0$. This will be done in multiple steps, corresponding to various terms in (5.11).

Step 1 (Angular part)**.** Let π_Θ denote the orthogonal projection onto the subspace of the tangent space spanned by the Θ angular directions. We will show that

$$h(\pi_\Theta\dot\sigma, \pi_\Theta\dot\sigma) \leq h(\pi_\Theta\dot\gamma, \pi_\Theta\dot\gamma).$$

An analogous statement holds for π_Φ. We proceed by computing $\pi_\Theta\dot\sigma$. Notice that shifting $r(s)^2$ to $r(s)^2-\epsilon(s)$ means following the flow $F\colon \mathcal{M}\times\mathbb{R}\to\mathcal{M}$ of the vector field ∇r for a certain time $t(s)>0$. Then

$$\dot\sigma(s) = DF(\gamma(s),t(s))\dot\gamma(s) + \frac{\partial F}{\partial t}(\gamma(s),t(s))\dot t. \tag{5.14}$$

Similary, shifting ρ^2 means following the flow of $-\nabla\rho$. Notice also that

$$\frac{\partial F}{\partial t}(\gamma(s),t(s)) = \nabla r(F(\gamma(s),t(s))).$$

Hence the second term on the RHS of (5.14) only adds a contribution to the r component of $\dot\sigma(s)$ (but not to the angular components). We can compute DF by solving the ODE

$$\frac{\partial}{\partial t}DF(x,t) = D(\nabla r)(F(x,t))DF(x,t)$$

with initial condition $DF(x,0)=\mathrm{Id}$. In Cartesian coordinates $D\nabla r$ takes a block diagonal form:

$$D\nabla r_{ij} = \begin{cases} a_i\delta_{ij} & \text{for } i,j=1,...,\lambda,\\ \delta_{ij} & \text{for } i,j=\lambda+1,...,n,\\ 0 & \text{otherwise,}\end{cases}$$

hence

$$DF(x,t)_{ij} = \begin{cases} e^{a_i t}\delta_{ij} & \text{for } i,j=1,...,\lambda,\\ \delta_{ij} & \text{for } i,j=\lambda+1,...,n,\\ 0 & \text{otherwise.}\end{cases} \tag{5.15}$$

Moreover, we have that $DF(x,t)\partial_r \propto \partial_r$ because F is the flow of a vector field collinear to ∂r. Therefore

$$\pi_\Theta DF\ V = \pi_\Theta DF\pi_\Theta V \text{ for any } V\in TM,$$

and thus we can write

$$\begin{aligned} h(\pi_\Theta\dot\sigma,\pi_\Theta\dot\sigma) &= h(\pi_\Theta DF\pi_\Theta\dot\gamma,\pi_\Theta DF\pi_\Theta\dot\gamma)\\ &\le h(DF\pi_\Theta\dot\gamma, DF\pi_\Theta\dot\gamma)\\ &\le h(\pi_\Theta\dot\gamma,\pi_\Theta\dot\gamma).\end{aligned}$$

Here we have first used that the orthogonal projection π_Θ cannot increase the norm, and then that DF cannot increase the norm either, because it does not increase any of the Cartesian components (5.15).

Step 2 (α direction)**.** We want to show that $\frac{\|\nabla f\|^2}{\|\nabla\alpha\|^2}$ does not increase when shifting r^2 and ρ^2 by ϵ, so that we do not get a larger contribution in (5.11). Thus in what follows we view r and ρ as functions of ϵ, in the sense that $r^2=r_0^2+\epsilon$ and $\rho^2=\rho_0^2+\epsilon$ with respect to some reference values r_0,ρ_0 (but we will omit the subscript 0 from the notation). From this point of view, what we want to show is

$$\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\frac{\|\nabla f\|^2}{\|\nabla\alpha\|^2}\ge 0.$$

We begin with some preliminary computations, where $\nu := 2 - \frac{1}{p}$, and all derivatives are evaluated at $\epsilon = 0$.

$$\begin{aligned}
\frac{\partial}{\partial\epsilon}\rho^2 &= 1, \\
\frac{\partial}{\partial\epsilon}\rho^4 &= 2\rho^2, \\
\frac{\partial}{\partial\epsilon}(r\rho)^{2\nu} &= \nu(r\rho)^{2\nu-2}(r^2+\rho^2), \\
\|\nabla f\|^2 &= \|df\|^2 = \|d(r^2)\|^2 + \|d(\rho^2)\|^2, \\
\|\nabla\alpha\|^2 &= \|d\alpha\|^2 = \frac{1}{(2p(r\rho)^\nu)^2}\left(\rho^4\|d(r^2)\|^2 + r^4\|d(\rho^2)\|^2\right).
\end{aligned}$$

Moreover, we need the following estimates,

$$\begin{aligned}
a\rho^2 \le \|d(\rho^2)\|^2 &\le A\rho^2, \\
a \le \frac{\partial}{\partial\epsilon}\|d(\rho^2)\|^2 &\le A,
\end{aligned}$$

where $a = 2\min_{i=1,\dots,n} a_i$, $A = 2\max_{i=1,\dots,n} a_i$. These are easily proven in Cartesian coordinates.

Applying the chain rule and substituting the previous computations and estimates, we get, after a lengthy but trivial computation, the estimate

$$\frac{\partial}{\partial\epsilon}\frac{\|\nabla f\|^2}{\|\nabla\alpha\|^2} \ge \frac{(\nu a^2 - A^2)(r^6+\rho^6) + \left((3\nu+2)a^2 - 5A^2\right)(\rho^4 r^2 + \rho^2 r^4)}{(2p(r\rho)^\nu)^2\|\nabla\alpha\|^4},$$

where the RHS is guaranteed to be positive if

$$\frac{A^2}{a^2} \le \min\{\nu, \frac{3\nu+2}{5}\}.$$

Since $\nu \in (1,2)$ only enters in our choice of coordinates, we can freely choose it. In particular, as long as

$$\frac{A^2}{a^2} \le \frac{8}{5},$$

we can choose ν close enough to 2 so that $\frac{\partial}{\partial\epsilon}\frac{\|\nabla f\|^2}{\|\nabla\alpha\|^2} \ge 0$.

Step 3. (Final argument). It simply remains to compare $g(\dot\sigma,\dot\sigma)$ to $g(\dot\gamma,\dot\gamma)$, term by term, according to (5.11). By step 1, we have

$$\begin{aligned}
h_\Theta(\dot\sigma,\dot\sigma) &\le h_\Theta(\dot\gamma,\dot\gamma), \\
h_\Phi(\dot\sigma,\dot\sigma) &\le h_\Phi(\dot\gamma,\dot\gamma).
\end{aligned}$$

Moreover, by computing $\|df\|^2$ in Cartesian coordinates, and using the fact that under the flow of ∇r and $-\nabla\rho$, the Cartesian coordinates are non-increasing (in absolute value), it is easy to check that

$$\|df\|^2(\sigma(s)) \le \|df\|^2(\gamma(s)).$$

By step 2, we have that

$$\frac{\|\nabla f\|^2}{\|\nabla \alpha\|^2}(\sigma(s)) \leq \frac{\|\nabla f\|^2}{\|\nabla \alpha\|^2}(\gamma(s)),$$

and since the ∂_α component of $\dot{\sigma}$ is the same as that of $\dot{\gamma}$ (because $\alpha_{\gamma(s)}$ and $\alpha_{\sigma(s)}$ only differ by a constant),

$$d\alpha(\dot{\sigma}) = d\alpha(\dot{\gamma}).$$

Finally, because $f_{\gamma(s)} = f_{\sigma(s)}$, we have

$$df(\dot{\sigma}) = df(\dot{\gamma}).$$

Plugging all of the above into (5.11), we conclude that

$$g(\dot{\sigma}, \dot{\sigma}) \leq g(\dot{\gamma}, \dot{\gamma}),$$

as desired. □

The following lemma is an easy consequence of Lemma 5.9, and from it we can derive Lemma 5.8.

Lemma 5.10. *Suppose Condition 1 is satisfied, and let $q \in I^-_{\mathcal{M}}(p_c)$. Then there exists a point $\tilde{q} \in J^+_M(q)$ such that $\alpha_{\tilde{q}} = 0$ and $f_{\tilde{q}} < 0$. Equivalently, $\rho_{\tilde{q}} = 0$ and $r_{\tilde{q}} > 0$.*

Proof. The equivalence of the two statements follows simply by definition (5.10). Now for the proof of existence: If $\alpha_q = 0$, choose $\tilde{q} = q$, and we are done because $q \in I^-_{\mathcal{M}}(p_c)$ implies $f_q < f_{p_c} = 0$. Otherwise, choose a f.d. timelike curve $\gamma\colon [0,1] \to \mathcal{M}$ from $\gamma(0) = q$ to $\gamma(1) = p_c$, and write it in components as in (5.12). If $\alpha(1/2) = 0$, choose $\tilde{q} = \gamma(1/2)$, noting that $\gamma(1/2) \in I^-_{\mathcal{M}}(p_c)$ and therefore $f_{\gamma(1/2)} < f_{p_c} = 0$. If $\alpha(1/2) \neq 0$, then since $\gamma(1/2) \in I^+_M(q)$, we can choose $0 < \varepsilon < \alpha(1/2)$ small enough so that $\hat{q} := (f(1/2), \alpha(1/2) - \varepsilon, \Theta(1/2), \Phi(1/2)) \in I^+_M(q)$. Then, by Lemma 5.9, there exists a f.d. timelike curve σ from $\hat{q}$ until some point $\tilde{q} := \sigma(s_\epsilon)$ such that $\alpha_{\tilde{q}} = 0$. Moreover, $f_{\tilde{q}} = f(s_\epsilon) < f(1) = 0$. □

Proof of Lemma 5.8. Let $q \in I^-_{\mathcal{M}}(p_c)$. Then we can choose $\tilde{q} \in J^+_M(q)$ as in Lemma 5.10. We claim that $\tilde{q} \in I^-_{\mathcal{M}}(p_c)$, not only with respect to our metric g, but even with respect to the metric g_{iso} (see Lemma 5.5). To prove this claim, simply note that we can reach p_c from $\tilde{q}$ by following the integral curve of ∇f through $\tilde{q}$ (which has $\rho = 0$ initially, hence $\rho = 0$ on the whole integral curve, while r must decrease, thus we reach p_c). By Lemma 5.19, $\tilde{q} \in I^-_{\mathcal{M}}(p_c, g_{\text{iso}})$ implies that there exists a neighborhood U of p_c such that $U \setminus \{p_c\} \subseteq I^+_M(\tilde{q}, g_{\text{iso}})$. By Lemma 5.5, $I^+_M(\tilde{q}, g_{\text{iso}}) \subseteq I^+_M(\tilde{q})$, and since $\tilde{q} \in J^+_M(q)$, it follows that $U \setminus \{p_c\} \subseteq I^+_M(q)$, as desired. □

5.2.3 Limit curves, push-up and proof of Theorem 5.3

Having established the crucial Lemma 5.8, the rest of the proof of Theorem 5.3 does not require any computations in coordinates. Yet it follows the same philosophy of showing that the causal relation on $\mathcal{M}$ has some of the same (good) properties that it would have on a non-degenerate spacetime.

The next lemma is a sort of limit curve theorem, but can also be interpreted as telling us that $(\mathcal{M}, h, f, \zeta)$ is causally simple (see [BEE96, p. 65] for causal simplicity of non-degenerate spacetimes).

Lemma 5.11 ($\overline{I^+_{\mathcal{M}}} \subseteq J^+_{\mathcal{M}}$)**.** *Suppose Condition 2 is satisfied. Let $(p_i)_i$, $(q_i)_i$ be sequences of points in $\mathcal{M}$ such that $q_i \in I^+_{\mathcal{M}}(p_i)$. If $p_i \to p$ and $q_i \to q$, then $q \in J^+_{\mathcal{M}}(p)$.*

Proof. We first show the case $p_i, q_i, p, q \neq p_c$. Then, by Lemma 5.7, there exists a sequence of f.d. timelike curves $\gamma_i \colon [a_i, b_i] \to M$ such that $\gamma_i(a_i) = p_i$ and $\gamma_i(b_i) = q_i$. The idea is quite simple: we claim that (γ_i), up to a subsequence, converges to a causal curve $\gamma \colon [a, b] \to \mathcal{M}$. We show this by applying the usual limit curve theorem [Min08a, Thm. 3.1] on the spacetime M. It is necessary to distinguish between the case when the limit curve is also in M, and the case when the limit curve crosses over the singular point p_c (then, technically speaking, there are two limit curves in M, which can be joined in $\mathcal{M}$).

- **Case 1:** The sequence γ_i converges uniformly to a causal curve $\gamma \colon [a, b] \to M$, or to a single point. Either way, $q \in J^+_M(p) \subseteq J^+_{\mathcal{M}}(p)$, and we are done.
- **Case 2:** There exist reparametrizations $\gamma_i^p \colon [0, b_i^p) \to M$ of γ_i and a future endless (in M) causal curve $\gamma^p \colon [0, \infty) \to M$ with $\gamma(0) = p$ such that $\gamma_i^p \to \gamma^p$ uniformly on compact subsets. Analogously, there exist reparametrizations $\gamma_i^q \colon (-b_i^q, 0] \to M$ of γ_i and a past endless causal curve $\gamma^q \colon (-\infty, 0] \to M$ with $\gamma^q(0) = y$ such that $\gamma_i^q \to \gamma^q$ uniformly on compact subsets.

In case 2, we claim that $\lim_{t\to\infty} \gamma^p(t) = \lim_{s\to-\infty} \gamma^q(s) = p_c$. This is a direct consequence of Lemma 5.6 and the fact that f is bounded away from $0, 1$ on γ^p and γ^q, hence γ^p, γ^q cannot run into the boundary $\partial\mathcal{M}$. But if $\lim_{t\to\infty} \gamma^p(t) = \lim_{s\to-\infty} \gamma^q(s) = p_c$, then (after suitable reparametrization) we can extend γ^p, γ^q to p_c and concatenate them, forming a causal curve in $\mathcal{M}$ that joins p with q, as desired.

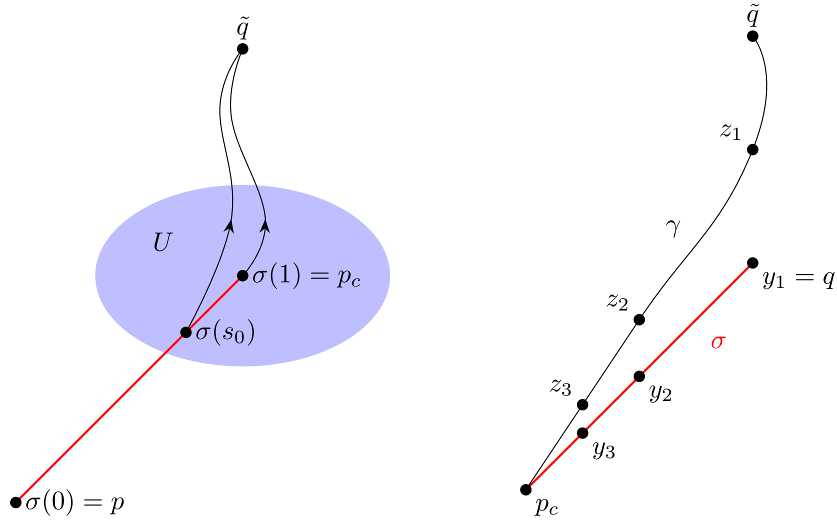
In case that some of p_i, q_i, p, q equal p_c, we can proceed with an analogous proof, but we have to add a third case, where p_c is an endpoint of the limit curve. $\square$

The next lemma is well-known for non-degenerate spacetimes.

Lemma 5.12 (Push-up)**.** *Suppose Condition 2 is satisfied. If $q \in J^+_{\mathcal{M}}(p)$, then $I^+_{\mathcal{M}}(q) \subseteq I^+_{\mathcal{M}}(p)$.*

Proof. If $p = q$, the result is trivial, so assume $p \neq q$, and let $\sigma \colon [0, 1] \to \mathcal{M}$ be a causal curve from p to q.

Case 1: $q = p_c$**.** Let $\tilde{q} \in I^+_M(p_c)$ (see Figure 5.1a). Then, by Condition 2, there exists a neighborhood U of p_c such that $U \backslash \{p_c\} \subseteq I^-_M(\tilde{q})$. Because $\sigma(1) = p_c$, there



(a) Case 1. (b) Case 2.

Figure 5.1: An illustration of the proof of Lemma 5.12. The red line represents σ, and the black curves represent future-directed causal curves.

must exist some $0 < s_0 < 1$ such that $\sigma(s_0) \in U \setminus \{p_c\}$. But then $\tilde{q} \in I^+_M(\sigma(s_0))$, and since also $\sigma(s_0) \in J^+_M(p)$, we conclude by the standard push-up lemma in M that $\tilde{q} \in I^+_M(p) \subseteq I^+_{\mathcal{M}}(p)$. Since $\tilde{q}$ was arbitrary, we are finished with this case.

Case 2: $p = p_c$**.** The argument is similar to the one in [AGH98, Prop. 2.1]. Let $\tilde{q} \in I^+_{\mathcal{M}}(q) = I^+_M(q)$ (see Figure 5.1b). We construct a timelike curve γ from p_c to $\tilde{q}$. Let $y_n := \sigma(1/n)$, and choose a point $z_1 \in I^+_M(y_1) \cap I^-_M(\tilde{q})$. By openness of $I^-_M(z_1)$, and since, by the usual push-up lemma in M, $y_2 \in J^-_M(y_1) \subset I^-_M(z_1)$, we may choose $z_2 \in I^-_M(z_1) \cap I^+_M(y_2) \cap B^h_{1/2}(y_2)$. Here $B^h_{1/2}(y_2)$ denotes the ball of radius $1/2$ around y_2, measured with respect to the Riemannian metric h. Iterating this procedure, we obtain a sequence $(z_l)_l$ such that $z_l \in I^-_M(z_{l-1}) \cap I^+_M(y_l) \cap B^h_{1/l}(y_l)$. Then we construct γ by joining all the timelike segments going from z_l to z_{l+1}. Since, by construction, $\lim_{l\to\infty} z_l = \lim_{l\to\infty} y_l = p_c$, the timelike curve γ connects p_c and $\tilde{q}$.

Case 3: $p, q \neq p_c$**.** If σ lies entirely in M, the result follows from the standard theory. Therefore, we assume w.l.o.g. that $\sigma(\frac{1}{2}) = p_c$. Then, in particular, $q \in J^+_{\mathcal{M}}(p_c)$, so by case 2, we have that $I^+_{\mathcal{M}}(q) \subseteq I^+_{\mathcal{M}}(p_c)$. But since also $p_c \in J^+_{\mathcal{M}}(p)$, by case 1 it follows that $I^+_{\mathcal{M}}(p_c) \subseteq I^+_{\mathcal{M}}(p)$, and we are done. □

Lemma 5.13 below, together with Lemma 5.8, completes the proof of Theorem 5.3.

Lemma 5.13. *If Condition 2 is satisfied, then the Morse spacetime (M, g) is causally continuous.*

Proof. By Definition 5.20 (in Appendix 5.6), (M, g) is causally continuous if it is distinguishing and reflecting. Because f is a time function, (M, g) must be distinguishing [Bor+99, Sec. 2]. Thus we only need to prove reflectivity. Let $p, q \in M$ be such that $I^-_M(p) \subseteq I^-_M(q)$ (the future case is analogous). We need to prove that $I^+_M(q) \subseteq I^+_M(p)$. By the time-reverse of Lemma 5.7, $I^-_{\mathcal{M}}(p) \subseteq I^-_{\mathcal{M}}(q)$, and then, since $p \in \overline{I^-_{\mathcal{M}}(p)} \subseteq \overline{I^-_{\mathcal{M}}(q)}$, Lemma 5.11 tells us that $p \in J^-_{\mathcal{M}}(q)$. But then, by Lemma 5.12, $I^+_{\mathcal{M}}(q) \subseteq I^+_{\mathcal{M}}(p)$, which again by Lemma 5.7 implies $I^+_M(q) \subseteq I^+_M(p)$. □

Remark 5.14. One may even say that the Morse geometry $(\mathcal{M}, h, f, \zeta)$ is globally hyperbolic. Firstly, it follows from Lemma 5.6 that f is a Cauchy time function, in the sense that any causal curve that is inextendible in $\mathcal{M}$, must start at one boundary component and end at the other, crossing each level set exactly once. Secondly, by compactness of $\mathcal{M}$, it is easy to see that the causal diamonds $J^+_{\mathcal{M}}(p) \cap J^-_{\mathcal{M}}(q)$ are compact, for all $p, q \in \mathcal{M}$. However, both arguments are also true when (M, g) is causally discontinuous, such as in the index $1, n-1$ case. The point we would like to make here, is that one should additionally require Condition 2 to hold, and then the causal structure of $\mathcal{M}$ is very well-behaved.

We can turn this remark into a mathematically precise statement by using the language of Lorentzian length spaces, introduced by Kunzinger and Sämann [KS18] (see also [BGH21], where topology change is discussed in this context). Lorentzian length spaces are topological spaces equipped with a notion of causal order and satisfying a set of axioms which, in particular, imply a version of Condition 2 [KS18, Lem. 2.12]. A somewhat related point is that (M, g) is semi-globally hyperbolic, meaning that it can be divided into globally hyperbolic pieces, which in our case are separated by the critical level sets of f. The notion of semi-globally hyperbolic spacetime was introduced by Janssen in [Jan22], with the goal of defining quantum field theories on them (note the connection to Conjecture 5.1).

5.3 Towards a full resolution of the Borde–Sorkin Conjecture

Throughout this section, we employ the same notational conventions as in Section 5.2, except that we allow our Morse functions to have multiple critical points. The current status of Conjecture 5.2 is summarized in Theorem 5.4 in the Introduction. What still remains open is the case when f has critical points of index $\lambda = 2, ..., n-2$, and h is arbitrary. In other words, we do not know what happens if we drop Condition 1.

Notice that Condition 1 is basically telling us two things:

(i) We can find a coordinate neighborhood $\mathcal{U}$ of p_c where both h and f take a specified standard form.

(ii) We have the bounds $\zeta_c < \zeta$ and $\zeta_c \leq 8/5$ (see (5.7)), which can be interpreted as a bound on how much anisotropy is allowed.

In the first part of this section, we show that the neighborhood $\mathcal{U}$ can always be found, the only difference being that in the general case, we need to add a perturbation to h that vanishes at p_c. In the second part of this section, we give a candidate counterexample to Conjecture 5.2, which suggests that $\zeta > \zeta_c$ is a necessary condition for causal continuity. We then conclude by proposing a modified version of the conjecture which takes this into account.

5.3.1 Generalized standard neighborhoods

The statement of the next proposition can be seen as a weaker version of the first part of Condition 1.

Proposition 5.15. *Let $(\mathcal{M}, h, f, \zeta)$ be a Morse geometry and p_c be a critical point of f. Then there exists an open neighborhood $\mathcal{U} \subseteq \mathcal{M}$ of p_c, an open ball $\mathcal{B} \in \mathbb{R}^n$ around the origin, and a coordinate chart $\phi : \mathcal{U} \to \mathcal{B}$ such that $\phi(p_c) = 0$ and*

$$f \circ \phi^{-1} = \frac{1}{2}\sum_i a_i (x^i)^2, \qquad h \circ \phi^{-1} = \sum_i (dx^i)^2 + \sum_{k,l} h^{(1)}_{kl}(x) dx^k dx^l,$$

for some real constants $a_i \neq 0$ and some tensor $h^{(1)}$ satisfying $h^{(1)}(p_c) = 0$.

The proof relies on the following two results from the literature.

Simultaneous Diagonalization Theorem [GH17, Thm. 13.4.3]**.** *Let H, D be two real symmetric $n \times n$ matrices, and let H be positive definite. Then there exists a real non-singular matrix Λ such that both $\Lambda^T H \Lambda$ and $\Lambda^T D \Lambda$ are diagonal.*

Morse Lemma [Mil63, Lem. 2.2]**.** *Let $\mathcal{M}$ be a manifold, $f \colon \mathcal{M} \to \mathbb{R}$ be a Morse function and p_c be a critical point of f. Then there exists an open neighborhood $\mathcal{U} \subseteq \mathcal{M}$ of p_c, an open ball $\mathcal{B} \subseteq \mathbb{R}^n$ around the origin, and a coordinate chart $\phi : \mathcal{U} \to \mathcal{B}$ such that $\phi(p_c) = 0$ and*

$$f \circ \phi^{-1} = \frac{1}{2}\sum_i a_i (x^i)^2,$$

for some real constants $a_i \neq 0$.

Proof of Proposition 5.15. By the Morse Lemma, we can find coordinates where f already has the desired form. Then we apply a linear change of coordinates in order to simultaneously diagonalize the bilinear forms $h(p_c)$ and $\operatorname{hess} f(p_c)$. Finally, we scale each coordinate, in order to normalize our new basis with respect to $h(p_c)$. □

We use Proposition 5.15 to formulate a relaxed version of Condition 1. Let $(\mathcal{M}, h, f, \zeta)$ be a Morse geometry, and suppose we have chosen a critical point p_c.

Condition 3. *The constants a_i appearing in Proposition 5.15 (applied to p_c) satisfy*

$$\zeta > \zeta_c := \max_{i,j} \left| \frac{a_i}{a_j} \right| .$$

We do not include the bound $\zeta_c \leq 8/5$, because it will not be relevant in the upcoming examples, and it seems likely to not be a necessary condition for causal continuity. Note also that in this paper, we take the point of view that ζ is specified as part of the Morse geometry. If, instead, we only specify h and f, then we can always choose $\zeta > \zeta_c$ at a given critical point (hence also at any finite number of critical points). In this sense, Condition 3 is not very restrictive. Note, in any case, that ζ_c depends on both h and f in a neighborhood of p_c.

Since the coordinate system that we get from Proposition 5.15 is not necessarily unique, the question arises whether the truth or falsehood of Condition 3 depends on any coordinate choices. The answer is no. To see this, note that $\max_i a_i$ is the maximum value of the quadratic form hess $f(p_c)$ applied to the $h(p_c)$-unit ball. Similarly, $\min_i a_i$ is the minimum. These maxima and minima are independent of the basis, so we conclude that the value of ζ_c is the same among all coordinate bases satisfying the properties listed in Proposition 5.15.

One is left to wonder how much our proofs in Section 5.2 are affected when adding the perturbation $h^{(1)}$ that appears in Proposition 5.15. We can say the following:

- If Condition 3 is satisfied, then a generalization of Lemma 5.5 holds. The only difference is that in (5.9), we need to replace ζ by $1 < \hat{\zeta} < \zeta$. This makes the lightcones of g_{iso} a bit narrower, compensating for the fact that the perturbation $h^{(1)}$ might have also made the lightcones of g a bit narrower.

- The proofs in Section 5.2.2 are no longer valid.

- If we can prove that Condition 3 implies Condition 2 (compare with Lemma 5.8), then causal continuity follows by the same arguments as in Section 5.2.3.

- On general spacetimes, causal continuity is not stable under perturbations of the metric. This remains true even if we require said perturbations to always widen or narrow the lightcones with respect to the original metric (see Examples 5.21 and 5.22 in Appendix 5.6).

5.3.2 A potential counterexample

The discussion in Section 5.3.1 leads to a natural question: is Condition 3 necessary in order to have causal continuity? If the answer is yes, it would mean that Conjecture 5.2 is false in its original form. We believe that this is indeed so. Constructing examples that violate Condition 3 is easy, but showing that they are causally discontinuous is not (and we do not achieve it in this paper).

The following examples are meant to illustrate what happens when Condition 3 is not satisfied. For convenience, we take $\mathcal{M}$ non-compact and without boundary, the idea being that it represents a neighborhood of a critical point in some larger Morse geometry.

Example 5.16. Let $\mathcal{M}$ be an open ball in $\mathbb{R}^2$, centered around the origin $p_c := 0$, and equipped with coordinates (x, y). Define

$$h := dx^2 + dy^2, \qquad\qquad f := -\frac{1}{2}\left(x^2 + by^2\right),$$

so that

$$g = (x^2 + b^2y^2)(dx^2 + dy^2) - \zeta(xdx + bydy)^2.$$

The case $b = 1$ is considered in Appendix 5.5: it is a *neighborhood spacetime* in the sense of Borde et al. [Bor+99]. We will refer to is as an *isotropic* neighborhood, while in the case of $b \neq 1$, we will talk about *anisotropic* neighborhoods.

Consider a radial line $\gamma(s) = (s, ms)$, for $m \in \mathbb{R}$. Then

$$g\left(\dot\gamma(s), \dot\gamma(s)\right) = \left((1-\zeta)b^2m^4 + (1 + b^2 - 2\zeta b)m^2 + 1 - \zeta\right)s^2. \tag{5.16}$$

If $b = 1$, this quantity reduces to

$$g\left(\dot\gamma(s), \dot\gamma(s)\right) = (1-\zeta)\left(m^2+1\right)^2 s^2,$$

which is negative for all m, hence all radial lines are timelike (see Figure 5.2a). In this case, the past of p_c, which is the whole Morse spacetime $M = \mathcal{M} \setminus \{p_c\}$, can be written as a single TIP, $M = I^-_{\mathcal{M}}(p_c) = I^-_M(\gamma)$, for any future directed timelike curve γ that ends at p_c (see Appendix 5.5).

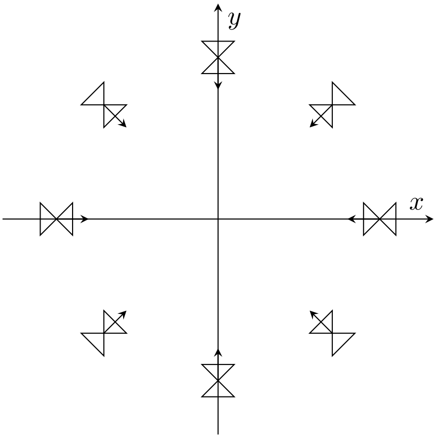
Taking the limit $b \to 0$ in (5.16), we get an expression that is positive whenever $m^2 > 1 - \zeta$. By continuity, we conclude that for $0 < b < 1$ small enough, $g(\dot\gamma(s), \dot\gamma(s))$ can be negative, zero or positive, depending on m (the dependence, however, is more complicated than in the $b \to 0$ limit). Concretely, for b small (relative to ζ):

- There exist null radial lines, which form the boundaries of the future sets[3] $\mathcal{F}_1, \mathcal{F}_2$ and the pasts sets $\mathcal{P}_1, \mathcal{P}_2$. Intuitively, this happens because the lightcones tilt much faster when moving in the x direction, compared to moving in the y direction (see Figure 5.2b). Regardless, we still have that $M = I^-_{\mathcal{M}}(p_c)$.

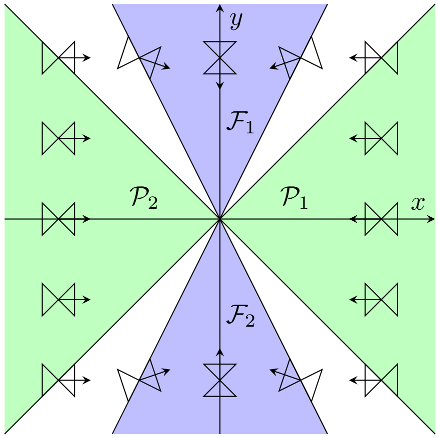
- $I^+_{\mathcal{M}}(q)$ is not open, for any $q \in M$. This is because $p_c \in I^+_{\mathcal{M}}(q)$, but no neighborhood of p_c is entirely contained in $I^+_{\mathcal{M}}(q)$ (compare with Lemma 5.8).

Note that there also exists an intermediate case, when b has the exact value so that setting (5.16) equal to zero has degenerate solutions, and then the boundaries of the sets $\mathcal{F}_1, \mathcal{F}_2, \mathcal{P}_1, \mathcal{P}_2$ overlap. The case of $b > 1$ very large can be reduced to the case of $b < 1$ small by rescaling both x and y by a factor of $1/b$ (then also h is rescaled, but this does not affect g).

Since the critical point in this example has index 0, the Morse spacetime (M, g) is causally continuous, no matter how small we choose b (see Theorem 5.4). In fact, (M, g) is even globally hyperbolic, with f being a Cauchy time function. Note however that in this example, $I^+_{\mathcal{M}}(p_c) = \emptyset$, which simplifies things a lot.



(a) $b = 1$



(b) $b \ll 1$

Figure 5.2: The causal structure of Example 5.16.

Building upon the previous example, we propose our candidate counterexample to Conjecture 5.2.

Example 5.17. Let $\mathcal{M}$ be an open ball in $\mathbb{R}^4$, centered around the origin $p_c := 0$, and equipped with coordinates (x, y, z, w). Define

$$h := dx^2 + dy^2 + dz^2 + dw^2, \qquad f := \frac{1}{2}\left(-x^2 - by^2 + z^2 + w^2\right),$$

and g, as usual, by (5.3). We claim that reflectivity (see Definition 5.20) is violated by the pair of points

$$p := (1, 0, 0, 0), \qquad q := (m, 0, 1, 0),$$

where $m := \frac{\sqrt{\zeta}-1}{\sqrt{\zeta-1}}$. From the causal analysis of the punctured (x, z)-plane (see Figure 5.3 and Appendix 5.5) it follows that $I^+_M(q) \subseteq I^+_M(p)$. However, $I^-_{XZ}(q) \not\supseteq I^-_{XZ}(p)$, where the subscript XZ means that we are only considering causal curves in the punctured (x, z)-plane. Nonetheless, it is possible that $I^-_M(q) \supseteq I^-_M(p)$ when also considering causal curves that leave said plane. In particular, if $b = 1$ (or close enough to 1), we see from the analysis in Example 5.16 (see also Figure 5.4b) that from p we can reach the negative x-axis via a future-directed timelike curve contained in the (x, y)-plane. Then, from the negative x-axis, we can reach q. If b is too small, however, this construction is no longer possible (see Figure 5.2b), suggesting that probably $I^-_M(q) \not\supseteq I^-_M(p)$. This is not a bulletproof argument, of course, because we are ignoring all timelike curves that are not contained in any coordinate plane.

[3] Here by *future set* we mean a set $\mathcal{F}$ such that $I^+_M(\mathcal{F}) = \mathcal{F}$. Analogously, a *past set* $\mathcal{P}$ satisfies $I^-_M(\mathcal{P}) = \mathcal{P}$.

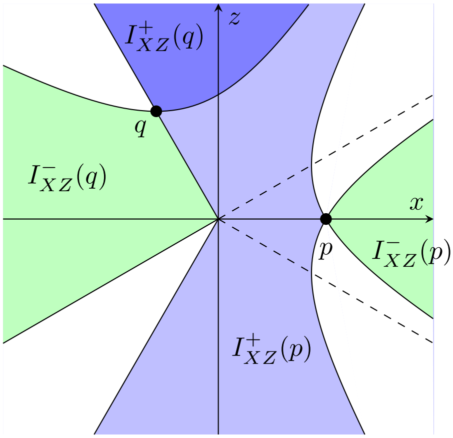


Figure 5.3: The points p, q of Example 5.17 and their futures and pasts (restricted to the (x, z)-plane XZ).

5.4 Conclusions

With the proof of Theorem 5.3, we have established a new case of the Borde–Sorkin conjecture (Conjecture 5.2 in the Introduction). In doing so, we have advanced the current status of the conjecture to that summarized in Theorem 5.4. Along the way, we have developed a notion of causal structure for Morse geometries that includes the critical points. This supports the view that degenerate metrics are physically reasonable, and is a first step towards understanding quantum fields on Morse geometries (in view of Conjecture 5.1).

Let us briefly mention here three recent approaches to quantum field theory that are specially relevant for topology change. The algebraic approach of Janssen [Jan22] has been developed specifically with topology change as one of its applications, but the existence of states in this approach is still an open problem. Another approach is that of Sorkin and Johnston [Joh09; Sor18]; so far it has been applied to the trousers topology change [Buc+17] (confirming the energy divergences there), but, to our knowledge, not to any of the causally continuous Morse geometries (where Conjecture 5.1 predicts a well-behaved QFT). Lastly, the recent paper of Kontsevich and Segal [KS21] defines QFTs on the category of cobordisms with certain complex metrics. Real Lorentzian metrics arise as a limit case of these complex metrics, and so do Morse metrics [LS97; Wit22]. So far, however, it is only known that a QFT is induced on the limit spacetime when the latter is non-degenerate and globally hyperbolic [KS21, Thm. 5.2]. It remains to be seen if this result can be generalized to Morse geometries.

Another important conclusion of the present paper is that we have found a potential counterexample to Conjecture 5.2 (Example 5.17 in Section 5.3.2). Despite being of dimension 4 and having only a critical point of index 2, we believe that our example is causally discontinuous, due to being highly anisotropic (i.e.

because it has $\zeta_c > \zeta$). The lack of symmetries and good coordinate choices has prevented us from proving this fully. Regardless, we propose the following refinement of Conjecture 5.2, which incorporates a bound on the anisotropy (Condition 3).

Conjecture 5.18. *Let $(\mathcal{M}, h, f, \zeta)$ be a Morse geometry of dimension n, and (M, g) the corresponding Morse spacetime. Assume f has a single critical point for each critical value. Then (M, g) is causally continuous if and only if the following hold:*

(i) None of the critical points has index 1 or $n-1$,

(ii) Condition 3 is satisfied at every critical point of index different from $0, n$.

In order to prove Conjecture 5.18, two steps remain. One is to show that Condition 3 is really necessary, by showing that Example 5.17 (where it is violated) is causally discontinuous. The other remaining step is to generalize Theorem 5.3 by adding a perturbation to h that vanishes at the critical points, and by removing the requirement that $\zeta_c \leq 8/5$. This is nontrivial, because causal continuity is, in general, not stable under perturbations (see Examples 5.21 and 5.22 in Appendix 5.6). Yet the second half of our proof (Section 5.2.3) is robust under perturbations, and does not require $\zeta_c \leq 8/5$, so it would suffice to prove openness of the chronological relation in $\mathcal{M}$ (Condition 2), and then causal continuity would follow.

5.5 Appendix: Neighborhood spacetimes

In this appendix to Chapter 5, we review the causal structure of isotropic neighborhood spacetimes, as studied in Borde et al. [Bor+99]. At the end, we prove Lemma 5.19, which is new, although it follows quite straightforwardly from the analysis in [Bor+99].

Let $\mathcal{M} \subseteq \mathbb{R}^n$ be an open neighborhood of the origin, equipped with a coordinate system $(x^1, ..., x^\lambda, y^1, ..., y^{n-\lambda})$, where $\lambda \neq 0, 1, n-1, n$, and

$$h := \sum_{i=1}^{\lambda} (dx^i)^2 + \sum_{j=1}^{n-\lambda} (dy^j)^2, \qquad f := -\frac{1}{2}\sum_{i=1}^{\lambda} (x^i)^2 + \frac{1}{2}\sum_{j=1}^{n-\lambda} (y^j)^2. \tag{5.17}$$

Note that the origin $p_c = (0, ..., 0)$, is the only critical point of f in this case. We write $M := \mathcal{M} \setminus \{p_c\}$, as usual (here $\mathcal{M}$ has no boundary, but can be thought of as a neighborhood of a critical point in some cobordism). It is convenient to change to polar coordinates (ρ, Θ, r, Φ), where

$$\rho := \sum_{j=1}^{n-\lambda} (y^j)^2, \qquad\qquad r := \sum_{i=1}^{\lambda} (x^i)^2,$$

and where Θ, Φ denote the angular coordinates corresponding to the subspaces $\{r = 0\}$ and $\{\rho = 0\}$ respectively (thus each of Θ, Φ is a collection of angular variables, rather than a single one). The Lorentzian metric (5.2) is then given by

$$\begin{aligned} g = &\left(r^2 - (\zeta-1)\rho^2\right) d\rho^2 + \left(\rho^2 - (\zeta-1)r^2\right) dr^2 + 2\zeta\rho r d\rho dr \\ &+ \left(\rho^2 + r^2\right)\left(\rho^2 d\Theta^2 + r^2 d\Phi^2\right). \end{aligned} \tag{5.18}$$

Because the coefficients in front of $d\Theta$ and $d\Phi$ are positive, any null geodesic in the $\Theta, \Phi =$ const. plane (with respect to the restricted metric), must also be a null geodesic in the full spacetime.

Thus we start by commenting on the situation with constant angles. Recall also that in 2 dimensions, any null curve is automatically a null geodesic. The implicit equations for any null geodesic can thus be found from (5.18):

$$\sqrt{\zeta-1}(x^2-y^2) = \pm 2xy + c_\pm, \tag{5.19}$$

where $c_\pm$ are constants. In particular, the case of $c_\pm = 0$ corresponds to geodesics that bound $I^\pm_{\mathcal{M}}(p_c)$, which are radial lines of a certain slope (depicted as dashed lines in Figure 5.4a). We can use this information to find the future and past sets of any point (depicted as colored regions in Figure 5.4a).

Next we analyse the case when $\rho = 0$ and $\Theta =$ const. (the case $r = 0$ and $\Phi =$ const. is analogous). We restrict to the case where $\Phi = \phi$ is just a single angular variable, hence again reducing the problem to two dimensions. The null geodesics on the $\rho = 0, \Theta =$ const. plane (with respect to the induced metric) are then given by

$$r(\phi) = r_0\, \mathrm{e}^{\pm\phi/\sqrt{\zeta-1}}.$$

Again, we can use this to find the future and past sets of any point on the plane, with respect to the induced metric (see Figure 5.4b). In this case, there are no null geodesics going through the origin p_c, and all points with $\rho = 0$ lie in the past of p_c.

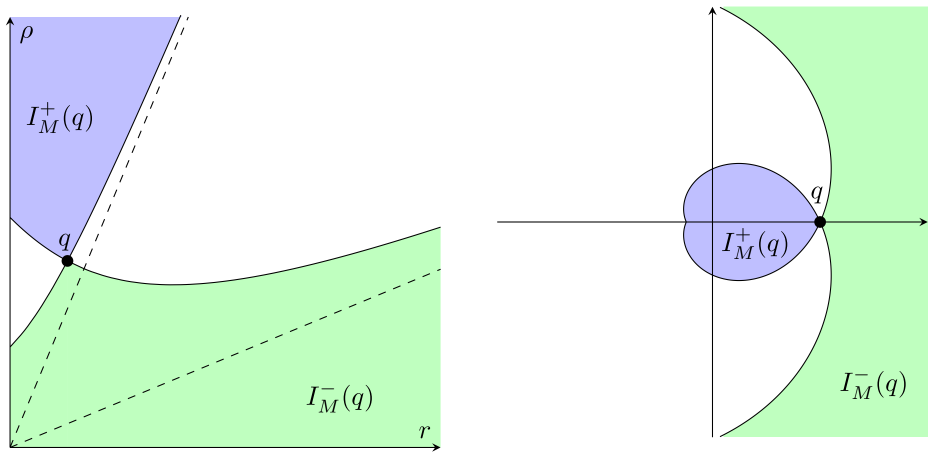
Lemma 5.19. *Let $(\mathcal{M}, h, f, \zeta)$ be an isotropic Morse neighborhood, with h, f given by* (5.17)*. For every $q \in I^\pm_{\mathcal{M}}(p_c)$, there exists a neighborhood U of p_c such that $U \setminus \{p_c\} \subseteq I^\mp_M(q)$.*

Proof. We show the case when $q \in I^-_{\mathcal{M}}(p_c)$. W.l.o.g. we may choose our coordinates such that $q = (\rho_q, 0, r_q, 0)$, where necessarily $\rho_q \neq 0$ (otherwise q cannot be in $I^-_{\mathcal{M}}(p_c)$). We want to find ρ_0, r_0 such that all points $(\rho, \Theta, r, \Phi) \neq p_c$ with $\rho < \rho_0$, $r < r_0$ and Θ, Φ arbitrary, are contained in $I^+_M(q)$. By symmetry, we may choose our coordinates such that at most one of the Θ- and one of the Φ-angles may be different from zero, hence effectively reducing the problem to four dimensions.

Our argument now resembles the one in the proof of [Bor+99, Claim 1]. Let $\ll$ denote the chronological relation in M. If $\rho_q, r_q \neq 0$, then

$$x = (\rho_x, 0, r_x, 0) \ll (\epsilon_1, 0, 0, 0) \ll (\epsilon_1\delta, \theta, 0, 0) \ll (\epsilon_1\delta\epsilon_2, \theta, \epsilon_3, \phi).$$

In every step where we have added an ϵ, we have used our analysis of the causal structure in the case θ, ϕ constant. In the step where we have added δ, it is using our analysis of the $\rho = 0$ and $\theta =$ const. case. In principle, δ depends on θ. We



(a) (r, ρ)-plane

(b) (r, ϕ)-plane

Figure 5.4: The causal structure of an isotropic neighborhood.

see, however, that the "worst case scenario" (when δ has to be the smallest) is when $\theta = \pi$. Thus we can choose this largest value, so that the procedure works in all cases. Note also that in the last step, since we start from the origin of the (r, ϕ)-plane, we can choose any value for ϕ that we want. Setting $\rho_0 := \epsilon_1 \delta \epsilon_2$ and $r_0 := \epsilon_3$, and again considering the causal structure in the case θ, ϕ constant, we are done. □

5.6 Appendix: Causal continuity

Let (M, g) be a non-degenerate spacetime. We refer to [BEE96, Chap. 3] for the basic concepts and notation of causality theory. The idea is that (M, g) is causally continuous if the set valued functions $q \mapsto I_M^{\pm}(q)$ are continuous. There are various equivalent ways to make this precise [BEE96, pp. 59-71]. In this paper, we use the following definition, which is perhaps the most standard one, even though it does not directly capture the intuition behind the concept.

Definition 5.20. A spacetime (M, g) is called

(i) *distinguishing* if

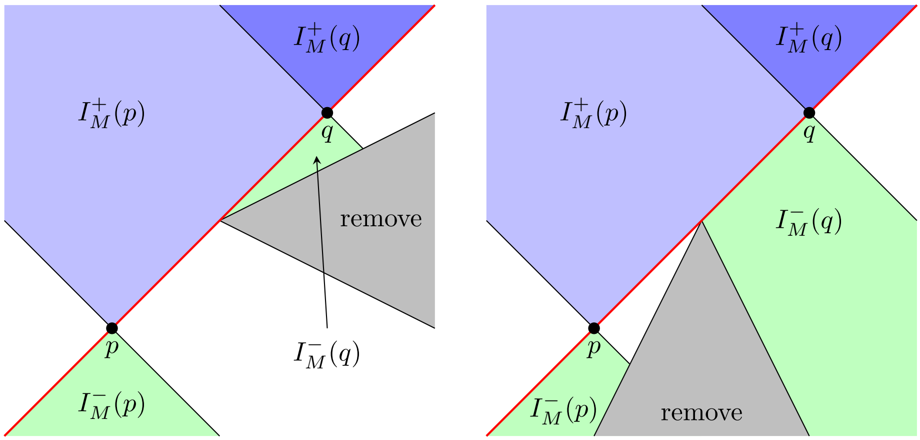
$$I_M^-(p) = I_M^-(q) \iff p = q \iff I_M^+(p) = I_M^+(q)$$

for all $p, q \in M$,

(ii) *reflecting* if

$$I_M^-(p) \subseteq I_M^-(q) \iff I_M^+(p) \supseteq I_M^+(q)$$

for all $p, q \in M$,



(a) Example 5.21 with $\alpha = 1$. (b) Example 5.22 with $\alpha = 1$.

Figure 5.5: A pair of points p, q for which reflectivity is violated.

(iii) *causally continuous* if it is distinguishing and reflecting.

The following example shows that causal continuity is not stable under perturbations of the metric g, even if we only allow perturbations that make the lightcones narrower.

Example 5.21. Let $M := \mathbb{R}^2 \setminus \{(x,t) \mid x \geq 2|t|\}$ and $g_\alpha = -\alpha dt^2 + dx^2$. Then (M, g_α) is causally continuous for $\alpha \geq 2$ and causally discontinuous for $\alpha < 2$. This can be seen in Figure 5.5a: Reflectivity is violated for pairs of points lying on the diagonal red line, one above and one below the origin (such as the depicted points p, q). The red line has slope $1/\alpha$, hence if $\alpha \geq 2$, half of the red line lies inside the removed wedge, and there is no violation of reflectivity anymore.

The next example shows that causal continuity is not stable under widening of the lightcones, either.

Example 5.22. Let $M := \mathbb{R}^2 \setminus \{(x,t) \mid t \leq -2|x|\}$ and $g_\alpha = -\alpha dt^2 + dx^2$. Then (M, g_α) is causally continuous for $\alpha \leq \frac{1}{2}$ and causally discontinuous for $\alpha > \frac{1}{2}$. The argument is similar to the one in Example 5.21 (see Figure 5.5b).

Chapter 6

Global hyperbolicity through the eyes of the null distance

This chapter is based on the preprint [BGH22], written in collaboration with Annegret Burtscher and submitted for publication.[1]

6.1 Introduction

The notion of global hyperbolicity was introduced by Leray [Ler53] in 1952 to prove the uniqueness of solutions for hyperbolic partial differential equations. Shortly thereafter, global hyperbolicity entered the field of General Relativity through the proof of the global well-posedness of the Einstein equations of Choquet–Bruhat and Geroch [CBG69; FB52] and the Singularity Theorems of Penrose and Hawking [HE73; Pen65]. Via the topological splitting result of Geroch [Ger70] globally hyperbolic spacetimes manifestly settled in Lorentzian Geometry in 1970.

Spacetimes are time-oriented Lorentzian manifolds (M, g). They are the geometric objects needed for formulating gravitation in General Relativity. Throughout this manuscript, we use the sign convention $(-, +, \dots, +)$ for g. Global hyperbolicity establishes a deep link between the topology of M and the causal structure induced by the metric tensor g. The causal structure is induced on M by causal curves, i.e., locally Lipschitz (with respect to any Riemannian metric [Chr20, Sec. 2.3], [Bur15, Thm. 4.5]) curves γ with $g(\dot\gamma, \dot\gamma) \leq 0$, as follows. If q can be reached by a future-directed causal curve from p we say that q is in the causal future of p, and write $q \in J^+(p)$ (dually for the causal past $J^-(p)$) or $(p, q) \in J^+$. Leray's original definition of global hyperbolicity was based on the C^0-compactness of the set of causal curves between any two points in a spacetime. The modern definition

[1]This manuscript was completed during an extended research stay at the Fields Institute for Research in Mathematical Sciences in Toronto, in connection with the thematic program "Nonsmooth Riemannian and Lorentzian Geometry" that took place in the Fall of 2022. Both authors gratefully acknowledge funding by the Fields Institute during their stay. AB's research is in part also supported by the Dutch Research Council (NWO), Project number VI.Veni.192.208.

of global hyperbolicity requires compactness of causal diamonds $J^+(p) \cap J^-(q)$ akin to the Heine–Borel property for complete Riemannian manifolds (see [HE73, Sec. 6.6] and [BS07]).

Definition 6.1. A spacetime (M, g) is called *globally hyperbolic* if it is *causal* (there is no closed causal curve) and all causal diamonds $J^+(p) \cap J^-(q)$, $p, q \in M$, are compact.

If (M, g) is a noncompact spacetime of dimension greater than 2 then the causal condition can be dropped [HM19].

A landmark result concerning global hyperbolicity is Geroch's Topological Splitting Theorem [Ger70], later promoted to a smooth orthogonal splitting by Bernal and Sánchez [BS05]. It states that any globally hyperbolic spacetime (M, g) admits a *Cauchy orthogonal splitting*, i.e., an isometry

$$(M, g) \cong (\mathbb{R} \times \Sigma, -\alpha d\tau^2 + \bar{g}_\tau), \tag{6.1}$$

for $\alpha \colon \mathbb{R} \times \Sigma \to (0, \infty)$ a smooth function and $(\bar{g}_\tau)_\tau$ identifiable with a family of Riemannian metrics on the Cauchy slices $\{\tau\} \times \Sigma$, smoothly varying in τ. The proof of this splitting result is rooted in the construction of a suitable time function τ.

Definition 6.2. Let (M, g) be a spacetime. A continuous function $\tau \colon M \to \mathbb{R}$ is said to be a *time function* if

$$q \in J^+(p) \setminus \{p\} \Longrightarrow \tau(p) < \tau(q).$$

Every *stably causal* spacetime admits a (non-unique) time function [Haw68; Min09; Min10]. The class of globally hyperbolic spacetimes can conveniently be characterized by the special type of time functions they admit.

Theorem 6.3 (Geroch [Ger70], Bernal–Sánchez [BS05])**.** *A spacetime (M, g) is globally hyperbolic if and only if there exists a* Cauchy time function τ *on M, meaning that each of its level sets $\tau^{-1}(s)$, $s \in \mathbb{R}$, is a* Cauchy surface*, i.e., intersected (exactly once) by every inextendible causal curve.*

From the compactness condition in Definition 6.1 it follows by a result obtained independently by Avez [Ave63] and Seifert [Sei77] that there exists a length-maximizing geodesic between any two causally related points (length-maximizing with respect to the Lorentzian distance, which is then also finite-valued and continuous [BEE96, Ch. 4]). In that sense globally hyperbolic spacetimes again resemble complete Riemannian manifolds. But here the analogy ends. The Lorentzian distance is far from inducing a metric space structure. Even more troubled is the relationship with geodesic completeness. Neither does global hyperbolicity imply geodesic completeness (the famous Penrose Singularity Theorem [Pen65] actually shows incompleteness under additional curvature bounds) nor the other way round (anti-de Sitter space). In both cases these are actually features rather than bugs of Lorentzian manifolds, and physically highly desired, for instance, for the mathematical existence of black holes. Nonetheless, even the physically undesired assumption of compactness does not guarantee geodesic completeness (Clifton–Pohl torus). Altogether these properties render any Hopf–Rinow type statement for spacetimes virtually a lost cause (see the early works of Busemann [Bus67], Beem [Bee76]; and [BE85; CY76; Eg16; Har88] for work on completeness of spacelike submanifolds). We reopen the case and characterize global hyperbolicity in a profoundly new way.

Theorem 6.4. *A spacetime (M, g) is globally hyperbolic if and only if there exists a time function τ such that $(M, \hat{d}_\tau)$ is a complete metric space.*

Here $\hat{d}_\tau$ is the null distance of Sormani and Vega [SV16], defined in 2016 with the purpose of studying geometric stability problems in General Relativity by means of a metric (measure) convergence theory, and to develop robust tools for spacetimes of low regularity (partly already realized in [AB22; BGH21; KS22; SS23; SV16; Veg21]). The $\hat{}$ in $\hat{d}_\tau$ indicates the dependence on the causal cone structure and τ the link to the time function.

Definition 6.5. Let (M, g) be a spacetime with time function τ. A *piecewise causal path* $\beta\colon [a, b] \to M$ is given by a partition $a = s_0 < s_1 < \dots s_{k-1} < s_k = b$ on which each restriction $\beta|_{[s_{i-1}, s_i]}$ is either a future- or past-directed causal curve. The *null length* of β is given by

$$\hat{L}_\tau(\beta) := \sum_{i=1}^{k} |\tau(\beta(s_i)) - \tau(\beta(s_{i-1}))|.$$

The *null distance* between two points $p, q \in M$ is

$$\hat{d}_\tau(p, q) := \inf\{\hat{L}_\tau(\beta) \mid \beta \text{ piecewise causal path between } p \text{ and } q\}.$$

Clearly, $\hat{d}_\tau$ is symmetric and satisfies the triangle inequality, but positive definiteness does not hold for all τ. For locally anti-Lipschitz time functions one indeed obtains a conformally invariant length-metric space $(M, \hat{d}_\tau)$ that induces the manifold topology (see [SV16, Thm. 4.6] and [AB22, Thm. 1.1]). Throughout most of this manuscript we will assume slightly more, namely that τ is a (weak) temporal function.

Definition 6.6. Let (M, g) be a spacetime with time function $\tau\colon M \to \mathbb{R}$. Let h be any Riemannian metric on M and d_h the associated distance function. If for every point x there exists a neighborhood U of x and $C \geq 1$ such that

$$(p, q) \in J^+ \cap (U \times U) \Longrightarrow \frac{1}{C} d_h(p, q) \leq \tau(q) - \tau(p) \leq C d_h(p, q), \tag{6.2}$$

then we say that τ is a *weak temporal function.* If τ satisfies only the first $\leq$ in (6.2) it is called *locally anti-Lipschitz*, if only the second $\leq$ it is called *locally Lipschitz.*

Based on our extension of (6.2) to an entire open set in Section 6.2 we show that weak temporal functions are indeed locally Lipschitz (in the usual sense) and have a timelike gradient almost everywhere. The standard (smooth) temporal functions, and also regular cosmological time functions à la Andersson–Galloway–Howard [AGH98] and Wald–Yip [WY81], are weak temporal functions. In fact, working with temporal functions is no restriction since every smooth spacetime that admits a time function also admits a temporal function [BS05, Thm. 1.2]. The true advantage of (smooth) temporal functions over other time functions is the orthogonal decomposition $g = -\alpha d\tau^2 + \bar{g}_\tau$ (although no product splitting of the manifold, see [MS11, Lem. 3.5]). This property is key in the proof of the subsequent results. Time functions are, however, also a useful tool in weaker nonsmooth geometric settings (see, for instance, [Bor+99; BGH21; KS22]) and also regular cosmological time functions are, in general, not smooth even on smooth spacetimes. Therefore, we decided to prove our results for the optimal regularity class.

Theorem 6.7. *Let (M, g) be a spacetime, $\tau\colon M \to \mathbb{R}$ be a weak temporal function, and h be a Riemannian metric on M. Then, for each compact set $K \subseteq M$, there exists a constant $C \geq 1$ such that for all $p, q \in K$,*

$$\frac{1}{C} d_h(p,q) \leq \hat{d}_\tau(p,q) \leq C d_h(p,q). \tag{6.3}$$

Note that the lower bound follows from the locally anti-Lipschitz property of τ, and the upper bound from the corresponding locally Lipschitz bound. Theorem 6.7 immediately implies that two stably causal spacetimes are metrically equivalent on compact sets in the following sense.

Corollary 6.8. *Let M be a smooth manifold and g and $\tilde{g}$ be spacetime metrics on M with weak temporal functions τ and $\tilde{\tau}$ (and corresponding null distances $\hat{d}_\tau$ and $\hat{d}_{\tilde{\tau}}$), respectively. Then, for each compact set $K \subseteq M$, there exists a constant $C \geq 1$ such that for all $p, q \in K$,*

$$\frac{1}{C} \hat{d}_\tau(p,q) \leq \hat{d}_{\tilde{\tau}}(p,q) \leq C \hat{d}_\tau(p,q).$$

Amongst others, the proof of Theorem 6.7 requires semiglobal techniques to go from open sets to compact sets. On a global scale the situation is even more involved and it is here where globally hyperbolic spacetimes really shine. We prove the following.

Theorem 6.9. *Let (M, g) be a globally hyperbolic spacetime and τ a locally anti-Lipschitz time function such that all nonempty level sets are Cauchy. Then the null distance* encodes causality*, that is, for any $p, q \in M$,*

$$q \in J^+(p) \Longleftrightarrow \hat{d}_\tau(p,q) = \tau(q) - \tau(p). \tag{6.4}$$

The $\Longrightarrow$ direction in Theorem 6.9 is trivial. Sormani and Vega [SV16, Thm. 3.25] showed that the converse holds for warped product spacetimes with complete Riemannian fiber and suitable temporal functions. It remained an open problem to determine under which general circumstances causality is encoded. Our Theorem 6.9 provides a sharp answer both in terms of regularity as well as the causality class (see counterexamples in Section 6.3.3). Initially we proved this result for Cauchy temporal functions in Theorem 6.25. Independently and simultaneously, Sakovich and Sormani [SS23, Thm. 4.1] have obtained a different global causality encoding result where they allow for general anti-Lipschitz time functions, but require them to be proper. This properness assumption, in fact, implies that the spacetime must be globally hyperbolic with *compact* Cauchy level sets (see Section 6.3.2). Both approaches yield a *local* encoding of causality on any stably causal spacetime (see Theorem 6.26 and [SS23, Thm. 1.1]). Upon studying the proofs of [SS23] we noticed that by combining part of their local arguments [SS23, Thm. 1.1] with our global proof of Theorem 6.25 we can obtain Theorem 6.9 which is optimal both in view of regularity as well as causality. It is precisely this optimality, together with the observation that τ having future (or past) Cauchy level sets is actually sufficient, that allows us to conclude with the following application.

Corollary 6.10. *Let (M, g) be a spacetime that admits a regular cosmological time function τ. Then $\hat{d}_\tau$ encodes causality globally.*

The manuscript is structured as follows. In Section 6.2 we prove Theorem 6.7 and Corollary 6.8. In the proof we make use of the orthogonal decomposition of the spacetime metric with respect to a temporal function and techniques developed in [Bur15]. We also present counterexamples that show that the weak temporal condition cannot be relaxed. In Section 6.3 we prove Theorem 6.9 and provide counterexamples for non-Cauchy locally anti-Lipschitz functions. Nonetheless, a local result for any temporal function on any stably causal spacetime is also obtained. In Section 6.4 we prove Theorem 6.4 and show that completely uniform temporal functions (very recently introduced in [BS18; BS20]) guarantee completeness of $(M, \hat{d}_\tau)$.

6.2 Bi-Lipschitz bounds

In this section we prove Theorem 6.7 and Corollary 6.8 of the introduction.

Since weak temporal functions satisfy $\hat{d}_\tau(p,q) = \tau(q) - \tau(p)$ for causally related points, the condition (6.2) can be viewed as a restricted local metric equivalence. In order to extend this property to a true local metric equivalence we make use of several technical results related to temporal functions obtained in Section 6.2.1, some of which are also used in Section 6.3. To extend the corresponding local result to weak temporal functions and to compact sets we employ semiglobal techniques similar to those in [Bur15]. These final steps of the proof of Theorem 6.7 are carried out in Section 6.2.2.

Theorem 6.7 implies that we can compare the null distances of any two spacetime metrics with entirely different causal cones as stated in Corollary 6.8 of the introduction. In particular, the null distance structures with respect to weak temporal functions are equivalent on compact sets of a fixed spacetime.

Corollary 6.11. *Let τ_1, τ_2 be two weak temporal functions on a spacetime (M, g), and $\hat{d}_{\tau_1}$, $\hat{d}_{\tau_2}$ their associated null distances. Then, for every compact set $K \subseteq M$,* *there exists a constant $C \geq 1$ such that for all $p, q \in K$,*

$$\frac{1}{C}\hat{d}_{\tau_1}(p,q) \leq \hat{d}_{\tau_2}(p,q) \leq C\hat{d}_{\tau_1}(p,q). \tag{6.5}$$

Note that Corollary 6.11 was already announced in [AB22, p. 7739] with an alternative direct proof. The advantage of such a proof is that it does not require the use of temporal functions which, for instance, do not exist in theory of Lorentzian length spaces [BGH21] (while weak temporal functions only require local Lipschitz conditions and can still be considered). We therefore provide this short alternative proof of Corollary 6.11 in Section 6.2.3.

The question remains how optimal all the results just mentioned are with respect to the regularity class of time functions considered. The lower bound in (6.2) is the standard local anti-Lipschitz assumption on τ and needed to even obtain a sensible metric space $(M, \hat{d}_\tau)$ [SV16, Thm. 4.6]. In Section 6.2.4 we show that the locally Lipschitz assumption (the upper bound in (6.2)) cannot be dropped either. Besides, both assumptions are necessary to prove that weak temporal functions have a timelike gradient almost everywhere and are thus the right generalization of temporal functions (Section 6.2.5). Moreover, we show that even under the best circumstances, a general global version of our weakest result presented in this section, namely Corollary 6.11, cannot be expected. This implies that also Theorem 6.7 and Corollary 6.8 in general do not hold globally.

6.2.1 Temporal functions and Wick-rotated metrics

The aim of this section is to compare the null distance to the distance obtained with respect to the Wick-rotated metric that exists for a temporal function. Recall that a spacetime admits a smooth temporal function whenever it admits a time function by a well-known result of Bernal and Sánchez [BS05, Thm. 1.2].

Definition 6.12. Let (M, g) be a spacetime. A *temporal function* is a smooth function $\tau\colon M \to \mathbb{R}$ with past-directed timelike gradient $\nabla\tau$.

In the local proofs in this section we make use of a weaker splitting result for temporal functions (compare to (6.1) in the globally hyperbolic case).

Lemma 6.13 (Müller–Sánchez [MS11, Lem. 3.5])**.** *If a spacetime (M, g) admits a temporal function τ, then the metric g admits an orthogonal decomposition*

$$g = -\alpha d\tau^2 + \bar{g}, \tag{6.6}$$

where $\alpha = |g(\nabla\tau, \nabla\tau)|^{-1} > 0$ and $\bar{g}$ is a symmetric 2-tensor which vanishes on $\nabla\tau$ and is positive definite on the complement.

The temporal function τ is said to be *steep* if $g(\nabla\tau, \nabla\tau) \leq -1$, and hence $\alpha = |g(\nabla\tau, \nabla\tau)|^{-1} \leq 1$. Not every (even causally simple) spacetime admits a steep temporal function [MS11, Thm. 1.1 and Ex. 3.3], but we can always rewrite (6.6) as $g = -\alpha(d\tau^2 + \tilde{g})$, for $\tilde{g} = \alpha^{-1}\bar{g}$. Since the null distance is conformally invariant we assume from now on, without loss of generality, that g is of the form

$$g = -d\tau^2 + \bar{g}. \tag{6.7}$$

Next we perform a standard trick to obtain a Riemannian metric g_W from g and τ. This technique is called Wick-rotation in the physics literature. The *Wick-rotated metric* g_W is given by

$$g_W := d\tau^2 + \bar{g}. \tag{6.8}$$

We denote the associated norm, length functional, and Riemannian distance by $\|\cdot\|_W$, L_W, and d_W respectively. We proceed to compare d_W and $\hat{d}_\tau$.

Lemma 6.14. *Let (M, g) be a spacetime equipped with a smooth temporal function τ, so that g and g_W are given by* (6.7) *and* (6.8)*, respectively. Let $\beta\colon [a, b] \to M$ be a piecewise causal curve. Then*

$$\hat{L}_\tau(\beta) \leq L_W(\beta) \leq \sqrt{2}\hat{L}_\tau(\beta),$$

and thus for any $p, q \in M$,

$$d_W(p, q) \leq \sqrt{2}\hat{d}_\tau(p, q). \tag{6.9}$$

Proof. By assumption τ is smooth and also the tangent vector $\dot{\beta}$ exists almost everywhere, thus we can write both length functionals in terms of the integrals

$$\begin{aligned}\hat{L}_\tau(\beta) &= \int_a^b |(\tau \circ \beta)'(s)| ds,\\ L_W(\beta) &= \int_a^b \|\dot{\beta}(s)\|_W ds = \int_a^b \sqrt{|d\tau(\dot{\beta}(s))|^2 + \|\dot{\beta}(s)\|_{\bar{g}}^2} ds.\end{aligned}$$

Since $\dot{\beta}^\tau(s) := (\tau \circ \beta)'(s) = d\tau_{\beta(s)}(\dot{\beta}(s))$ and the curve β is piecewise causal, that is, $g(\dot{\beta}, \dot{\beta}) \leq 0$ almost everywhere, we have that

$$|\dot{\beta}^\tau| \geq \|\dot{\beta}\|_{\bar{g}},$$

and the inequalities for the lengths follow immediately.

The second inequality for lengths descends to the level of distances, because the class of piecewise causal curves considered (for $\hat{d}_\tau$) is contained in the class of all locally Lipschitz curves (strictly speaking, d_W is obtained via piecewise smooth curves, but [Bur15, Cor. 3.13] shows that it is the same as the intrinsic metric obtained via the class of absolutely continuous curves, each of which has a locally Lipschitz reparametrization [AGS08, Lem. 1.1.4]). □

The next lemma proves a reverse inequality to (6.9), albeit only locally and via a more involved proof (because d_W requires knowledge also about non-causal curves).

Lemma 6.15. *Let (M, g) be a spacetime equipped with a smooth temporal function τ, so that g and g_W are given by* (6.7) *and* (6.8)*, respectively. Then, around every $x \in M$ there is a neighborhood U such that for all $p, q \in U$,*

$$\hat{d}_\tau(p, q) \leq 4 d_W(p, q).$$

Proof. Let $x \in M$ and $\dim M = N+1$. The idea is to use coordinates adapted to the vector field $\nabla\tau$ and the hypersurface $\mathcal{S}_{\tau(x):=}\tau^{-1}(\tau(x))$ to explicitly construct approximating piecewise causal curves β_n in a neighborhood U of x.

Step 1. Construction of U and coordinates. The level set $\mathcal{S}_{\tau(x)}$ is a hypersurface in M (since $d\tau \neq 0$), therefore there exists a neighborhood $V \subseteq \mathcal{S}_{\tau(x)}$ of x and a coordinate map $\varphi\colon V \to \varphi(V) \subseteq \mathbb{R}^N$. Let $y \in \varphi(V) \subseteq \mathbb{R}^N$ and let c_y denote the unique integral curve of $-\nabla\tau$ in M with $c_y(0) \in V \subseteq \mathcal{S}_{\tau(x)} \subseteq M$ such that $\varphi(c_y(0)) = y$. Since we have assumed that $d\tau(-\nabla\tau) = -g(\nabla\tau, \nabla\tau) = 1$, we have that

$$\tau(c_y(t)) = \tau(x) + t. \tag{6.10}$$

The Flow Box Theorem [Lee09, Thm. 2.91] guarantees the existence of a suitable $a > 0$ and a neighborhood of x (which we immediately restrict to our chart neighborhood on $\mathcal{S}_{\tau(x)}$ and again denote by V) such that the map

$$\begin{aligned} \phi\colon (-a,a) \times \varphi(V) &\to M, \\ (t,y) &\mapsto c_y(t), \end{aligned}$$

is well-defined. By the Inverse Function Theorem, ϕ is a local diffeomorphism, hence one obtains a coordinate system (t,y) on a neighborhood U of x in M. At any point $z = (t,y) \in \phi^{-1}(U) \subseteq (-a,a) \times \mathbb{R}^N$, the differential $D_z\phi\colon \mathbb{R} \times \mathbb{R}^N \to T_{\phi(z)}M$ maps $(1,0,\ldots,0)$ to $\dot{c}_y(t) = \nabla\tau|_{\phi(z)}$. Thanks to (6.10), $D_z\phi$ also maps $\{0\} \times \mathbb{R}^N$ to $T_{\phi(z)}\mathcal{S}_{\tau(\phi(z))} \subseteq T_{\phi(z)}M$. Since $T_{\phi(z)}\mathcal{S}_{\tau(\phi(z))} = (\operatorname{span} \nabla\tau|_{\phi(z)})^\perp$, pulling back the metric tensor g with ϕ, we see that its components in the (t,y) coordinates are

$$g_{00} = -1 \qquad\qquad g_{0i} = 0 \quad \text{for all} \quad i \neq 0, \tag{6.11}$$

where the subindex 0 corresponds to the t-component and the subindex i to any of the y-components. Finally, by further shrinking U, we may assume that it is g_W-convex.

Step 2. Construction of piecewise causal curves β_n approximating $d_W(p,q)$. Let $p, q \in U$ be given. By assumption on U, there exists a length-minimizing g_W-geodesic $\gamma\colon [0,L] \to U$ from p to q, parametrized by arclength. In particular,

$$L = L_W(\gamma) = d_W(p,q).$$

We write $\gamma = (\gamma^\tau, \gamma^1, ..., \gamma^N)$ to denote the coordinate expression of γ. Consider the sequence $(\beta_n)_n$ of curves given in coordinates by

$$\beta_n(s) := \left(\gamma^\tau(s) + 3f_n(s), \gamma^1(s), ..., \gamma^N(s)\right), \tag{6.12}$$

for $|f_n|$ sufficiently small so that β_n is contained in U. Furthermore, we choose $(f_n)_n$ in a way that β_n is piecewise causal, that is,

$$g(\dot\beta_n, \dot\beta_n) \leq 0 \quad \text{almost everywhere},$$

and so that $\hat{L}_\tau(\beta_n) \leq 4L_W(\gamma)$ for n sufficiently large.

We show that the functions $f_n\colon [0,L] \to \mathbb{R}$, given by

$$f_0(s) := \begin{cases} s & \text{for } s \in [0, \frac{L}{2}], \\ L - s & \text{for } s \in [\frac{L}{2}, L], \end{cases} \qquad \text{and}$$

$$f_n(s) := \frac{1}{2} \begin{cases} f_{n-1}(2s) & \text{for } s \in [0, \frac{L}{2}], \\ f_{n-1}(2s - L) & \text{for } s \in [\frac{L}{2}, L], \end{cases}$$

as depicted in Figure 6.1a satisfy all these properties. The curves β_n given by (6.12) are shown in Figure 6.1b. By definition, $\dot\beta_n$ exists almost everywhere, and by (6.11) it follows that

$$\begin{aligned} g(\dot\beta_n, \dot\beta_n) &= -|\dot\beta_n^\tau|^2 + \bar g(\dot\beta_n, \dot\beta_n) \\ &= -|\dot\gamma_n^\tau + 3f_n'|^2 + (\bar g_{ij} \circ \beta_n)\dot\gamma^i \dot\gamma^j. \end{aligned} \tag{6.13}$$

Since $|f_n'| = 1$ almost everywhere, and $|\dot\gamma_n^\tau| \leq g_W(\dot\gamma, \dot\gamma) = 1$, it follows that almost everywhere

$$2 \leq 3|f_n'| - |\dot\gamma_n^\tau| \leq |\dot\gamma_n^\tau + 3f_n'| \leq 3|f_n'| + |\dot\gamma_n^\tau| \leq 4. \tag{6.14}$$

We consider the second term in (6.13). Since γ is a g_W-geodesic it is smooth and so are all coordinate components $\dot\gamma^i$ of tangent vector. Due to compactness of γ the $\dot\gamma^i$ must therefore also be bounded, say $|\dot\gamma^i| \leq M$ for some $M > 0$. Due to $|f_n(s)| \leq \frac{L}{2^{n+1}} \to 0$ uniformly as $n \to \infty$ it follows that $\beta_n^\tau \to \gamma_n^\tau$ uniformly, and thus $\beta_n \to \gamma$ uniformly in coordinates. The coordinate functions $\bar g_{ij} \circ \beta_n \to h_{ij} \circ \gamma$ converge uniformly on $[0, L]$ as $n \to \infty$ as well. Thus for any $\varepsilon \in (0, \frac{1}{M^2})$ there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, for all $s \in [0, L]$,

$$\begin{aligned} &\bar g_{ij}(\beta_n(s))\dot\gamma^i(s)\dot\gamma^j(s) \\ &\quad\leq \bar g_{ij}(\gamma(s))\dot\gamma^i(s)\dot\gamma^j(s) + |\bar g_{ij}(\beta_n(s)) - \bar g_{ij}(\gamma(s))||\dot\gamma^i(s)||\dot\gamma^j(s)| \\ &\quad\leq g_W(\dot\gamma(s), \dot\gamma(s)) + \varepsilon M^2 \leq 2. \end{aligned} \tag{6.15}$$

Together, (6.13)–(6.15) imply

$$g(\dot\beta_n, \dot\beta_n) \leq -2^2 + 2 \leq 0,$$

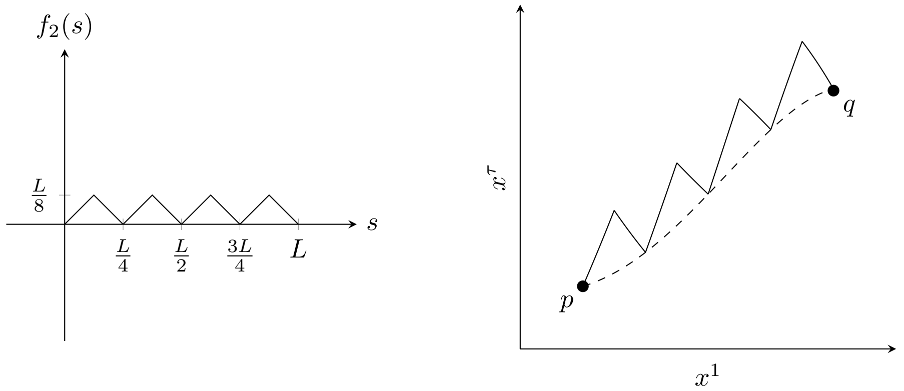
hence all β_n for $n \geq n_0$ are piecewise causal.

Using again that (6.14) holds almost everywhere, it follows that for all n

$$\hat L_\tau(\beta_n) = \int_0^L |(\tau \circ \beta_n)'(s)| ds = \int_0^L |\dot\gamma^\tau(s) + 3f_n'(s)| ds \leq 4L = 4L_W(\gamma),$$

and since γ is a length-minimizing g_W-geodesic we conclude that

$$\hat d_\tau(p, q) \leq 4L = 4d_W(p, q). \qquad \square$$



(a) The function f_2.

(b) The curves γ (dashed) and β_2 (solid).

Figure 6.1: Illustration of the proof of Lemma 6.15.

6.2.2 Proof of Theorem 6.7

In Section 6.2.1 we have already shown that the bi-Lipschitz estimates (6.3) of Theorem 6.7 hold globally (the first inequality, by Lemma 6.14) and locally (the second inequality, by Lemma 6.15) if we choose τ to be a temporal function and $h = g_W$ to be the corresponding Wick-rotated Riemannian metric. We first extend these local results to weak temporal functions and arbitrary Riemannian metrics, and then prove Theorem 6.7 on compact sets.

Lemma 6.16. *Let (M, g) be a spacetime equipped with a weak temporal function τ. Suppose h is a Riemannian metric on M. Then, for every $x \in M$ there exists a neighborhood U of x and a constant $C \geq 1$ such that for all $p, q \in U$*

$$\frac{1}{C} d_h(p, q) \leq \hat{d}_\tau(p, q) \leq C d_h(p, q). \tag{6.16}$$

Proof. Let $x \in M$. We consider both inequalities in (6.16) separately for different Riemannian metrics and choose U to be intersection of the neighborhoods V_1 and V_2 derived in each step. More precisely, the first inequality follows from the local anti-Lipschitz assumption for τ in Definition 6.6, and the second inequality from Lemma 6.15 for an auxiliary temporal functions. Without loss of generality we furthermore assume that U is relatively compact. Then by a result of Burtscher [Bur15, Thm. 4.5] the bi-Lipschitz estimate extends to any Riemannian metric.

Step 1. Lower bound. We show that $\frac{1}{C} d_h(p, q) \leq \hat{d}_\tau(p, q)$ for C and h as in Definition 6.6.

Let U_1 be the neighborhood of x from Definition 6.6. Since $\hat{d}_\tau$ induces the manifold topology there is a radius $r > 0$ such that for the open ball $\hat{B}^\tau_{3r}(x) \subseteq U_1$. Consider $p, q \in V_1 := \hat{B}^\tau_r(x)$. There exists a sequence $(\beta_n)_n$ of piecewise causal curves in M between p and q such that

$$\hat{L}_\tau(\beta_n) \leq \hat{d}_\tau(p,q) + \frac{1}{n}.$$

For all $n > \frac{1}{r}$ the curves β_n cannot leave U_1. Consider one such $\beta_n\colon [0,1] \to U_1$ and its partition $0 = s_0 < s_1 < \ldots < s_{k-1} < s_k = 1$ in causal pieces. Then for C and h as in Definition 6.6

$$\begin{aligned}\frac{1}{C} d_h(p,q) &\leq \frac{1}{C}\sum_{i=1}^{k} d_h(\beta_n(s_i),\beta_n(s_{i-1})) \\ &\leq \sum_{i=1}^{k} \hat{d}_\tau(\beta_n(s_i),\beta_n(s_{i-1})) = \hat{L}_\tau(\beta_n) \leq \hat{d}_\tau(p,q) + \frac{1}{n}.\end{aligned}$$

Thus the result follows as $n \to \infty$.

Step 2. Upper bound. We show that $\hat{d}_\tau(p,q) \leq 5Cd_W(p,q)$ for g_W the Riemannian metric (6.8) with respect to an auxiliary temporal function and corresponding C of Definition 6.6.

Every spacetime which admits a time function also admits a temporal function $\tilde{\tau}$ by Bernal and Sánchez [BS05, Thm. 1.2]. We consider the Wick-rotated Riemannian metric g_W defined in (6.8) with respect to such a fixed $\tilde{\tau}$. Let V_2 be a geodesically convex neighborhood of x with respect to g_W. Without loss of generality we assume that V_2 is contained in the neighborhoods U_2 of Definition 6.6 and Lemma 6.15. Suppose $p,q \in V_2$ and γ is the g_W-length minimizing geodesic γ between p and q in V_2. In the proof of Lemma 6.15 a piecewise causal curve β in V_2, sufficiently close to γ, was constructed for which

$$\begin{aligned}L_W(\beta) &= \int_0^L \|\dot{\beta}(s)\|_W ds = \int_0^L \sqrt{|\dot{\beta}^\tau|^2 + h(\dot{\beta},\dot{\beta})} ds \\ &\leq \int_0^L \sqrt{4^2+2} ds < 5L = 5L_W(\gamma). \end{aligned} \tag{6.17}$$

Since β is piecewise causal, there is a partition $0 = s_0 < s_1 < \ldots < s_{k-1} < s_k = L$ such that

$$\hat{L}_\tau(\beta) = \sum_{i=1}^{k} \hat{d}_\tau(\beta(s_i),\beta(s_{i-1})),$$

and τ being a weak temporal function implies that there is a constant $C \geq 1$ (Definition 6.6) such that

$$\hat{L}_\tau(\beta) \leq C\sum_{i=1}^{k} d_W(\beta(s_i),\beta(s_{i-1})) \leq CL_W(\beta).$$

Together with (6.17) and the fact that γ is g_W-length minimizing we thus obtain

$$\hat{d}_\tau(p,q) \leq \hat{L}_\tau(\beta) \leq CL_W(\beta) \leq 5CL_W(\gamma) = 5Cd_W(p,q). \qquad \square$$

Lemma 6.16 shows that the original local bi-Lipschitz bound (6.2) for causality related points (as required for weak temporal functions in Definition 6.6) extends to the bi-Lipschitz bound (6.16) on a whole neighborhood of each point. Finally, we adapt the proof of Burtscher [Bur15, Thm. 4.5] to extend this local estimate to compact sets.

Proof of Theorem 6.7. Let K be a compact subset of M. The argument proceeds by contradiction. Suppose that for any $n \in \mathbb{N}$, there exist points $p_n, q_n \in K$ such that

$$\hat{d}_\tau(p_n, q_n) > n d_h(p_n, q_n). \tag{6.18}$$

By passing to subsequences, we can assume that $p_n \to p$ and $q_n \to q$ (in one and hence both metrics, since they both induce the manifold topology). Since K is compact it is bounded with respect to the null distance $\hat{d}_\tau$ and the inequality (6.18) furthermore implies that

$$d_h(p_n, q_n) \to 0 \quad \text{as} \quad n \to \infty,$$

and hence $p = q$. Thus, for n large enough, p_n, q_n are contained in neighborhood U of p that by Lemma 6.16 is small enough so that (6.16) holds for some $\tilde{C} \geq 1$. But then we conclude that

$$\tilde{C} d_h(p_n, q_n) \geq \hat{d}_\tau(p_n, q_n) > n d_h(p_n, q_n),$$

a contradiction for n large. This proves that there is a constant $C \geq 1$ such that $\hat{d}_\tau(p, q) \leq C d_h(p, q)$, and the reverse inequality is obtained in the same way. □

Remark 6.17. Bi-Lipschitz maps are a crucial tool in Metric Geometry. In particular, they are useful in the study of Gromov–Hausdorff convergence of metric spaces. Moreover, the estimates of Theorem 6.7 can be used to equip suitable spacetimes with a (local) integral current space structure (see [AK00; FF60; JL21; Lan11; LW11; SW11]) via the approach laid out in Allen and Burtscher [AB22, Sec. 2.5, 2.6, and 4]. Thanks to our Completeness Theorem 6.4 and [AB22, Thm. 1.3] we know, for instance, that any globally hyperbolic spacetime viewed as the metric space $(M, \hat{d}_\tau)$ with completely uniform temporal function τ is a (local) integral current space (see Section 6.4). This also makes it possible to use spacetime intrinsic flat convergence to study geometric stability questions in General Relativity, as proposed by Sormani.

6.2.3 Direct proof of Corollary 6.11

Corollary 6.11 is a very special case of Theorem 6.7. Here we show that it can also be obtained directly without the use of temporal functions. We first prove a local result.

Lemma 6.18. *Let τ_1, τ_2 be two weak temporal functions (see Definition 6.6) on a spacetime (M, g), and $\hat{d}_{\tau_1}$, $\hat{d}_{\tau_2}$ their associated null distances. Then, for every* *point $x \in M$, there exists a neighborhood U of x and a constant $C \geq 1$ such that for all $p, q \in U$,*

$$\frac{1}{C}\hat{d}_{\tau_1}(p,q) \leq \hat{d}_{\tau_2}(p,q) \leq C\hat{d}_{\tau_1}(p,q).$$

Proof. Fix a point $x \in M$ and an arbitrary Riemannian metric h on M. Let $U_x^{\tau_1}$ and $U_x^{\tau_2}$ be the neighborhoods around x where both τ_1 and τ_2 are Lipschitz and anti-Lipschitz with positive constants C_1, C_2 with respect to h as in (6.2) of Definition 6.6. Since M is locally compact there exists a relatively compact neighborhood $V \subseteq U_x^{\tau_1} \cap U_x^{\tau_2}$ of x. Since both null distances $\hat{d}_{\tau_1}, \hat{d}_{\tau_2}$ induce the manifold topology [SV16, Thm. 4.6], we find $r > 0$ sufficiently small so that $\hat{B}_{4r}^{\tau_1}(x) \cup \hat{B}_{4r}^{\tau_2}(x) \subseteq V$. Define a neighborhood of x by

$$U := \hat{B}_r^{\tau_1}(x) \cap \hat{B}_r^{\tau_2}(x).$$

We claim that for $i = 1, 2$ and for all $p, q \in U$,

$$\hat{d}_{\tau_i}(p,q) = \inf\{\hat{L}_{\tau_i}(\beta) \mid \beta \text{ a p.w. causal curve in } V \text{ between } p \text{ and } q\}. \tag{6.19}$$

What we are saying is that the null distances in U can be approximated by sequences of curves that never leave V. To prove this, let $p, q \in U$, consider an arbitrary $\varepsilon \in (0, r)$, and for each $i = 1, 2$ choose a piecewise causal curve $\beta_\varepsilon^i \colon I \to M$ from p to q such that

$$\hat{L}_{\tau_i}(\beta_\varepsilon^i) < \hat{d}_{\tau_i}(p,q) + \varepsilon.$$

Then, for all $t \in I$,

$$\begin{aligned}\hat{d}_{\tau_i}(x, \beta_\varepsilon^i(t)) &\leq \hat{d}_{\tau_i}(x,p) + \hat{d}_{\tau_i}(p, \beta_\varepsilon^i(t)) \leq \hat{d}_{\tau_i}(x,p) + \hat{L}_{\tau_i}(\beta_\varepsilon^i) \\ &\leq \hat{d}_{\tau_i}(x,p) + \hat{d}_{\tau_i}(p,q) + \varepsilon < 4r,\end{aligned}$$

hence β_ε^i lies entirely in $\hat{B}_{4r}^{\tau_i}(x) \subseteq V$, proving the claim (6.19) for $i = 1, 2$.

For any causally related points $(\tilde{p}, \tilde{q}) \in J^+$ in $U_x^{\tau_1} \cap U_x^{\tau_2}$, by (6.2), we have for $C := C_1 C_2 \geq 1$ that

$$\frac{1}{C}(\tau_1(\tilde{q}) - \tau_1(\tilde{p})) \leq \tau_2(\tilde{q}) - \tau_2(\tilde{p}) \leq C(\tau_1(\tilde{q}) - \tau_1(\tilde{p})). \tag{6.20}$$

This implies that for all piecewise causal curves β contained in $U_x^{\tau_1} \cap U_x^{\tau_2}$,

$$\frac{1}{C}\hat{L}_{\tau_1}(\beta) \leq \hat{L}_{\tau_2}(\beta) \leq C\hat{L}_{\tau_1}(\beta). \tag{6.21}$$

It remains to be shown that the bi-Lipschitz bounds on the null lengths extends to the null distances between any two points p and q in the smaller set U. We have already seen that for each $n \in \mathbb{N}$ exists a piecewise causal curve β_n^1 between p and q (by (6.19) β_n^1 is entirely contained in V!) such that

$$\hat{L}_{\tau_1}(\beta_n^1) \leq \hat{d}_{\tau_1}(p,q) + \frac{1}{n},$$

and by (6.21) therefore

$$\hat{d}_{\tau_2}(p,q) \leq \hat{L}_{\tau_2}(\beta_n^1) \leq C\hat{L}_{\tau_1}(\beta_n^1) \leq C\hat{d}_{\tau_1}(p,q) + \frac{C}{n}.$$

Since this inequality holds for all n and fixed C, we have that $\hat{d}_{\tau_2}(p,q) \leq C\hat{d}_{\tau_1}(p,q)$ on U. In the same way, for piecewise causal curves β_n^2 approximating $\hat{d}_{\tau_2}(p,q)$, we obtain $\hat{d}_{\tau_1}(p,q) \leq C\hat{d}_{\tau_2}(p,q)$ on U, which proves the Lemma. □

Proof of Corollary 6.11. Having established the local result in Lemma 6.18, the proof on compact sets is now the same as that of Theorem 6.7. □

6.2.4 Counterexamples

Finally, we construct two examples showing that the assumptions and statement of Corollary 6.11, and hence also that of Theorem 6.7 and Corollary 6.8, are sharp.

Example 6.19 (Corollary 6.11 is false globally)**.** Let $M = \mathbb{R}^{1,n}$ be the Minkowski spacetime, and $\tau_1 = t$, $\tau_2 = \exp(t)$ two temporal functions (which even have the same level sets!). Since the exponential function is *not* globally Lipschitz, the time functions τ_1 and τ_2 are not equivalent in the sense that the bi-Lipschitz condition (6.20) for causally related points does not hold globally. Since $\tau(q)-\tau(p) = \hat{d}_\tau(p,q)$ for all $q \in J^+(p)$, it follows that if (6.20) fails to hold globally, then the null distances $\hat{d}_{\tau_1}$, $\hat{d}_{\tau_2}$ are globally inequivalent.

Example 6.19 shows that although τ_1 and τ_2 are both temporal functions on the *same* spacetime (M,g), there is a clear distinction *globally* between the metric structures induced by the null distances $\hat{d}_{\tau_1}$ and $\hat{d}_{\tau_2}$. In particular, in Section 6.4 we show that $(M,\hat{d}_{\tau_1})$ is complete while $(M,\hat{d}_{\tau_2})$ is not.

Example 6.20 (Corollary 6.11 is false for time functions that are not locally Lipschitz)**.** Consider the Minkowski spacetime $M = \mathbb{R}^{1,n}$ with respect to the usual global coordinates (t,x). Let $\tau_1(t,x) := t$ be the standard (smooth and locally anti-Lipschitz) time function and

$$\tau_2(t,x) := \operatorname{sgn}(t)\sqrt{|t|}.$$

Clearly, τ_2 is continuous and a time function. It is, however, *not* locally Lipschitz at $t=0$ because $\partial_t\tau_2(t,x) \to \infty$ as $t \to 0$. Nonetheless, τ_2 is locally anti-Lipschitz (and therefore $\hat{d}_{\tau_2}$ induces the manifold topology): Consider a point (t,x) with $t > 0$ and the neighborhood $U = (0,2t)\times\mathbb{R}^n$. Suppose $q \in J^+(p)$ with $p = (t_p,x_p)$, $q = (t_q,x_q)$. By the Mean Value Theorem for $\sqrt{\ }$ on $[0,2t]$ we have

$$\sqrt{t_q} - \sqrt{t_p} \geq \left(\inf_{r\in(0,2t)} \frac{d}{dr}\sqrt{r}\right)(t_q - t_p) = \frac{1}{2\sqrt{2t}}|t_q - t_p|.$$

Hence due to the causal relation on Minkowski spacetime we obtain with respect to the Euclidean distance d and with $C := \frac{1}{4\sqrt{2t}} > 0$ that on U

$$q \in J^+(p) \Longrightarrow \tau_2(q) - \tau_2(p) \geq \frac{1}{4\sqrt{2t}}\sqrt{(t_q-t_p)^2 + |x_q - x_p|^2} \geq Cd(q,p).$$

In the same fashion, τ_2 is anti-Lipschitz in the neighborhood $U = (2t, 0) \times \mathbb{R}$ of a point (t, x) with $t < 0$. If $t = 0$, we can simply use the neighborhood $U = (-1, 1) \times \mathbb{R}^n$ and $C = \frac{1}{4}$ (since the infimum is at $|r| = 1$). Hence τ_2 is locally anti-Lipschitz everywhere (but clearly not globally anti-Lipschitz with respect to the Euclidean distance).

Now assume that $p = (0, 0)$ and let $q_t = (t, 0)$, $t \in (0, 1)$, be arbitrarily close to p. Then

$$\hat{d}_{\tau_1}(p, q_t) = t \leq \sqrt{t} = \hat{d}_{\tau_2}(p, q_t)$$

but the bi-Lipschitz estimate (6.3) of Theorem 6.11 does not hold because

$$\frac{\hat{d}_{\tau_2}(p, q_t)}{\hat{d}_{\tau_1}(p, q_t)} = \frac{1}{\sqrt{t}} \to \infty \qquad \text{as } t \to 0.$$

6.2.5 Basic properties of weak temporal functions

The following result justifies the notion of weak temporal functions.

Proposition 6.21. *Let (M, g) be a spacetime and $\tau \colon M \to \mathbb{R}$ be a weak temporal function. Then τ is locally Lipschitz with past-directed timelike gradient $\nabla\tau$ almost everywhere.*

Proof. The local Lipschitz condition of Definition 6.6 implies the upper bound in Lemma 6.16, i.e., for each Riemannian metric h and each point of M there is an open neighborhood U and $C \geq 1$ such that for $p, q \in U$

$$|\tau(q) - \tau(p)| \leq \hat{d}_\tau(p, q) \leq C d_h(p, q).$$

In other words, τ is locally Lipschitz on M. By Rademacher's Theorem $\nabla\tau$ therefore exits almost everywhere.

Suppose that $\nabla\tau$ exists a point p. Consider a future-directed causal vector $v \in T_pM \setminus \{0\}$ and a smooth causal curve γ from $p = \gamma(0)$ in direction v. Then

$$d\tau(v) = \left.\frac{d}{ds}\right|_{s=0} (\tau \circ \gamma)(s) = \lim_{s \to 0} \frac{\tau(\gamma(s)) - \tau(p)}{s} = \lim_{s \to 0} \frac{\hat{d}_\tau(\gamma(0), \gamma(s))}{s},$$

which by the local anti-Lipschitz property of τ and [Bur15, Prop. 4.10] implies that

$$d\tau(v) \geq \frac{1}{C} \lim_{s \to 0} \frac{d_h(\gamma(0), \gamma(s))}{s} = \frac{1}{C}\|v\|_h > 0.$$

Hence $\nabla\tau$ is past-directed timelike whenever it exists. □

The converse is not true, in the sense that a function with almost everywhere timelike gradient is not necessarily weak temporal, unless one assumes local upper and lower bounds on $\|\nabla\tau\|_h$. Indeed, the time functions in Example 6.20 and [SV16, Ex. 3.4] have timelike gradient almost everywhere, but the first is not locally Lipschitz and the second not locally anti-Lipschitz.

The following important classes of time functions are weak temporal.

Lemma 6.22. *Temporal functions and regular cosmological time functions (à la Andersson–Galloway–Howard [AGH98] and Wald–Yip [WY81]) are weak temporal functions.*

Proof. For temporal functions, locally Lipschitz follows by smoothness and locally anti-Lipschitz by [SV16, Cor. 4.16]. For regular cosmological time functions, locally Lipschitz follows by [AGH98, Thm. 1.2(v)] and locally anti-Lipschitz by [SV16, Thm. 5.4]. □

6.3 Encoding causality

In this section we prove Theorem 6.9 and a corresponding local result. Recall that we defined the causal relation $J^+ \subseteq M \times M$ by

$$J^+ := \{(p,q) \mid \text{there exists a future-directed causal curve from } p \text{ to } q\}.$$

By definition the null distance is achieved for causal curves. We investigate when the converse holds, that is, when

$$(p,q) \in J^+ \iff \hat{d}_\tau(p,q) = \tau(q) - \tau(p), \tag{6.22}$$

in which case the null distance is said to *encode causality*, an open problem mentioned in [SV16, Sec. 1]. We introduce a new notation for the right hand side of (6.22).

Definition 6.23. Let (M,g) be a spacetime equipped with a time function τ. We define the *null distance relation* $\hat{R}^+_\tau \subseteq M \times M$ by

$$(p,q) \in \hat{R}^+_\tau :\iff \hat{d}_\tau(p,q) = \tau(q) - \tau(p).$$

Clearly, the relation $\hat{R}^+_\tau$ is reflexive and transitive. Antisymmetry requires definiteness of $\hat{d}_\tau$ which, for instance, follows if τ is a locally anti-Lipschitz time function [SV16, Thm. 4.6]. We can thus summarize the basic properties of $\hat{R}^+_\tau$ as follows.

Lemma 6.24. *Suppose (M,g) is a spacetime and τ is a locally anti-Lipschitz time function on M. Then the null distance relation $\hat{R}^+_\tau$ is a partial order on M satisfying $J^+ \subseteq \hat{R}^+_\tau$.* □

In Section 6.3.1 we prove that globally hyperbolic spacetimes encode causality when τ is carefully chosen, which can be reformulated in terms of $\hat{R}^+_\tau$ as follows.

Theorem 6.25. *Let (M,g) be a globally hyperbolic spacetime and let τ be a temporal function such that every nonempty level set is a Cauchy surface. Then causality is encoded in $\hat{d}_\tau$, meaning $J^+ = \hat{R}^+_\tau$.*

Note that the time functions $\tau\colon M \to (T_1, T_2)$ of Theorems 6.9 and 6.25 are Cauchy time functions (as per Theorem 6.3) up to composition with an increasing homeomorphism $f\colon (T_1,T_2) \to \mathbb{R}$. Our relaxed Cauchy assumption adds value in the weak context, notably because cosmological time functions take values only on $(0, \infty)$.

Subsequently we also prove a local version of this result for any spacetime with a temporal function.

Theorem 6.26. *Let (M, g) be a spacetime and τ be a temporal function. Then, every point $x \in M$ has a neighborhood U such that*

$$J^+ \cap (U \times U) = \hat{R}^+_\tau \cap (U \times U).$$

Upon completion of an earlier preprint of this manuscript and that of Sakovich and Sormani [SS23] we noticed that our different proofs can be combined to yield a stronger result of both Theorem 6.25 (to locally anti-Lipschitz τ) and [SS23] (to noncompact Cauchy slices). This stronger result is Theorem 6.9 and proven in Section 6.3.2.

In Section 6.3.3 we show that the assumptions for the global Theorem 6.9 are sharp in the sense that there are globally hyperbolic spacetimes and locally anti-Lipschitz functions τ for which $J^+ \subsetneq \hat{R}^+_\tau$.

6.3.1 Proofs of Theorems 6.25 and 6.26 for temporal functions

In this section, the Wick-rotated metric g_W of g introduced in Section 6.2.1 plays an important role again. Since g_W need not be complete even if the spacetime is globally hyperbolic (see Example 6.39) we furthermore make use of the existence of a conformally equivalent complete Riemannian metric

$$g_R := \Omega^2 g_W, \tag{6.23}$$

where $\Omega \colon M \to [1, \infty)$ is a smooth function [NO61, Thm. 1]. We denote the corresponding norm, length functional and distance by $\|\cdot\|_R$, L_R and d_R, respectively. The relation of g_R and g_W carries over to the distances as follows, that is, for any curve $\gamma \colon [a, b] \to M$,

$$L_W(\gamma) \leq L_R(\gamma),$$

and hence for any $p, q \in M$

$$d_W(p, q) \leq d_R(p, q).$$

In the following lemma, we obtain reverse inequalities on compact sets.

Lemma 6.27. *Let g_R and g_W be two conformally equivalent Riemannian metrics on M as in (6.23), and suppose K is a compact set in M. Then there exist constants $C_K \geq c_K \geq 1$ such that for any curve γ in K*

$$L_W(\gamma) \leq L_R(\gamma) \leq c_K L_W(\gamma).$$

and for any $p, q \in K$,

$$d_W(p, q) \leq d_R(p, q) \leq C_K d_W(p, q).$$

Proof. The first statement about the lengths follows immediately from the Definition (6.23), and we may pick $c_K := \max_{x \in K} \Omega(x)$. The first inequality for the distances is trivial and the second inequality follows from [Bur15, Thm. 4.5] (note that, in general, we need to pick $C_K > c_K$ as minimizing curves may leave K, but the proof in [Bur15] guarantees boundedness of C_K). □

With these tools, we proceed to prove the theorem.

Proof of Theorem 6.25. Trivially, $J^+ \subseteq \hat{R}_\tau^+$, so we only need to prove $\hat{R}_\tau^+ \subseteq J^+$. Suppose that $(p,q) \in \hat{R}_\tau^+$, that is, $\hat{d}_\tau(p,q) = \tau(q) - \tau(p)$. Since $(p,p) \in J^+$ is trivially satisfied, we can assume that $p \neq q$ and thus necessarily $\tau(q) > \tau(p)$. By definition of $\hat{d}_\tau$ there exists a sequence of piecewise causal curves $(\beta_n)_n$ such that

$$0 < \tau(q) - \tau(p) \leq \hat{L}_\tau(\beta_n) \leq \tau(q) - \tau(p) + \frac{1}{n}. \tag{6.24}$$

In what follows we construct a future-directed causal curve β between p and q, which then immediately implies $(p,q) \in J^+$. We proceed as follows: After a preliminary local estimate we construct a candidate limit curve near p in (M, g_R), and then show that it is both locally Lipschitz and future-directed causal. Finally, we show it naturally extends all the way up to q.

Step 1. A local version of (6.24)**.** We consider $\beta_n \colon [0, L_n] \to M$ on arbitrary subintervals of its domain. Suppose there exists a subinterval $[s,t] \subseteq [0, L_n]$ such that $\hat{L}_\tau(\beta_n|_{[s,t]}) > \tau(\beta_n(t)) - \tau(\beta_n(s)) + \frac{1}{n}$. Then by the additivity of null lengths

$$\begin{aligned}
\hat{L}_\tau(\beta_n) &= \hat{L}_\tau(\beta_n|_{[0,s]}) + \hat{L}_\tau(\beta_n|_{[s,t]}) + \hat{L}_\tau(\beta_n|_{[t,L_n]}) \\
&> |\tau(\beta_n(s)) - \tau(p)| + |\tau(\beta_n(t)) - \tau(\beta_n(s)) + \frac{1}{n}| + |\tau(q) - \tau(\beta_n(t))| \\
&\geq \tau(q) - \tau(p) + \frac{1}{n},
\end{aligned}$$

a contradiction to (6.24). Thus for all subintervals $[s,t]$ of the domain of β_n

$$|\tau(\beta_n(t)) - \tau(\beta_n(s))| \leq \hat{L}_\tau(\beta_n|_{[s,t]}) \leq \tau(\beta_n(t)) - \tau(\beta_n(s)) + \frac{1}{n}. \tag{6.25}$$

(Note that at this point we cannot yet rule out that $\tau(\beta_n(t)) - \tau(\beta_n(s)) < 0$, which is why we do not want to drop the absolute value on the left hand side.)

Step 2. Construction of a candidate limit curve β near p. Suppose, that each $\beta_n \colon [0, L_n] \to M$ is parametrized by g_R-arclength with $\beta_n(0) = p$ and $\beta_n(L_n) = q$. In addition, we attach a future-directed inextendible causal curve $\tilde{\beta}$ at q. Since g_R is complete, $\tilde{\beta}$ must have infinite g_R-length. Thus we can extend all β_n by $\tilde{\beta}$ to obtain future-indextendible piecewise causal curves with g_R-arclength parametrization, again denoted by $\beta_n \colon [0, \infty) \to M$ and satisfying (6.25). In particular, for each n and any $s, t \in [0, \infty)$,

$$d_R(\beta_n(s), \beta_n(t)) \leq L_R(\beta_n|_{[s,t]}) = |t - s|. \tag{6.26}$$

By construction, $\beta_n(0) = p$ for all n. Fix any other $t_0 \in (0, \infty)$. Then, due to the g_R-arclength parametrization, each curve segment $\beta_n|_{[0,t_0]}$ is contained in the bounded set (which, by the Hopf–Rinow Theorem for g_R is also compact)

$$B_{t_0}(p) := \{x \in M : d_R(p, x) \leq t_0\}.$$

In particular, the family $(\beta_n|_{[0,t_0]})_n$ is uniformly bounded and uniformly equicontinuous. Thus, by the Arzelà–Ascoli Theorem (see, for instance, [Mun00, Thm. 47.1] or [BEE96, Thm. 3.30]), there exists a continuous curve $\beta\colon [0, \infty) \to M$, such that a subsequence of $(\beta_n)_n$ converges uniformly to β on all compact subintervals. Since $\frac{1}{n_k} \leq \frac{1}{k}$ for any subsequence, we can denote this subsequence again by $(\beta_n)_n$. Moreover, (6.26) implies that β is a locally Lipschitz curve with

$$d_R(\beta(s), \beta(t)) \leq |t - s| \tag{6.27}$$

for all $s, t \in [0, \infty)$.

Step 3. The curve β is future-directed causal. Since β is locally Lipschitz, together with Rademacher's Theorem, we know that $\dot\beta$ exists almost everywhere. To conclude that β is a future-directed causal curve, it remains to be shown that $g(\dot\beta, \dot\beta) \leq 0$ and $(\tau \circ \beta)' > 0$ almost everywhere.

By [Bur15, Prop. 4.10] the g_R-norm of the analytic derivative and d_R-metric derivative of β exist and coincide almost everywhere. Combined with the Lipschitz estimate (6.27) for β we thus have for almost all $s \in [0, \infty)$

$$0 \leq \|\dot\beta(s)\|_R = \lim_{h\to 0} \frac{d_R(\beta(s+h), \beta(s))}{|h|} \leq \lim_{h\to 0} \frac{|h|}{|h|} = 1. \tag{6.28}$$

In order to show that $\dot\beta$ is almost everywhere causal we need to control the τ-component $\dot\beta^\tau = (\tau \circ \beta)'$ of the tangent vectors. Suppose $\dot\beta(s_0)$ exists for a fixed $s_0 \in (0, \infty)$, and consider the closed (and hence compact) d_R-ball B_ε of radius ε at $\beta(s_0)$ in M. By Lipschitz continuity (6.27) of β the whole interval $[s_0 - \varepsilon/2, s_0 + \varepsilon/2]$, is mapped into $B_{\varepsilon/2}$. We use the approximating sequence $(\beta_n)_n$ of β next, more precisely, that the β_n are piecewise causal and converge uniformly on compact intervals. Thus for sufficiently large n all $\beta_n([s_0 - \varepsilon/2, s_0 + \varepsilon/2]) \subseteq B_\varepsilon$. By the local estimate (6.25)

$$\lim_{n\to\infty} \hat L_\tau(\beta_n|_{[s_0-\varepsilon/2, s_0+\varepsilon/2]}) = \tau(\beta(s_0 + \varepsilon/2)) - \tau(\beta(s_0 - \varepsilon/2)). \tag{6.29}$$

On the other hand, due to the g_R-arclength parametrization of β_n and Lemmas 6.14 and 6.27 (with constant $c_\varepsilon = \max_{x\in B_\varepsilon} \Omega(x) \geq 1$)

$$\begin{aligned} \varepsilon = L_R(\beta_n|_{[s_0-\varepsilon/2, s_0+\varepsilon/2]}) &\leq c_\varepsilon L_W(\beta_n|_{[s_0-\varepsilon/2, s_0+\varepsilon/2]}) \\ &\leq \sqrt{2} c_\varepsilon \hat L_\tau(\beta_n|_{[s_0-\varepsilon/2, s_0+\varepsilon/2]}). \end{aligned}$$

Hence by (6.29)

$$\frac{1}{\sqrt{2} c_\varepsilon} \leq \frac{1}{\varepsilon} \lim_{n\to\infty} \hat L_\tau(\beta_n|_{[s_0-\varepsilon/2, s_0+\varepsilon/2]}) = \frac{\tau(\beta(s_0 + \varepsilon/2)) - \tau(\beta(s_0 - \varepsilon/2))}{\varepsilon}.$$

Due to the continuity of Ω, it follows that $c_\varepsilon \to \Omega(\beta(s_0))$ as $\varepsilon \to 0$, while the difference quotient of $\tau \circ \beta$ converges to the derivative $\dot\beta^\tau(s_0) = (\tau \circ \beta)'(s_0)$. Thus in the limit we obtain

$$\dot\beta^\tau(s_0) \geq \frac{1}{\sqrt{2}\Omega(\beta(s_0))} > 0.$$

Due to the Wick-rotation g_W of g as well as the g_R-bound (6.28)

$$\begin{aligned} g(\dot\beta(s_0), \dot\beta(s_0)) &= -2|\dot\beta^\tau(s_0)|^2 + \|\dot\beta(s_0)\|_W^2 \\ &\leq -\Omega^{-2}(\beta(s_0)) + \Omega^{-2}(\beta(s_0))\|\dot\beta(s_0)\|_R^2 \leq 0. \end{aligned}$$

This proves that $\dot\beta$ is future-directed causal almost everywhere.

Step 4. The point q lies on β. By construction $\beta(0) = p$, and by Step 3 we know that $\tau \circ \beta$ is strictly increasing on $[0, \infty)$. We distinguish two cases:

(i) If there is an $s_0 \in [0, \infty)$ such that $\tau(\beta(s_0)) = \tau(q)$, then the following argument implies that $\beta(s_0) = q$:

Let $\varepsilon > 0$ be arbitrary. Since both $\hat d_\tau$ and d_R induce the manifold topology, there exists a $\delta > 0$ such that $d_R(\beta(s_0), x) < \delta$ implies $\hat d_\tau(\beta(s_0), x) < \varepsilon$. Due to the convergence $\beta_n(s_0) \to \beta(s_0)$ (obtained with respect to d_R in Step 2), for any n sufficiently large

$$\hat d_\tau(\beta(s_0), \beta_n(s_0)) < \varepsilon. \tag{6.30}$$

Morever, due to the continuity of $\tau \circ \beta$, for all $n > \frac{1}{\varepsilon}$ sufficiently large

$$|\tau(\beta_n(s_0)) - \tau(q)| = |\tau(\beta_n(s_0)) - \tau(\beta(s_0))| < \varepsilon.$$

Since $\beta_n(L_n) = q$, the local estimate (6.25) on $[s_0, L_n]$ (or $[L_n, s_0]$ if $L_n < s_0$) yields

$$\begin{aligned} \hat d_\tau(\beta_n(s_0), q) &\leq \hat L_\tau(\beta_n|_{[s_0, L_n]}) \\ &\leq \tau(q) - \tau(\beta_n(s_0)) + \frac{1}{n} < 2\varepsilon. \end{aligned} \tag{6.31}$$

Combining (6.30) and (6.31) implies that $\hat d_\tau(\beta(s_0), q) < 3\varepsilon$ for any $\varepsilon > 0$. Thus $\beta(s_0) = q$.

(ii) The only obstruction to the desired conclusion is therefore that for all $s \in [0, \infty)$

$$\tau(\beta(s)) < \tau(q). \tag{6.32}$$

We show that this case cannot occur. Since the level set $\tau^{-1}(q)$ is a Cauchy surface and τ increases along β, (6.32) implies that the causal curve β is future extendible as a future-directed causal curve. In particular, the future endpoint $x := \lim_{s\to\infty} \beta(s)$ of β exists (since it is necessarily part of any extension) and by (6.32), $\tau(x) \leq \tau(q)$. If $x = q$ we are done, otherwise there exists a relatively compact open set W around $\beta \cup \{x\}$ such that $q \notin \overline{W}$.

We will use the approximating curves β_n to show that β must in fact leave W: Since $\beta_n(0) = p \in W$ and $\beta_n(L_n) = q \notin W$ there exists

$$b_n := \sup\{s \mid \beta_n(t) \in W \text{ for all } t \in [0,s]\} \in (0, L_n),$$

and $\beta_n(b_n) \in \partial W$. Since $\overline{W}$ is compact, by Lemmas 6.14 and 6.27, there exists a constant $C > 0$ such that

$$\begin{aligned} b_n = L_R(\beta_n|_{[0,b_n]}) &\leq \sqrt{2} C \hat{L}_\tau(\beta_n|_{[0,b_n]}) \leq \sqrt{2} C \hat{L}_\tau(\beta_n|_{[0,L_n]}) \\ &\leq \sqrt{2} C \left(\tau(q) - \tau(p) + \frac{1}{n}\right). \end{aligned}$$

Hence all b_n are uniformly bounded from above by a constant

$$a := \sqrt{2} C \left(\tau(q) - \tau(p) + 1\right).$$

In particular, a subsequence converges to $b := \limsup b_n \in [0, a)$. Let $\varepsilon > 0$. By uniform convergence $\beta_n \to \beta$ on $[0, a]$ for all sufficiently large n (along the previous subsequence), we have

$$d_R(\beta(b_n), \beta_n(b_n)) < \varepsilon. \tag{6.33}$$

Since $\beta_n(b_n) \in \partial W$ and ∂W is compact as closed subset of the compact set $\overline{W}$, there exists a subsequence of points $\beta_{n_k}(b_{n_k})$ that converges to a point $y \in \partial W$, i.e., for k sufficiently large

$$d_R(\beta_{n_k}(b_{n_k}), y) < \varepsilon. \tag{6.34}$$

Combining (6.33)–(6.34) yields

$$d_R(\beta(b), y) \leq d_R(\beta(b), \beta_{n_k}(b_{n_k})) + d_R(\beta_{n_k}(b_{n_k}), y) < 2\varepsilon.$$

In other words, $\beta(b) = y \in \partial W$, a contradiction to the assumption that β is entirely contained in the open set W. Hence the assumption (6.32) must be false, and thus by case (i) β indeed reaches q.

To sum up, we have constructed a future-directed causal (Step 3) curve β from p (Step 2) to q (Step 4). Therefore, $(p, q) \in J^+$, and thus $J^+ = \hat{R}^+_\tau$. $\square$

We can adapt the proof of Theorem 6.25 to show its local counterpart Theorem 6.26, which holds for every stably causal Lorentzian manifold (since every such manifold admits a smooth temporal function by [Min19a, Thm. 4.100]).

Proof of Theorem 6.26. Let $x \in M$. Since $J^+ \subseteq \hat{R}^+_\tau$ is always true, the $\subseteq$ inclusion is trivial. In order to show $\supseteq$ we construct a suitable relatively compact open neighborhood U of x and construct causal curves locally similar to the proof of Theorem 6.25. Only Step 2 made use of the complete Riemannian metric g_R and

Step 4 used global hyperbolicity of (M, g) and have to be carried out slightly different.

By local compactness of M there is an $r > 0$ sufficiently small such that $\hat{B}^{\tau}_{3r}(x)$ is relatively compact (the open ball with respect to $\hat{d}_\tau$ of radius $3r$). Consider $U := \hat{B}^{\tau}_{r}(x)$. Suppose now $p, q \in U$ and $(p, q) \in \hat{R}^{+}_{\tau}$, i.e.,

$$\hat{d}_\tau(p, q) = \tau(q) - \tau(p) > 0.$$

By definition of U and the triangle inequality also $\hat{d}_\tau(p, q) < 2r$. Let $(\beta_n)_n$ be a sequence of piecewise causal paths $\beta_n : [0, L_n] \to M$ that approximates $\hat{d}_\tau(p, q)$. Hence we may fix any $\varepsilon \in (0, \hat{d}_\tau(p, q))$ and assume without loss of generality that for all n,

$$\hat{L}_\tau(\beta_n) < \hat{d}_\tau(p, q) + \varepsilon < 2r. \tag{6.35}$$

In particular, contained in $\hat{B}^{\tau}_{3r}(x)$.

Since τ is assumed to be temporal the Wick-rotated metric g_W of g exists (see Section 6.2.1). We assume that the β_n are parametrized by g_W-arclength, and therefore $L_n = L_W(\beta_n)$. By Lemma 6.14 and (6.35) we then obtain the following estimates:

$$\hat{d}_\tau(p, q) \leq \hat{L}_\tau(\beta_n) \leq L_n \leq \sqrt{2}\hat{L}_\tau(\beta_n) \leq \sqrt{2}\left(\hat{d}_\tau(p, q) + \varepsilon\right). \tag{6.36}$$

It is more convenient to have all the β_n defined on the same interval $[0, L]$, which we achieve by extending each β_n as follows: Set $L := \sqrt{2}(\hat{d}_\tau(p, q) + \varepsilon)$. Then attach to each β_n a future-directed causal curve $\tilde{\beta}_n$ starting at q of g_W-length $\tilde{L}_n \leq L - L_n$. Notice that

$$\tilde{L}_n \leq L - L_n < (\sqrt{2} - 1)\hat{d}_\tau(p, q) + \sqrt{2}\varepsilon < 2(\sqrt{2} - 1)r + \sqrt{2}\varepsilon < r$$

for ε small enough. Since $\hat{L}_\tau(\tilde{\beta}_n) \leq \tilde{L}_n$ (by Lemma 6.14) and $\tilde{\beta}_n$ starts at $q \in \hat{B}^{\tau}_{r}(x)$, it follows that $\tilde{\beta}_n$ is contained in $\hat{B}^{\tau}_{3r}(x)$. This also proves that indeed a long enough extension up to $\tilde{L}_n = L - L_n$ exists, given that $\hat{B}^{\tau}_{3r}(x)$ is relatively compact and the spacetime (M, g) is non-imprisoning (because it admits a time function). We have thus obtained a sequence of piecewise causal curves, denoted again as $\beta_n : [0, L] \to \hat{B}^{\tau}_{3r}(x)$. The curves β_n start at $p = \beta_n(0)$, reach $q = \beta_n(L_n)$, and then continue to their endpoint $\beta_n(L)$. The local estimate (6.25) of Step 1 holds too.

Let $B := \hat{B}^{\tau}_{3r}(x)$. By the Arzelà–Ascoli Theorem applied on the compact metric space $(\overline{B}, d_W|_{\overline{B}})$, there exits a subsequence, again denoted by $(\beta_n)_n$, that uniformly converges to a 1-Lipschitz continuous limit curve $\beta : [0, L] \to \overline{B}$ as in Step 2 of the proof of Theorem 6.26.

Step 3 in the proof of Theorem 6.26 can be used verbatim with $g_R = g_W$ (and $\Omega \equiv 1$) since completeness of g_R was not needed here. Thus β is future-directed causal.

It remains to prove that β reaches the point q, which is now easier thanks to the fact that $(\beta_n)_n$ converges to β uniformly on $[0, L]$. Recall that $\beta_n(L_n) = q$

for all $n \in \mathbb{N}$. By passing to a subsequence, if necessary, we may assume that $L_n \to L_\infty \leq L$. But then, given $\delta > 0$, for all n large enough we obtain

$$\begin{aligned} &d_W(\beta(L_\infty), q) \\ &\qquad \leq d_W(\beta(L_\infty), \beta_n(L_\infty)) + d_W(\beta_n(L_\infty), \beta_n(L_n)) + d_W(\beta_n(L_n), q) \\ &\qquad \leq \delta + |L_\infty - L_n| + 0 < 2\delta, \end{aligned}$$

and therefore conclude that $\beta(L_\infty) = q$.

Thus β is indeed a future-directed causal curve from p through q and hence $(p, q) \in J^+$. □

6.3.2 Alternative approaches and proof of Theorem 6.9

In earlier work of Sormani and Vega the important class of warped product spacetimes $I \times_f \Sigma$ with interval $I \subseteq \mathbb{R}$ and complete Riemannian fibers Σ was already shown to encode causality for certain temporal functions [SV16, Thm. 3.25]. It is also easy to see that $\hat{d}_\tau$ encodes causality if all null distances are realized by piecewise causal curves and τ is locally anti-Lipschitz [AB22, Rem. 3.22]. It remained an open problem to understand causality encoding in the general case. Independently to our approach, Sakovich and Sormani [SS23] very recently obtained some results that are comparable to Theorem 6.25 and Theorem 6.26. We briefly discuss their setting and how it compares to ours.

The global causality encoding result [SS23, Thm. 4.1] of Sakovich and Sormani is formulated for spacetimes with *proper* locally anti-Lipschitz time functions, requiring that all time slabs $\tau^{-1}([\tau_1, \tau_2])$ with $[\tau_1, \tau_2] \subseteq \tau(M)$ are compact. The following argument shows that the level sets of a proper time function τ are (compact) *Cauchy* hypersurfaces and thus the spacetimes that Sakovich and Sormani consider are, in particular, *globally hyperbolic*: Suppose, for the sake of contradiction, that γ is an inextendible causal curve on M that does not intersect some τ-level set. Without loss of generality, suppose that $[0, 1] \in \tau(M)$ and that γ intersects $\{\tau = 0\}$ but not $\{\tau = 1\}$. Then the piece of γ lying in the compact set $\tau^{-1}([0, 1])$ is future inextendible, contradicting the fact that any spacetime with a time function is non-totally imprisoning. Hence γ must intersect every level set of τ.

Therefore, our global Theorem 6.25 is applicable to a wider class of spacetimes (also those having noncompact Cauchy surfaces) while the result [SS23, Thm. 4.1] of Sakovich and Sormani is applicable to a wider class of time functions (locally anti-Lipschitz instead of temporal).

In Section 6.3.3 we show that the assumption of Cauchy level sets can, in general, not be relaxed (see Remark 6.28 for a mild trivial extension). An example of Sakovich and Sormani [SS23, Ex. 2.2] that was constructed to show that noncompact level sets are problematic, in fact also already fails on a much more fundamental level because the spacetime is not globally hyperbolic.

Both proofs, that of Theorem 6.25 and that of Sakovich and Sormani [SS23, Thm. 4.1], rely on constructing a limit of a $\hat{d}_\tau$-minimizing sequence of piecewise causal curves, and showing that the limit is a (continuous) causal curve. Both proofs do this via the Arzelà–Ascoli Theorem, but while Sakovich and Sormani apply it using the null distance, we employ a complete Riemannian metric g_R (directly related to g via Wick-rotation, hence requiring τ temporal). This added regularity in our proof allows us to indeed obtain a locally Lipschitz (with respect to any Riemannian metric) limit curve β and compute $g(\dot{\beta}, \dot{\beta})$ explicitly. Sakovich and Sormani can work with locally anti-Lipschitz time functions by using special coordinate systems and not relying on the regularity of β to show that points are causally related. The important step to prove that the limit curve indeed reaches the desired endpoint is achieved by Sakovich and Sormani by the properness of τ (which implies that the whole sequence lies in a compact set) while we employ Cauchyness of the level sets (which can be non-compact, placing less restrictions on the spacetime, as discussed above).

In Theorem 6.9 we combine both approaches to obtain causality encodation for all locally anti-Lipschitz time functions with Cauchy level sets (neither required to be proper nor temporal). See Example 6.29 for a physically relevant case where our result applies.

Proof of Theorem 6.9. We use the notation of Theorem 6.25 and sequence $(\beta_n)_n$ satisfying (6.24). Note that Step 1 in the proof of Theorem 6.25 does not require any specific property of τ either, so (6.25) also holds. We can carry out Step 2 with respect to any complete Riemannian metric h on M (instead of g_R). Thus by the Arzelà–Ascoli Theorem we obtain a locally Lipschitz limit curve $\beta\colon [0,\infty) \to M$ from p.

Since d_h and $\hat{d}_\tau$ both induce the manifold topology, the uniform convergence with respect to d_h on compact subintervals implies pointwise convergence $\beta_n(t) \to \beta(t)$ with respect to the null distance $\hat{d}_\tau$ for all $t \in [0,\infty)$ as $n \to \infty$. Since the induced length structure of $\hat{d}_\tau$, i.e.,

$$L_{\hat{d}_\tau}(\gamma) = \sup \left\{ \sum_{i=1}^{k} \hat{d}_\tau(\gamma(s_i), \gamma(s_{i-1})) \,|\, a = s_0 < s_1 < \ldots < s_k = b \right\}$$

for rectifiable paths $\gamma\colon [a,b] \to M$, is lower semicontinuous [BBI01, Prop. 2.3.4] and agrees with $\hat{L}_\tau$ on the class of piecewise causal curves [AB22, Prop. 3.8] we obtain

$$L_{\hat{d}_\tau}(\beta|_{[s,t]}) \leq \lim_{n\to\infty} L_{\hat{d}_\tau}(\beta_n|_{[s,t]}) = \lim_{n\to\infty} \hat{L}_\tau(\beta_n|_{[s,t]}).$$

Together with property (6.25) and the continuity of τ we thus have that

$$L_{\hat{d}_\tau}(\beta|_{[s,t]}) \leq \tau(\beta(t)) - \tau(\beta(s)) \leq \hat{d}_\tau(\beta(s), \beta(t)) \leq L_{\hat{d}_\tau}(\beta|_{[s,t]}).$$

Hence β is not only a $\hat{d}_\tau$-minimizing curve but also satisfies for all $s, t \in [0,\infty)$

$$L_{\hat{d}_\tau}(\beta|_{[s,t]}) = \tau(\beta(t)) - \tau(\beta(s)).$$

As in the proof of Sakovich and Sormani [SS23, Thm. 4.3] their local causality encoding property thus implies that β is future-directed causal as continuous curve (in the sense of [HE73, p. 184]) starting at p. This completes Step 3.

It remains to be shown that β reaches q (in fact, at this point it could still be constant p). We proceed in a similar fashion as in Step 4 of the proof of Theorem 6.25 and distinguish the two cases (i) $\tau(\beta(s_0)) = \tau(q)$ for some $s_0 \in [0,\infty)$ and (ii) $\tau(\beta(s)) < \tau(q)$ for all s:

(i) extends verbatim (replacing d_R by d_h) and implies that $\beta(s_0) = q$.

(ii) requires us to prove that b_n is uniformly bounded from above. Note that, again with respect to the arbitrary complete Riemannian metric h chosen for the convergence in Step 2, by [Bur15, Thm. 4.11],

$$\begin{aligned} b_n &= L_h(\beta_n|_{[0,b_n]}) \\ &= \sup\left\{ \sum_{i=1}^{k} d_h(\beta_n(s_i), \beta_n(s_{i-1})) \,|\, 0 = s_0 < s_1 < \ldots < s_k = b_n \right\}. \end{aligned}$$

Since τ is locally anti-Lipschitz, by Step 1 in the proof of Lemma 6.16, for every point in M there exists a neighborhood U and a constant $C > 0$ such that

$$d_h(x,y) \le C\hat{d}_\tau(x,y) \tag{6.37}$$

for all $x, y \in U$. By the (reverse) argument in the proof of Theorem 6.7 on page 110, we can even assume that C is such that (6.37) holds on the entire compact set $\overline{W}$. Since all $\beta_n(s_i) \in \overline{W}$ by construction, and again by [AB22, Prop. 3.8], we have

$$\begin{aligned} b_n &\le \sup\left\{ \sum_{i=1}^{k} C\hat{d}_\tau(\beta_n(s_i), \beta_n(s_{i-1})) \,|\, 0 = s_0 < s_1 < \ldots < s_k = b_n \right\} \\ &\le CL_{\hat{d}_\tau}(\beta_n|_{[0,b_n]}) = C\hat{L}_\tau(\beta_n|_{[0,b_n]}) \le C(\tau(q) - \tau(p) + 1). \end{aligned}$$

Proceeding again as in the proof of Theorem 6.25 yields a contradiction.

Therefore, β is a locally Lipschitz future-directed causal curve from p to q, and $q \in J^+(p)$. □

Remark 6.28 (Future/Past Cauchy level sets)**.** Inspecting the proofs of Theorems 6.25 and 6.9 one observes that the assumptions that the level sets of τ are Cauchy is only used in Step 4(ii). In fact, it is only needed that the level sets are *past Cauchy* because we construct the limiting curve β from p to q (and $\tau(p) < \tau(q)$). We could have equally well constructed the curve from q in which case we would have used that the level sets are *future Cauchy* (see definition in [AGH98, p. 315]). Since either case yields the desired result, τ having future (or past) Cauchy level sets is already sufficient for $\hat{d}_\tau$ to encode causality.

Remark 6.28 allows us to immediately prove *global* causality encodation for a large and physically relevant class of time functions for which *local* causality encodation was already shown by Sakovich and Sormani [SS23, Cor. 1.2].

Proof of Corollary 6.10. By Lemma 6.22 the regular cosmological time function $\tau\colon M \to (0,\infty)$ is weak temporal. By [AGH98, Prop. 2.6] the level sets of τ are future Cauchy, hence by Remark 6.28 the result follows from the proof of Theorem 6.9. $\square$

We conclude with a basic cosmological example for which Theorem 6.9 is directly applicable.

Example 6.29 (Milne model)**.** Recall that the $n+1$-dimensional Milne model (M,g) is a globally hyperbolic spacetime which can be viewed as the chronological future $I^+(0)$ of the origin in the Minkowski spacetime $\mathbb{R}^{1,n}$ (see, for instance, [Lin20]). The cosmological time function $\tau\colon M \to (0,\infty)$ is the Lorentzian distance d_g from the origin, i.e.,

$$\tau(p) := \sup_{q\in J^-(p)} d_g(p,q).$$

By Lemma 6.22 τ is a weak temporal function and the level sets of τ are the noncompact hyperboloids (which are Cauchy). Thus by our Theorem 6.9 we see that $\hat{d}_\tau$ encodes causality globally. Due to noncompactness the result [SS23, Thm. 4.1] is not applicable. Since, however, the Milne model can be viewed as a warped product when expressed in the right coordinates [Lin20, Eq. 1.1], causality encodation already follows from an earlier result of Sormani and Vega [SV16, Thm. 3.25].

Remark 6.30 (Local causality-encoding results)**.** One can extend the local Theorem 6.26 to locally anti-Lipschitz time functions along the same lines. Note that in this local result a neighborhood U of x is constructed on which *any* two points $p,q\in U$ can be compared as in (6.22), while in the local result [SS23, Thm. 1.1] of Sakovich and Sormani the point $p=x$ is fixed and only q can be chosen freely.

6.3.3 Counterexamples

We conclude this section with a series of examples that show that the local Theorem 6.26 with respect to temporal functions cannot be promoted to a global statement in the spirit of Theorem 6.25, even if the spacetime is globally hyperbolic, but the τ-level sets are not Cauchy (Example 6.34).

In order to better contextualize our examples, we also consider the K^+ relation. Recall that K^+ is defined as the (unique) smallest closed and transitive relation containing J^+. The definition of K^+ is due to Sorkin and Woolgar [SW96]. Furthermore, a spacetime (M,g) is called *K^+-causal* if the K^+ relation is antisymmetric, a condition later shown to be equivalent to stable causality by Minguzzi [Min09], and hence also equivalent to the existence of time function [Haw68; Min10].

Example 6.31 ($J^+ \subsetneq K^+ = \hat{R}^+_\tau$)**.** Allen and Burtscher constructed examples [AB22, Ex. 3.23, 3.24] by removing points and lines from Minkowski space for which causality is not encoded, i.e., $J^+ \neq \hat{R}^+_\tau$. Notably, K-causality is still encoded, meaning that $K^+ = \hat{R}^+_\tau$.

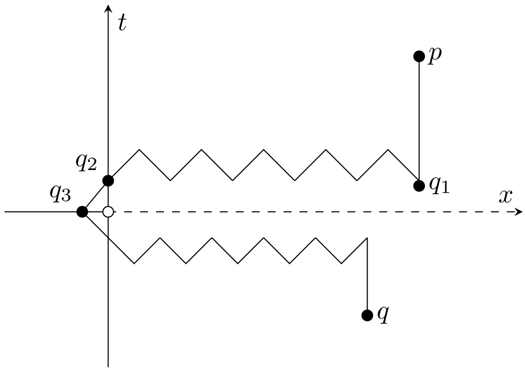


Figure 6.2: A piecewise causal curve that approximates the distance between the points p and q in Example 6.32, giving the upper bound (6.41).

Sormani and Vega gave another example [SV16, Prop. 3.4] for which one can check that $J^+ \subsetneq K^+ \subsetneq \hat{R}^+_\tau$. Their example is Minkowski space with the time function $\tau = t^3$, which is is not locally anti-Lipschitz, and the null distance is not definite. We modify said example to obtain a definite null distance for which neither J^+ nor K^+ are encoded.

Example 6.32 ($J^+ \subsetneq K^+ \subsetneq \hat{R}^+_\tau$)**.** Consider $M := \mathbb{R}^{1,1} \setminus \{(0,x) \mid x \geq 0\}$ equipped with the usual Minkowski metric $g := -dt^2 + dx^2$. We define the function $\tau \colon M \to \mathbb{R}$ by

$$\tau(t,x) := \begin{cases} t^3 & \text{if } x > 0, \\ t^3 + tx^2 & \text{if } x \leq 0. \end{cases}$$

First observe that τ is a temporal function on (M,g) because its gradient vector is timelike: For $x > 0$, this is trivial (note that then $t \neq 0$, by definition of M). For $x \leq 0$, since $(0,0) \notin M$, it follows from

$$\begin{aligned} g(\nabla\tau,\nabla\tau) &= -\left(\partial_t\tau\right)^2 + \left(\partial_x\tau\right)^2 \\ &= -\left(3t^2 + x^2\right)^2 + (2tx)^2 \\ &= -9t^4 - x^4 - 6t^2x^2 + 4t^2x^2 < 0. \end{aligned}$$

Next we show that every pair of points $p = (t_p, x_p)$, $q = (t_q, x_q)$ with $x_p, x_q > 0$ and $t_q < 0 < t_p$ satisfies

$$\hat{d}_\tau(p,q) = \tau(p) - \tau(q), \tag{6.38}$$

despite the fact that clearly not all such p and q are related by J^+ or K^+: Let $k \in \mathbb{N}$ be large enough so that $t_p > x_p/k$ and $|t_q| > x_q/k$. Consider the points (see Figure 6.2)

$$q_1 := (x_p/k, x_p), \quad q_2 := (x_p/k, 0), \quad q_3 := (0, -\min\{x_p, x_q\}/k).$$

Since $p \in J^+(q_1)$ and $q_2 \in J^+(q_3)$ trivially

$$\hat{d}_\tau(p, q_1) = t_p^3 - \left(\frac{x_p}{k}\right)^3, \qquad \hat{d}_\tau(q_2, q_3) = \left(\frac{x_p}{k}\right)^3. \tag{6.39}$$

Moreover, taking a piecewise null curve from q_1 to q_2 that consists of $2k$ segments between $t = x_p/k$ and $t = 2x_p/k$ (see Figure 6.2), we get that

$$\hat{d}_\tau(q_1, q_2) \leq 2k\left(\left(\frac{2x_p}{k}\right)^3 - \left(\frac{x_p}{k}\right)^3\right) = 14\frac{x_p^3}{k^2}. \tag{6.40}$$

Combining (6.39) and (6.40) with the triangle inequality proves

$$\hat{d}_\tau(p, q_3) \leq t_p^3 + 14\frac{x_p^3}{k^2},$$

and by symmetry we obtain the analogous estimate for $\hat{d}_\tau(q, q_3)$, thus

$$\hat{d}_\tau(p, q) \leq \hat{d}_\tau(p, q_3) + \hat{d}_\tau(q_3, q) \leq t_p^3 + 14\frac{x_p^3 + x_q^3}{k^2} + |t_q|^3, \tag{6.41}$$

and taking the limit $k \to \infty$ implies (6.38), as desired.

In the previous example, we have constructed a minimizing sequence of piecewise causal curves by choosing curves that are close to a "barrier" (the removed positive x axis). This barrier, however, also makes the spacetime causally discontinuous. Since the time function τ is perfectly regular (it is C^1 with timelike gradient, and can easily be smoothened out), one might suspect that causal discontinuity is the reason that causality is not encoded by $\hat{d}_\tau$ in Example 6.32. This motivates the next example, where we construct a causally simple spacetime with a temporal function τ but causality is still not encoded in $\hat{d}_\tau$. Recall that causal simplicity means that the causal relation J^+ is antisymmetric and closed, and sits only one step below global hyperbolicity on the causal ladder. In order to achieve this effect, instead of approaching a barrier, we construct an example with a minimizing sequence of piecewise null curves that runs off to infinity.

Example 6.33 ($J^+ = K^+ \subsetneq \hat{R}_\tau^+$)**.** Let $M := \mathbb{R}^3$ with coordinates (t, x, y) and warped product metric tensor

$$g := \cosh^2(x)\left(-dt^2 + dy^2\right) + dx^2. \tag{6.42}$$

Note that all $\{x = x_0\}$ planes are conformal to Minkowski space, while each $\{y = y_0\}$ plane is isometric to the universal cover of AdS^2. Therefore, the induced null geodesics $s \mapsto (t_\pm(s), x_\pm(s))$ in the $\{y = y_0\}$ plane going through a point $(t_0, 0)$ are given by

$$\begin{aligned} t_\pm(s) &= 2\arctan\left(\tanh\left(\frac{s}{2}\right)\right) + t_0, \\ x_\pm(s) &= \pm s, \end{aligned} \tag{6.43}$$

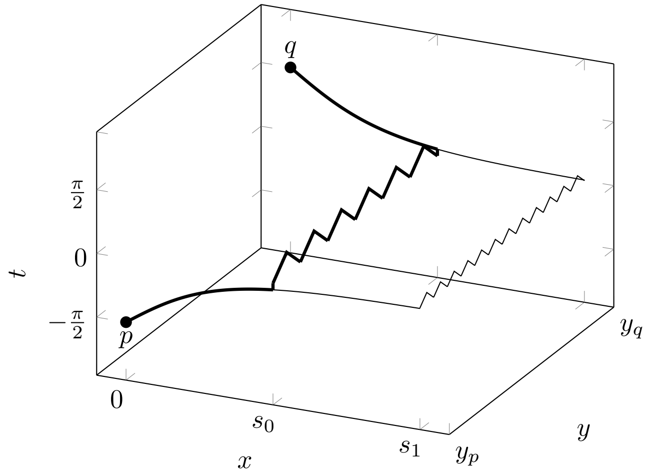


Figure 6.3: Two piecewise causal curves going trough the points p_{s_0} and q_{s_0} (thick) and p_{s_1} and q_{s_1} (thin), respectively, that approximate the null distance between p and q in Example 6.33. The thin curve yields a better approximation.

the subscript $+$ or $-$ indicating the right- or left-going geodesics respectively [Lan21, Sec. 5.10]. Moreover, the function

$$\tau(t, x, y) := \cosh^{-1}(x)t + t^3,$$

is a steep temporal function on (M, g) since

$$\begin{aligned} g(\nabla\tau, \nabla\tau) &= -\cosh^2(x)\left(\partial_t\tau\right)^2 + \left(\partial_x\tau\right)^2 \\ &= -\cosh^2(x)\left(\cosh^{-1}(x) + 3t^2\right)^2 + \left(-\cosh^{-1}(x)\tanh(x)t\right)^2 \\ &= -1 - \cosh^2(x)9t^4 - \left(6\cosh(x) - \cosh^{-2}(x)\tanh^2(x)\right)t^2 \leq -1, \end{aligned}$$

where in the last line, we have used that $|\tanh(x)| < 1 \leq \cosh(x)$ for all $x \in \mathbb{R}$.

Consider now two points $p, q \in M$ of the form

$$p := \left(-\frac{\pi}{2},\ 0,\ y_p\right), \qquad\qquad q := \left(\frac{\pi}{2},\ 0,\ y_q\right),$$

where $|y_p - y_q| > \pi$. We are going to show that $(p, q) \in \hat{R}^+_\tau$, i.e., that

$$\hat{d}_\tau(p, q) = \tau(q) - \tau(p), \tag{6.44}$$

despite the fact that $(p, q) \notin J^+$ (and since our spacetime is causally simple, $J^+ = K^+$). To see that $(p, q) \notin J^+$, note that by definition of g in (6.42) the $\{x = 0\}$ plane is isometric to $1 + 1$-Minkowski space, and that the projection $(t, x, y) \mapsto (t, y)$ of any causal curve from p to q in M to $\{x = 0\}$ remains causal in $\{x = 0\}$. The condition $|y_p - y_q| > \pi$ then implies that $(p, q) \notin J^+_{\{x=0\}}$, and therefore also $(p, q) \notin J^+$.

In order to show (6.44), we proceed in three steps. First, define

$$p_s = (0, s, y_p), \qquad\qquad q_s = (0, s, y_q),$$

where $s \in \mathbb{R}$ is arbitrary. By (6.43), a null geodesic in the $\{y = y_p\}$ plane starting at p eventually reaches the point $(t_+(s), s)$ where $t_+(s) < 0$. It follows that $(p, p_s) \in J^+$, and by the time reversed argument on the $\{y = y_q\}$ plane, that $(q_s, q) \in J^+$, altogether implying that

$$\hat{d}_\tau(p, p_s) = \tau(p_s) - \tau(p) = -\tau(p), \qquad \hat{d}_\tau(q_s, q) = \tau(q) - \tau(q_s) = \tau(q). \tag{6.45}$$

The second step is to estimate the null distance between p_s and q_s. Given that the $\{x = s\}$ plane is conformal to Minkowski space and the null distance is conformally invariant it is easy to construct piecewise causal curves between p_s and q_s. The null distance induced on each plane is different though, and since $\tau|_{\{x=x_s\}} \to t^3$ as $s \to \infty$ we have at the "boundary" of our spacetime an indefinite null distance that cannot distinguish any points in the $\{t = 0\}$ slice [SV16, Prop. 3.4]. We make this intuitive picture precise by constructing in each $\{x = x_s\}$ plane a piecewise null curve $\beta_{s,k}$ that bounces k times between $t = 0$ and $t = |y_p - y_q|/k$ (see Figure 6.3). Using the curves $\beta_{s,k}$ to estimate the null distance, we obtain the upper bound

$$\begin{aligned}
\hat{d}_\tau(p_s, q_s) &\leq \liminf_{k\to\infty} \hat{L}_\tau(\beta_{s,k}) \\
&= \liminf_{k\to\infty} 2k \left(\frac{|y_p - y_q|^3}{k^3} + \cosh^{-1}(s) \frac{|y_p - y_q|}{k} \right) \\
&= 2\cosh^{-1}(s)\, |y_p - y_q| \, .
\end{aligned}$$

Finally, the triangle inequality together with (6.45) and the above estimate yields

$$\hat{d}_\tau(p, q) \leq \lim_{s\to\infty} \left(\hat{d}_\tau(p, p_s) + \hat{d}_\tau(p_s, q_s) + \hat{d}_\tau(q_s, q) \right) = \tau(q) - \tau(p).$$

This finishes the proof of (6.44), since the opposite inequality is always true.

Note that in our proof, $|y_p - y_q|$ can be chosen arbitrarily large while $\hat{d}_\tau(p, q) = \pi + \frac{\pi^3}{4}$ remains the same. Therefore the $\hat{d}_\tau$-ball at p of radius $R > \pi + \frac{\pi^3}{4}$ is unbounded with respect to the usual Euclidean distance.

Given that in the causal ladder of spacetimes, causal simplicity comes just before global hyperbolicity, the previous Example 6.33 shows that the assumptions in Theorem 6.25 are sharp in view of the causal structure required. The following and final example shows that even on a globally hyperbolic spacetime, Cauchyness of the time function cannot simply be dropped in Theorem 6.9. This is not surprising, because non-Cauchy temporal functions on globally hyperbolic spacetimes can have a much wilder behaviour than their Cauchy counterparts, such as topology changes of the level sets [Sán23].

Example 6.34 ($J^+ = K^+ \subsetneq \hat{R}^+_\tau$ for non-Cauchy locally anti-Lipschitz time function in globally hyperbolic spacetime)**.** We show that the Cauchy assumption

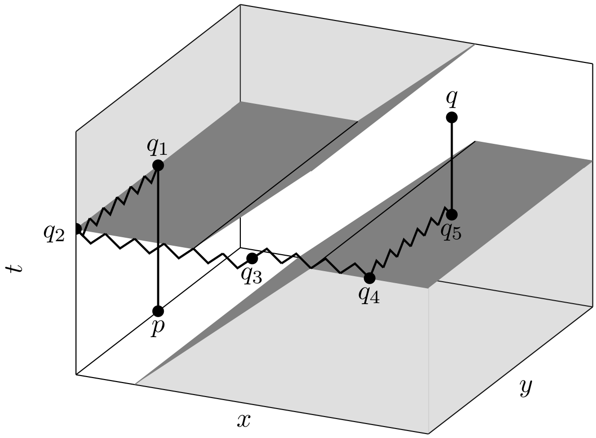


Figure 6.4: A piecewise causal curve that approximates the null distance between p and q in Example 6.34. Only the space in between the two grey surfaces is part of the spacetime. The curve comes ε-close to the boundary, but remains entirely within M.

in Theorem 6.9 cannot be relaxed. To this end we construct an example that combines aspects of Examples 6.32 and 6.33 in the sense that a $\hat{d}_\tau$-minimizing sequence of piecewise causal curves between certain points approaches both a barrier (in the x direction) *and* runs off to infinity (in the y direction).

The spacetime under consideration is

$$\begin{aligned} M := &\{(t,x,y) \mid t>0, x>t-1\} \\ &\cup \{(t,x,y) \mid t<0, x<t+1\} \\ &\cup \{(0,x,y) \mid -1<x<1\} \subseteq \mathbb{R}^{1,2}, \end{aligned}$$

considered as subset of the $(2+1)$-dimensional Minkowski space with metric $g := -dt^2+dx^2+dy^2$ (see Figure 6.4). Clearly, M is globally hyperbolic. We equip M with the continuous function

$$\tau(t,x,y) := t^3 + \Psi(t,x)\cosh^{-1}\left(\frac{y}{2}\right),$$

where

$$\Psi(t,x) := \begin{cases} \sqrt{(t+1)^2-x^2} & \text{if } |t+1|>|x|, \\ 0 & \text{otherwise.} \end{cases}$$

We show that (i) τ is a time function for (M,g) and (ii) that the corresponding null distance $\hat{d}_\tau$ does not encode causality globally. Theorem 6.9 thus implies that τ is not Cauchy, as can also be seen by considering causal curves in Minkowski spacetime which leave the region M.

(i) We show that the gradient vector field $\nabla\tau$ is timelike almost everywhere, which is a sufficient condition for being a time function. In the region where

$\Psi = 0$, our function is simply t^3, and since said region does not include $\{t = 0\}$, we have that $\nabla\tau$ is timelike there. It remains to consider the region where $\Psi \neq 0$. There the gradient of τ is given by

$$\begin{aligned}\nabla\tau = &-\left(3t^2 + \frac{t+1}{\sqrt{(t+1)^2 - x^2}}\cosh^{-1}\left(\frac{y}{2}\right)\right)\partial_t \\ &- \frac{x}{\sqrt{(t+1)^2 - x^2}}\cosh^{-1}\left(\frac{y}{2}\right)\partial_x \\ &- \frac{1}{2}\sqrt{(t+1)^2 - x^2}\cosh^{-1}\left(\frac{y}{2}\right)\tanh\left(\frac{y}{2}\right)\partial_y,\end{aligned}$$

and therefore its norm is

$$\begin{aligned}g(\nabla\tau,\nabla\tau) = &-9t^4 - 3t^2\frac{t+1}{\sqrt{(t+1)^2 - x^2}}\cosh^{-1}\left(\frac{y}{2}\right) \\ &- \cosh^{-2}\left(\frac{y}{2}\right) + \frac{1}{4}\left((t+1)^2 - x^2\right)\cosh^{-2}\left(\frac{y}{2}\right)\tanh^2\left(\frac{y}{2}\right).\end{aligned}$$

On the RHS, the terms on the first line are always negative (since $t+1 > 0$ in the region we are considering). If $t \leq 1$, then the second line is also negative, since then $(t+1)^2 - x^2 \leq 4$ and $|\tanh(z)| < 1$. If, on the other hand, $t \geq 1$, then the $9t^4$ term dominates the whole expression (since also $\cosh(z) > 1$). In either case, we have shown that $g(\nabla\tau,\nabla\tau) < 0$, as desired.

(ii) It remains to be shown that causality is not encoded in the null distance. Concretely, we show that for all $p = (t_p, x_p, y_p)$ and $q = (t_q, x_q, y_q)$ with $t_p < -2 < 2 < t_q$ and $x_p < t_p + 1$, $x_q > t_q + 1$,

$$\hat{d}_\tau(p,q) = \tau(q) - \tau(p),$$

despite the fact that clearly not every two such points are causally related. The argument is depicted in Figure 6.4, and we omit some computations that are analogous to the ones in the previous examples. Choose $\varepsilon > 0$ and follow a causal segment from p to $q_1 = (-\varepsilon, x_p, y_p)$, so that

$$\hat{d}_\tau(p,q_1) = \tau(p) - \varepsilon^3.$$

Note that q_1 lies in a region where $\tau = t^3$. Therefore, given any (arbitrarily large) $R > 0$, for $q_2 = (-\varepsilon, x_p, R)$ we have

$$\hat{d}_\tau(q_1,q_2) \to 0 \text{ as } \varepsilon \to 0,$$

similarly to the situation in Example 6.32. Next, let $q_3 = (0, 0, R)$. Then

$$\hat{d}_\tau(q_2,q_3) \sim \varepsilon \text{ for } R \text{ such that } \cosh^{-1}(R) \leq \varepsilon^2,$$

similar to what happens in Example 6.33. Finally, do a similar procedure backwards to get from q_3 to $q_4 = (\varepsilon, x_q, R)$ to $q_5 = (\varepsilon, x_q, y_q)$ (with arbitrarily small length) and then to q (with length $\tau(q) - \varepsilon^3$).

In conclusion, by choosing $R(\varepsilon)$ such that $\cosh^{-1}(R(\varepsilon)) \leq \varepsilon^2$ (as required above), we have that

$$\lim_{\varepsilon \to 0} \sum_{i=1}^{4} \hat{d}_\tau(q_i, q_{i+1}) = 0,$$

and by the triangle inequality

$$\hat{d}_\tau(p,q) \leq \lim_{\varepsilon \to 0} \left[\hat{d}_\tau(p, q_1) + \sum_{i=1}^{4} \hat{d}_\tau(q_i, q_{i+1}) + \hat{d}_\tau(q_5, q) \right] = \tau(q) - \tau(p),$$

as claimed (the opposite inequality always holds).

We end this section noting that the temporal function in Example 6.33 is *steep*, a notion already discussed at the beginning of Section 6.2.1. Since any temporal function is steep for a conformal transformation of g (which leaves the null distance invariant) steepness is unrelated to causality encodation. The situation is different for completely uniform temporal functions (also called h-steep), because they are special Cauchy temporal functions. We define and make use of them in the following section.

6.4 Completeness

In this final section we prove our Main Theorem 6.4 which characterizes global hyperbolicity of (M, g) by metric completeness of $(M, \hat{d}_\tau)$. Completeness is a global property, therefore we cannot expect $(M, \hat{d}_\tau)$ to be complete for all choices of τ (even though they are locally equivalent by Section 6.2). In Section 6.3 we have observed that locally anti-Lipschitz Cauchy functions encode causality globally. Therefore, it comes at no surprise that this is a necessary ingredient for a completeness result. We use the following class of time functions recently introduced by Bernard and Suhr [BS18; BS20] to study closed cone fields.

Definition 6.35. Let (M, g) be a spacetime. A smooth function $\tau \colon M \to \mathbb{R}$ is called a *completely uniform temporal function* if there exists a complete Riemannian metric h on M such that for all causal vectors $v \in TM$

$$d\tau(v) \geq \|v\|_h. \tag{6.46}$$

We call τ a *completely uniform weak temporal function* if it is weak temporal and (6.46) holds almost everywhere.

Originally these functions were called steep with respect to a (complete) Riemannian metric in [BS18, p. 473] and later renamed in [BS20, Def. 1.2]. Subsequently, f-steep functions with respect to any positive homogenous C^1 function (not just $f = \|.\|_h$ for h a complete Riemannian metric) were also used by Minguzzi [Min18, p. 2] in the analysis of Lorentz–Finsler spaces.

It was shown by Bernard and Suhr [BS18, Thm. 3] and later also by Minguzzi [Min18, Thm. 3.1] that the existence of completely uniform temporal function is equivalent to global hyperbolicity of the spacetime.

We prove the following refined version of Theorem 6.4.

Theorem 6.36. *Let (M, g) be a spacetime.*

(i) If τ is a time function such that $(M, \hat{d}_\tau)$ is a complete metric space, then τ is a Cauchy time function. In particular, (M, g) is globally hyperbolic.

(ii) If (M, g) is globally hyperbolic then there exists a completely uniform weak temporal function τ, and for every such τ, $(M, \hat{d}_\tau)$ is a complete metric space.

Theorem 6.4 is a direct corollary of Theorem 6.36.

Proof. (i) Assume that $(M, \hat{d}_\tau)$ is complete but τ is not a Cauchy time function. Then there exists, without loss of generality, a future-directed future-inextendible causal curve $\gamma\colon \mathbb{R} \to M$ such that $\lim_{s\to\infty} \tau(\gamma(s)) < \infty$. Consider the sequence $(p_n)_n$ of points given by $p_n = \gamma(n)$. Since the p_n are causally related among each other $\hat{d}_\tau(p_n, p_m) = |\tau(p_n) - \tau(p_m)|$. Then the fact that $\tau \circ \gamma\colon \mathbb{R} \to \mathbb{R}$ is strictly increasing and bounded from above implies that $(p_n)_n$ is a Cauchy sequence in $(M, \hat{d}_\tau)$. By completeness there exists a limit point p, and since γ is continuous, $p \in \overline{\gamma}$, a contradiction to the inextendibility of γ. Hence τ must be a Cauchy time function, and (M, g) globally hyperbolic by Theorem 6.3.

(ii) By [BS18; Min18] (M, g) is globally hyperbolic if and only if there is a completely uniform temporal function τ which with respect to a complete Riemannian metric h satisfies (6.46). We show that any such (even only weak temporal) τ is anti-Lipschitz with respect to the (complete) distance d_h induced by h, i.e., there is a $C > 0$ such that for all $p, q \in M$

$$(p, q) \in J^+ \implies \tau(q) - \tau(p) \geq C d_h(p, q). \tag{6.47}$$

Pick any $q \in J^+(p)$ and $\gamma\colon [0, 1] \to M$ a causal curve with $\gamma(0) = p$, $\gamma(1) = q$. Then by (6.46)

$$\begin{aligned} \tau(q) - \tau(p) &= \int_0^1 d\tau\left(\dot\gamma(s)\right) ds \\ &\geq \int_0^1 \sqrt{h\left(\dot\gamma(s), \dot\gamma(s)\right)} ds = L_h(\gamma) \geq d_h(p, q). \end{aligned}$$

Thus (6.47) holds globally, and a theorem of Allen and Burtscher [AB22, Thm. 1.6] implies that $(M, \hat{d}_\tau)$ is complete (and definite). ☐

Remark 6.37. Recall that $(M, \hat{d}_\tau)$ is always a locally compact length-metric space [AB22, Thm. 1.1]. If $(M, \hat{d}_\tau)$ is also complete, the Hopf–Rinow–Cohn-Vossen Theorem implies that any pair of points can be joined by a $\hat{d}_\tau$-length minimizing curve. Beware that the minimizer is, in general, only $\hat{d}_\tau$-rectifiable, but not necessarily piecewise causal [AB22, Ex. 3.17]. If τ is besides Cauchy also locally anti-Lipschitz, thanks to Theorem 6.9, we still do know that the null distance between two points is their difference in time precisely when there is a causal curve between them.

Applying Theorem 6.25(ii) and then (i) proves the following result, originally shown by Bernard and Suhr [BS18, Thm. 3] for temporal functions (see also [BS20, Lemma 1.3]).

Corollary 6.38. *If a weak temporal function τ is completely uniform, then τ is Cauchy.* □

Since the cosmological time function does not attain negative values it is not Cauchy, and hence by Theorem 6.4(i) the corresponding null distance is not complete. We conclude our paper with a counterexample that shows that non-completely uniform temporal functions on globally hyperbolic spacetimes do, in general, not imply metric completeness.

Example 6.39 (Cauchy temporal function with incomplete null distance)**.** In [Sán22, Sec. 6.4], Sánchez constructs a globally hyperbolic spacetime $(M, g) = (\mathbb{R}^2, -dt^2 + f^2(t,x)dx^2)$ with a certain piecewise defined L^1-function $f\colon M \to (0,\infty)$ and such that t is a Cauchy temporal function, but the Riemannian slice $\{t = 0\} = (\mathbb{R}, f^2(0,x)dx^2)$, that is, the Riemannian manifold $(\mathbb{R}, f^2(0,x)dx^2)$, is geodesically incomplete. Let $(0,x)$, $(0,y)$ be two points on the $\{t = 0\}$ slice. Then we can estimate their null distance by a sequence of piecewise null curves $\gamma_n(s) = (\gamma_n^t(s), s)$ satisfying $0 \leq \gamma_n^t(s) \leq \frac{1}{n}$. We obtain

$$\hat{d}_t((0,x),(0,y)) \leq \hat{L}_t(\gamma_n) = \int_x^y |\dot{\gamma}_n^t(s)| ds = \int_x^y f(\gamma_n^t(s), s) ds.$$

Applying dominated convergence to the right hand side yields

$$\lim_{n\to\infty} \int_x^y f(\gamma_n^t(s), s) ds = \int_x^y f(0,s) ds \leq \|f(0,\cdot)\|_{L^1(\mathbb{R})} < \infty.$$

This implies that the sequence $(n)_n$ is Cauchy because for any $\varepsilon > 0$, assuming that $m \leq n$ sufficiently large,

$$\hat{d}_t((0,m),(0,n)) \leq \int_m^n f(0,s) ds \leq \int_m^\infty f(0,s) ds < \varepsilon.$$

The hypothetical limit point at ∞, however, is not in M. Therefore, $(M, \hat{d}_t)$ is incomplete.

Summary

Understanding cause-and-effect relationships lies at the heart of science. Naïvely, one would think that an event happening at a time t_0 can, in principle, influence any event happening at a later time $t > t_0$. However, this does not take into account the physical law that nothing can travel faster than the speed of light, not even causal influences. The fact that the speed of light is the same for all observers (even if moving with respect to one another) further complicates things. In spite of these issues, the theory of special relativity has successfully clarified our understanding of cause-and-effect relationships, thanks to its key insight that time is not absolute, but instead it is "mixed" with space, forming *spacetime*. In the theory of general relativity, spacetime is further upgraded from being a fixed background structure to having geometrical properties that influence, and are influenced by, the matter in our Universe. Notably, general relativity is able to describe the physical phenomenon of gravitation as a consequence of the geometrical curvature of spacetime.

In classical works on general relativity, spacetime is usually assumed to be smooth. In mathematical terms, spacetime is a manifold equipped with an smooth (meaning infinitely differentiable) Lorentzian metric tensor. Note that a Riemannian metric would be suitable to describe space alone, but not spacetime. The smoothness assumption, while useful to simplify things, is not clearly justified. It could well be that there is matter in our Universe that causes spacetime to be non-smooth, especially when we think of the quantum mechanical properties that matter has on small scales. In this thesis, it is investigated how robust our understanding of cause-and-effect relationships is when taking into account this possible non-smoothness of spacetime.

Different scenarios of what a non-smooth spacetime could be are considered. Examples are given of manifolds equipped with merely continuous Lorentzian metrics, which behave pathologically compared to smooth ones. For the Lorentzian length spaces of Kunzinger and Sämann (a Lorentzian version of the usual length spaces), existence results for time functions are proven. Manifolds with degenerate metric tensors are also investigated, leading to progress on a conjecture of Borde and Sorkin motivated by quantum gravity. Finally, it is established how the causal relationships on a (smooth) Lorentzian manifold can be encoded in a metric space structure, by using the null distance of Sormani and Vega.

Samenvatting

Het begrijpen van oorzaak-en-gevolgrelaties vormt de kern van de wetenschap. Naïef zou men denken dat een gebeurtenis die op een tijdstip t_0 plaatsvindt, in principe elke gebeurtenis op een later tijdstip $t > t_0$ kan beïnvloeden. Dit houdt echter geen rekening met de natuurkundige wet dat niets sneller kan reizen dan de snelheid van het licht, zelfs causale invloeden kunnen dit niet. Het feit dat de lichtsnelheid voor alle waarnemers gelijk is (zelfs als ze ten opzichte van elkaar bewegen) maakt de zaken nog ingewikkelder. Met het oog op deze kwesties heeft de speciale relativiteitstheorie ons begrip van oorzaak-en-gevolgrelaties verduidelijkt, dankzij het belangrijke inzicht dat tijd niet absoluut is, maar "vermengd" is met ruimte, waardoor *ruimtetijd* ontstaat. In de algemene relativiteitstheorie wordt ruimtetijd verder opgewaardeerd van een vaste achtergrondstructuur naar een geometrisch object dat de materie in ons universum beïnvloedt en erdoor wordt beïnvloed. Met name kan de algemene relativiteitstheorie de zwaartekracht beschrijven als gevolg van de geometrische kromming van de ruimtetijd.

In klassieke werken over algemene relativiteit wordt meestal aangenomen dat ruimtetijd glad is. In wiskundige termen is ruimtetijd een variëteit die is uitgerust met een gladde (wat oneindig vaak differentieerbaar betekent) Lorentz-metriek. Merk op dat een Riemann-metriek geschikt zou zijn om ruimte te beschrijven, maar niet ruimtetijd. De aanname van gladheid, hoewel nuttig om de theorie te vereenvoudigen, is niet duidelijk gerechtvaardigd. Het zou heel goed kunnen dat er materie in ons heelal is die ervoor zorgt dat ruimtetijd niet-glad is, vooral als we denken aan de kwantummechanische eigenschappen die materie op kleine schaal heeft. In dit proefschrift wordt onderzocht hoe robuust ons begrip van oorzaak-en-gevolgrelaties is als we rekening houden met deze mogelijke niet-gladheid van ruimtetijd.

Er zijn verschillende scenario's bekeken van wat een niet-gladde ruimtetijd zou kunnen zijn. Er worden voorbeelden gegeven van variëteiten die zijn uitgerust met slechts continue Lorentz-metrieken, die zich pathologisch gedragen in vergelijking met gladde. Voor de Lorentz-lengteruimten van Kunzinger en Sämann (een Lorentz-versie van de gebruikelijke lengteruimten) zijn bestaansresultaten voor tijdfuncties bewezen. Variëteiten met gedegenereerde metrische tensoren worden ook onderzocht, wat leidt tot vooruitgang op een vermoeden van Borde en Sorkin gemotiveerd door kwantumgravitatie. Ten slotte wordt vastgesteld hoe de causale relaties op een (gladde) Lorentz-variëteit kunnen worden gecodeerd in een metrische ruimtestructuur, door de nulafstand van Sormani en Vega te gebruiken.

Acknowledgements

I want to start by thanking my PhD advisor Annegret. First of all, for hiring me despite my limited background in mathematical general relativity: I am now extremely happy to be working in this field. Then, during my whole PhD trajectory, I have felt very much supported and listened to. I take away very positive memories, such as our road trips in Canada and our viewing sessions of online seminars. Annegret also helped me find an accommodation in Nijmegen (thanks also to Ruben for that, and to the de Kruijffs for helping me move). I am also very grateful to my promotor Klaas, who has always been there when we needed him. Being his teaching assistant for the singularities and black holes course has shaped my vision of mathematical relativity (thanks also to all our students!).

My first spontaneous collaboration happened with Teri, leading to Chapter 3, so I want to thank him as well. More generally, I have come to feel very welcome in the research community. In this regard, special thanks go to Lydia Bieri, Melanie Graf, Michael Kunzinger, Miguel Sánchez, Stefan Suhr and Eric Woolgar for inviting me to their respective universities and/or helping me during my hunt for postdoc positions. I am also indebted to the organizers and participants of the many conferences and workshops that I attended during the last four years, especially those of the thematic program on nonsmooth Riemannian and Lorentzian geometry at the Fields Institute in Toronto. In addition to the two members already mentioned above, I am also grateful to Robert McCann, Renate Loll and Anna Sakovich for being on the manuscript committee, and Gert Heckmann, Ioan Mărcuț and Walter van Suijlekom for joining the examination board.

I have become very fond of Nijmegen during my time here, and the truth is that it is because of the people. Therefore, many thanks go to all my friends and colleagues: everyone at the math department, the bouldering team, the book club, the vierdaagse walkers and supporters, my not-stressed friends, and my friends in Utrecht. In particular, I want to mention Mostafa for being at the intersection of almost all of the above, Tommy for being my first friend here, and Gastón and Jordi, who will be my paranymphs. During these four years, I have also spent quite some time back home, and am grateful to the *pseudo-físics*, to all my new and old friends in Barcelona and Igualada, and to Agathe for lending me the bunker.

Last but not least, I want to thank Vedika for all the fun times, the updates on the local fauna, and the emotional support during some stressful moments while finishing the thesis. And, of course, my family, who have always been there for me: *Us estimo! ¡Os quiero! Ich hab euch lieb!*

About the author

Leonardo García Heveling was born on the 10th of January 1995 in Vilassar de Mar to a German mother and a Spanish father. As a child growing up in rural Catalonia, he wanted to become an engineer, thinking of it as the grown-up version of building with LEGO. While attending secondary school at Pere Vives Vich in Igualada, he became more interested in fundamental science, and later enrolled in the physics bachelor at the University of Barcelona. There he also developed a passion for mathematics, leading him to complete a minor in this subject, and later to continue his trajectory with the theoretical and mathematical physics master at the Ludwig-Maximilians-Universität München.

In 2019, Leonardo started as a doctoral student at the Mathematics Department of the Radboud University, with Annegret Burtscher as his PhD advisor. During the fall of 2022, he stayed for three months at the Fields Institute for Research in the Mathematical Sciences (Toronto), attending the thematic program on nonsmooth Riemannian and Lorentzian geometry. At the Dutch mathematical congress of 2023, he was awarded the KWG PhD prize of the Royal Dutch Mathematical Society. Leonardo has also participated in the *vierdaagse* marches in Nijmegen, obtaining the cross for completing the 200 km distance.

After defending his PhD thesis, Leonardo will be moving to a postdoc position in the group of Melanie Graf at the University of Hamburg.

Research data management

This thesis research has been carried out under the institute research data management policy of IMAPP, Radboud University. No data has been produced or analysed in this project.

Bibliography

[AH+22] L. Ake Hau et al. "Causal completions as Lorentzian pre-length spaces". *Gen. Rel. Grav.* 54 (9) (2022). [Erratum: Gen.Rel.Grav. 54, 141 (2022)], 108.

[ACS20] L. Aké Hau, A. J. Cabrera Pacheco, and D. A. Solis. "On the causal hierarchy of Lorentzian length spaces". *Classical Quantum Gravity* 37 (21) (2020), 215013.

[Ale+19] S. B. Alexander et al. "Generalized cones as Lorentzian length spaces: Causality, curvature, and singularity theorems" (2019). To appear in *Comm. Anal. Geom.* arXiv: `1909.09575 [math.MG]`.

[AB22] B. Allen and A. Burtscher. "Properties of the null distance and spacetime convergence". *Int. Math. Res. Not. IMRN* (10) (2022), 7729–7808.

[AJL00] J. Ambjørn, J. Jurkiewicz, and R. Loll. "A nonperturbative Lorentzian path integral for gravity". *Phys. Rev. Lett.* 85 (2000), 924–927.

[AL98] J. Ambjørn and R. Loll. "Nonperturbative Lorentzian quantum gravity, causality and topology change". *Nucl. Phys. B* 536 (1998), 407–434.

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Second. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008, x+334.

[AGS14] L. Ambrosio, N. Gigli, and G. Savaré. "Metric measure spaces with Riemannian Ricci curvature bounded from below". *Duke Math. J.* 163 (7) (2014), 1405–1490.

[AK00] L. Ambrosio and B. Kirchheim. "Currents in metric spaces". *Acta Math.* 185 (1) (2000), 1–80.

[AD86] A. Anderson and B. S. DeWitt. "Does the Topology of Space Fluctuate?" *Found. Phys.* 16 (1986), 91–105.

[AGH98] L. Andersson, G. J. Galloway, and R. Howard. "The cosmological time function". English. *Classical Quantum Gravity* 15 (2) (1998), 309–322.

[Ave63] A. Avez. "Essais de géométrie riemannienne hyperbolique globale. Applications à la relativité générale". *Ann. Inst. Fourier (Grenoble)* 13 (fasc. 2) (1963), 105–190.

[BGP07] C. Bär, N. Ginoux, and F. Pfäffle. *Wave equations on Lorentzian manifolds and quantization.* English. Zürich: European Mathematical Society Publishing House, 2007, viii + 194.

[BOS22] W. Barrera, L. M. de Oca, and D. A. Solis. "Comparison theorems for Lorentzian length spaces with lower timelike curvature bounds". *Gen. Rel. Grav.* 54 (9) (2022), 107.

[BOS23] W. Barrera, L. M. de Oca, and D. A. Solis. "On the space of compact diamonds of Lorentzian length spaces" (2023). arXiv: `2302.11819 [math-ph]`.

[Bar88] R. Bartnik. "Remarks on cosmological spacetimes and constant mean curvature surfaces". *Comm. Math. Phys.* 117 (4) (1988), 615–624.

[Bee76] J. K. Beem. "Globally hyperbolic space-times which are timelike Cauchy complete". *Gen. Relativity Gravitation* 7 (4) (1976), 339–344.

[BE85] J. K. Beem and P. E. Ehrlich. "Geodesic completeness of submanifolds in Minkowski space". *Geom. Dedicata* 18 (2) (1985), 213–226.

[BEE96] J. K. Beem, P. E. Ehrlich, and K. L. Easley. *Global Lorentzian geometry.* Second. Vol. 202. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 1996.

[BNR23] T. Beran, L. Napper, and F. Rott. "Alexandrov's Patchwork and the Bonnet-Myers Theorem for Lorentzian length spaces" (2023). arXiv: `2302.11615 [math.DG]`.

[BR22] T. Beran and F. Rott. "Gluing constructions for Lorentzian length spaces" (2022). arXiv: `2201.09695 [math.DG]`.

[BS22] T. Beran and C. Sämann. "Hyperbolic angles in Lorentzian length spaces and timelike curvature bounds" (2022). arXiv: `2204.09491 [math.DG]`.

[Ber+22] T. Beran et al. "The splitting theorem for globally hyperbolic Lorentzian length spaces with non-negative timelike curvature" (2022). arXiv: `2209.14724 [math.DG]`.

[BS03] A. N. Bernal and M. Sánchez. "On smooth Cauchy hypersurfaces and Geroch's splitting theorem". *Comm. Math. Phys.* 243 (3) (2003), 461–470.

[BS05] A. N. Bernal and M. Sánchez. "Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes". *Comm. Math. Phys.* 257 (1) (2005), 43–50.

[BS06a] A. N. Bernal and M. Sánchez. "Further results on the smoothability of Cauchy hypersurfaces and Cauchy time functions". *Lett. Math. Phys.* 77 (2) (2006), 183–197.

[BS07] A. N. Bernal and M. Sánchez. "Globally hyperbolic spacetimes can be defined as 'causal' instead of 'strongly causal'". *Classical Quantum Gravity* 24 (2007), 745–750.

[BS18] P. Bernard and S. Suhr. “Lyapounov functions of closed cone fields: from Conley theory to time functions”. *Comm. Math. Phys.* 359 (2) (2018), 467–498.

[BS20] P. Bernard and S. Suhr. “Cauchy and uniform temporal functions of globally hyperbolic cone fields”. *Proc. Amer. Math. Soc.* 148 (11) (2020), 4951–4966.

[BN04] L. Bombelli and J. Noldus. “The moduli space of isometry classes of globally hyperbolic space-times”. *Classical Quantum Gravity* 21 (2004), 4429–4454.

[Bom+87] L. Bombelli et al. “Space-Time as a Causal Set”. *Phys. Rev. Lett.* 59 (1987), 521–524.

[BS90] H.-J. Borchers and R. N. Sen. “Theory of ordered spaces”. *Comm. Math. Phys.* 132 (3) (1990), 593–611.

[BS06b] H.-J. Borchers and R. N. Sen. *Mathematical implications of Einstein-Weyl causality.* Vol. 709. Lecture Notes in Physics. Springer-Verlag, Berlin, 2006, xii+189.

[Bor+99] A. Borde et al. “Causal continuity in degenerate spacetimes”. *Classical Quantum Gravity* 16 (11) (1999), 3457–3481.

[Bra22] M. Braun. “Rényi’s entropy on Lorentzian spaces. Timelike curvature-dimension conditions” (2022). arXiv: `2206.13005 [math-ph]`.

[Bra23] M. Braun. “Good geodesics satisfying the timelike curvature-dimension condition”. *Nonlinear Analysis* 229 (2023), 113205.

[BC22] M. Braun and M. Calisti. “Timelike Ricci bounds for low regularity spacetimes by optimal transport” (2022). arXiv: `2209.03802 [gr-qc]`.

[Buc+17] M. Buck et al. “The Sorkin–Johnston state in a patch of the trousers spacetime”. *Classical Quantum Gravity* 34 (5) (2017), 055002.

[BBI01] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry.* Vol. 33. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2001, xiv+415.

[BGH21] A. Burtscher and L. García-Heveling. “Time functions on Lorentzian length spaces” (2021). arXiv: `2108.02693 [gr-qc]`.

[BGH22] A. Burtscher and L. García-Heveling. “Global hyperbolicity through the eyes of the null distance” (2022). arXiv: `2209.15610 [math.DG]`.

[Bur15] A. Y. Burtscher. “Length structures on manifolds with continuous Riemannian metrics”. *New York J. Math.* 21 (2015), 273–296.

[BL14] A. Y. Burtscher and P. G. LeFloch. “The formation of trapped surfaces in spherically-symmetric Einstein–Euler spacetimes with bounded variation”. *Journal de Mathématiques Pures et Appliquées* 102 (6) (2014), 1164–1217.

[Bus67] H. Busemann. “Timelike spaces”. *Dissertationes Math. (Rozprawy Mat.)* 53 (1967), 52.

[Car71] B. Carter. "Causal structure in space-time". *Gen. Rel. Grav.* 1 (1971), 349–391.

[CM20] F. Cavalletti and A. Mondino. "Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications" (2020). arXiv: `2004.08934 [math.MG]`.

[CM22] F. Cavalletti and A. Mondino. "A review of Lorentzian synthetic theory of timelike Ricci curvature bounds". *Gen. Relativity Gravitation* 54 (11) (2022), Paper No. 137, 39.

[CY76] S. Y. Cheng and S. T. Yau. "Maximal space-like hypersurfaces in the Lorentz-Minkowski spaces". *Ann. of Math. (2)* 104 (3) (1976), 407–419.

[CBG69] Y. Choquet-Bruhat and R. Geroch. "Global aspects of the Cauchy problem in general relativity". *Comm. Math. Phys.* 14 (1969), 329–335.

[Chr95] D. Christodoulou. "Self-gravitating relativistic fluids: a two-phase model". *Arch. Rational Mech. Anal.* 130 (4) (1995), 343–400.

[Chr99] D. Christodoulou. "On the global initial value problem and the issue of singularities". *Classical Quantum Gravity* 16 (12A) (1999), A23–A35.

[Chr20] P. T. Chruściel. *Geometry of black holes.* English. Vol. 169. Int. Ser. Monogr. Phys. Oxford University Press, 2020.

[CG12] P. T. Chruściel and J. D. E. Grant. "On Lorentzian causality with continuous metrics". *Classical Quantum Gravity* 29 (14) (2012), 145001, 32.

[CGM16] P. T. Chruściel, J. D. E. Grant, and E. Minguzzi. "On differentiability of volume time functions". *Ann. Henri Poincaré* 17 (10) (2016), 2801–2824.

[CEMS01] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger. "A Riemannian interpolation inequality à la Borell, Brascamp and Lieb". *Invent. Math.* 146 (2) (2001), 219–257.

[DL17] M. Dafermos and J. Luk. "The interior of dynamical vacuum black holes I: The C^0-stability of the Kerr Cauchy horizon" (2017). arXiv: `1710.01722 [gr-qc]`.

[Den21] Q. Deng. *Hölder Continuity of Tangent Cones and Non-Branching in RCD(K,N) Spaces.* Thesis (Ph.D.), University of Toronto (Canada). 2021.

[Die87] J. Dieckmann. "Volumenfunktionen in der allgemeinen Relativitätstheorie". (German). PhD thesis. Berlin, 1987.

[Die88a] J. Dieckmann. "Cauchy surfaces in a globally hyperbolic space-time". *J. Math. Phys.* 29 (3) (1988), 578–579.

[Die88b] J. Dieckmann. "Volume functions in general relativity". *Gen. Relativity Gravitation* 20 (9) (1988), 859–867.

[Dow02] F. Dowker. "Topology change in quantum gravity". *Workshop on Conference on the Future of Theoretical Physics and Cosmology in Honor of Steven Hawking's 60th Birthday.* 2002, 436–452.

[DS98] F. Dowker and S. Surya. "Topology change and causal continuity". *Phys. Rev. D* 58 (1998), 124019.

[DG98] H. F. Dowker and R. S. Garcia. "A Handlebody calculus for topology change". *Classical Quantum Gravity* 15 (1998), 1859–1879.

[DGS00a] H. F. Dowker, R. S. Garcia, and S. Surya. "K-causality and degenerate spacetimes". *Classical Quantum Gravity* 17 (21) (2000), 4377–4396.

[DGS00b] H. F. Dowker, R. S. Garcia, and S. Surya. "Morse index and causal continuity. A criterion for topology change in quantum gravity". *Classical Quantum Gravity* 17 (3) (2000), 697–712.

[EM17] M. Eckstein and T. Miller. "Causality for nonlocal phenomena". *Ann. Henri Poincaré* 18 (9) (2017), 3049–3096.

[EPS72] J. Ehlers, F. A. E. Pirani, and A. Schild. "The geometry of free fall and light propagation". *General relativity (papers in honour of J. L. Synge).* Clarendon Press, Oxford, 1972, 63–84.

[Eg16] F. Erkekoğlu. "A survey on sufficient conditions for geodesic completeness of nondegenerate hypersurfaces in Lorentzian geometry". *Int. J. Geom. Methods Mod. Phys.* 13 (3) (2016), 1630003, 9.

[Esc88] J.-H. Eschenburg. "The splitting theorem for space-times with strong energy condition". *J. Differential Geom.* 27 (3) (1988), 477–491.

[Fat15] A. Fathi. "Time functions revisited". *Int. J. Geom. Methods Mod. Phys.* 12 (8) (2015), 1560027, 12.

[FS12] A. Fathi and A. Siconolfi. "On smooth time functions". *Math. Proc. Cambridge Philos. Soc.* 152 (2) (2012), 303–339.

[FF60] H. Federer and W. H. Fleming. "Normal and integral currents". *Ann. of Math. (2)* 72 (1960), 458–520.

[Fey48] R. P. Feynman. "Space-time approach to nonrelativistic quantum mechanics". *Rev. Mod. Phys.* 20 (1948), 367–387.

[Fin18] F. Finster. "Causal Fermion Systems: A Primer for Lorentzian Geometers". *Journal of Physics Conference Series.* Vol. 968. Journal of Physics Conference Series. 2018, 012004.

[FB52] Y. Fourès-Bruhat. "Théorème d'existence pour certains systèmes d'èquations aux dérivées partielles non linéaires". *Acta Math.* 88 (1952), 141–225.

[Gal89] G. J. Galloway. "The Lorentzian splitting theorem without the completeness assumption". *J. Differential Geom.* 29 (2) (1989), 373–387.

[GLS18] G. J. Galloway, E. Ling, and J. Sbierski. "Timelike completeness as an obstruction to C^0-extensions". *Comm. Math. Phys.* 359 (3) (2018), 937–949.

[GH17] S. R. Garcia and R. A. Horn. *A Second Course in Linear Algebra.* 1st ed. Cambridge Mathematical Textbooks. Cambridge University Press, 2017.

[GH21] L. García-Heveling. "Causality theory of spacetimes with continuous Lorentzian metrics revisited". *Classical Quantum Gravity* 38 (14) (2021), 145028, 13.

[GH22] L. García-Heveling. "Topology change with Morse functions: progress on the Borde-Sorkin conjecture" (2022). To appear in *Adv. Theor. Math. Phys.* arXiv: `2202.09833 [gr-qc]`.

[GHS22] L. García-Heveling and E. Soultanis. "Causal bubbles in globally hyperbolic spacetimes". *Gen. Relativity Gravitation* 54 (12) (2022), Paper No. 155, 7.

[Ger70] R. Geroch. "Domain of dependence". *J. Math. Phys.* 11 (1970), 437–449.

[Ger67] R. P. Geroch. "Topology in general relativity". *J. Math. Phys.* 8 (1967), 782–786.

[Ger68] R. P. Geroch. "What is a singularity in general relativity?" *Annals Phys.* 48 (1968), 526–540.

[GT87] R. P. Geroch and J. H. Traschen. "Strings and other distributional sources in general relativity". *Phys. Rev. D* 36 (1987), 1017–1031.

[GKS19] J. D. E. Grant, M. Kunzinger, and C. Sämann. "Inextendibility of spacetimes and Lorentzian length spaces". *Ann. Global Anal. Geom.* 55 (1) (2019), 133–147.

[Gra+20] J. D. E. Grant et al. "The future is not always open". *Lett. Math. Phys.* 110 (1) (2020), 83–103.

[Har88] S. G. Harris. "Closed and complete spacelike hypersurfaces in Minkowski space". *Classical Quantum Gravity* 5 (1) (1988), 111–119.

[HE73] S. W. Hawking and G. F. R. Ellis. *The large scale structure of space-time.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1973.

[HKM76] S. W. Hawking, A. R. King, and P. J. McCarthy. "A New Topology for Curved Space-Time Which Incorporates the Causal, Differential, and Conformal Structures". *J. Math. Phys.* 17 (1976), 174–181.

[HS74] S. W. Hawking and R. K. Sachs. "Causally continuous spacetimes". *Comm. Math. Phys.* 35 (1974), 287–296.

[Haw68] S. Hawking. "The existence of cosmic time functions". *Proc. Roy. Soc. Lond. A* A308 (1968), 433–435.

[Haw14] S. Hawking. "Singularities and the geometry of space-time". *EPJ H* 39 (2014). Reprint of the 1966 Adams prize essay, 413–503.

[Hor91] G. T. Horowitz. "Topology change in classical and quantum gravity". *Classical Quantum Gravity* 8 (1991), 587–602.

[HM19] R. A. Hounnonkpe and E. Minguzzi. “Globally hyperbolic spacetimes can be defined without the ‘causal’ condition”. *Classical Quantum Gravity* 36 (19) (2019), 197001, 9.

[JM19] J. Jang and T. Makino. “On rotating axisymmetric solutions of the Euler–Poisson equations”. *Journal of Differential Equations* 266 (7) (2019), 3942–3972.

[Jan22] D. W. Janssen. “Quantum Fields on Semi-globally Hyperbolic Space-Times”. *Comm. Math. Phys.* 391 (2022), 669—705.

[JL21] J. L. Jauregui and D. A. Lee. “Lower semicontinuity of ADM mass under intrinsic flat convergence”. *Calc. Var. Partial Differential Equations* 60 (5) (2021), Paper No. 193, 42.

[Joh09] S. Johnston. “Feynman propagator for a free scalar field on a causal set”. *Phys. Rev. Lett.* 103 (18) (2009), 180401, 4.

[Jos05] J. Jost. *Postmodern analysis.* Third. Universitext. Springer-Verlag, Berlin, 2005, xvi+371.

[Ket23] C. Ketterer. “Characterization of the null energy condition via displacement convexity of entropy” (2023). arXiv: `2304.01853 [math.DG]`.

[KS20] E.-A. Kontou and K. Sanders. “Energy conditions in general relativity and quantum field theory”. *Classical Quantum Gravity* 37 (19) (2020), 193001, 63.

[KS21] M. Kontsevich and G. Segal. “Wick rotation and the positivity of energy in quantum field theory”. *Q. J. Math.* 72 (1-2) (2021), 673–699.

[KP08] S. G. Krantz and H. R. Parks. *Geometric integration theory.* Cornerstones. Birkhäuser Boston, 2008.

[KP67] E. H. Kronheimer and R. Penrose. “On the structure of causal spaces”. *Proc. Cambridge Philos. Soc.* 63 (1967), 481–501.

[Kun67] W. Kundt. “Non-existence of trouser-worlds”. *Comm. Math. Phys.* 4 (2) (1967), 143–144.

[KS18] M. Kunzinger and C. Sämann. “Lorentzian length spaces”. *Ann. Global Anal. Geom.* 54 (3) (2018), 399–447.

[KS22] M. Kunzinger and R. Steinbauer. “Null distance and convergence of Lorentzian length spaces”. *Ann. Henri Poincaré* (2022).

[Lan21] K. Landsman. *Foundations of General Relativity. From Einstein to Black Holes.* Nijmegen: Radboud University Press, 2021, 394.

[Lan11] U. Lang. “Local currents in metric spaces”. *J. Geom. Anal.* 21 (3) (2011), 683–742.

[LW11] U. Lang and S. Wenger. “The pointed flat compactness theorem for locally integral currents”. *Comm. Anal. Geom.* 19 (1) (2011), 159–189.

[Lee09] J. M. Lee. *Manifolds and differential geometry.* Vol. 107. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2009, xiv+671.

[Ler53] J. Leray. *Hyperbolic differential equations.* Institute of Advanced Study, 1953, reprinted 1955.

[Lev83] V. L. Levin. “Continuous utility theorem for closed preorders on a metrizable σ-compact space”. *Soviet Math. Dokl.* 28 (3) (1983), 715–718.

[Lin20] E. Ling. “Aspects of C^0 causal theory”. *Gen. Relativity Gravitation* 52 (6) (2020), Paper No. 57, 40.

[Lol20] R. Loll. “Quantum Gravity from Causal Dynamical Triangulations: A Review”. *Classical Quantum Gravity* 37 (1) (2020), 013002.

[LV09] J. Lott and C. Villani. “Ricci curvature for metric-measure spaces via optimal transport”. *Ann. of Math. (2)* 169 (3) (2009), 903–991.

[LS97] J. Louko and R. D. Sorkin. “Complex actions in two-dimensional topology change”. *Classical Quantum Gravity* 14 (1997), 179–204.

[Mag22] M. Magnabosco. “Example of a highly branching CD space”. *J. Geom. Anal.* 32 (6) (2022), Paper No. 173, 32.

[MCD88] C. Manogue, E. Copeland, and T. Dray. “The trousers problem revisited”. *Pramana* 30 (1988), 279—292.

[MS48] E. Marczewski and R. Sikorski. “Measures in non-separable metric spaces”. *Colloq. Math.* 1 (1948), 133–139.

[McC20] R. J. McCann. “Displacement convexity of Boltzmann’s entropy characterizes the strong energy condition from general relativity”. *Camb. J. Math.* 8 (3) (2020), 609–681.

[McC23] R. J. McCann. “A synthetic null energy condition” (2023). arXiv: `2304.14341 [gr-qc]`.

[MS22a] R. J. McCann and C. Sämann. “A Lorentzian analog for Hausdorff dimension and measure”. *Pure and Applied Analysis* 4 (2) (2022), 367–400.

[Mil63] J. W. Milnor. *Morse theory. Based on lecture notes by M. Spivak and R. Wells.* Vol. 51. Annals of Mathematics Studies. Princeton University Press, 1963.

[Min08a] E. Minguzzi. “Limit curve theorems in Lorentzian geometry”. *J. Math. Phys.* 49 (2008), 092501–092518.

[Min08b] E. Minguzzi. “Weak distinction and the optimal definition of causal continuity”. *Classical Quantum Gravity* 25 (2008), 075015–075021.

[Min09] E. Minguzzi. “K-causality coincides with stable causality”. *Comm. Math. Phys.* 290 (1) (2009), 239–248.

[Min10] E. Minguzzi. “Time functions as utilities”. *Comm. Math. Phys.* 298 (3) (2010), 855–868.

[Min13] E. Minguzzi. “Topological ordered spaces as a foundation for a quantum spacetime theory”. *J. Phys. Conf. Ser.* 442 (2013). Ed. by L. Diosi et al., 012034.

[Min18] E. Minguzzi. “The representation of spacetime through steep time functions”. *J. Phys. Conf. Ser.* 968 (2018), 012009.

[Min19a] E. Minguzzi. “Lorentzian causality theory”. *Living Rev. Rel.* 22 (1) (2019), 3–204.

[Min23] E. Minguzzi. “Further observations on the definition of global hyperbolicity under low regularity” (2023). arXiv: `2302.09284 [gr-qc]`.

[MS22b] E. Minguzzi and S. Suhr. “Lorentzian metric spaces and their Gromov-Hausdorff convergence” (2022). arXiv: `2209.14384 [math.DG]`.

[Min19b] E. Minguzzi. “Causality theory for closed cone structures with applications”. *Rev. Math. Phys.* 31 (5) (2019), 1930001, 139.

[MTW73] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman and Co., San Francisco, Calif., 1973, ii+xxvi+1279+iipp.

[MS22c] A. Mondino and S. Suhr. “An optimal transport formulation of the Einstein equations of general relativity”. *J. Eur. Math. Soc.* (2022). DOI 10.4171/JEMS/1188.

[MS11] O. Müller and M. Sánchez. “Lorentzian manifolds isometrically embeddable in $\mathbb{L}^N$”. *Trans. Amer. Math. Soc.* 363 (10) (2011), 5367–5379.

[Mül22a] O. Müller. “Gromov-Hausdorff distances for Lorentzian length spaces” (2022). arXiv: `2209.12736 [math.DG]`.

[Mül22b] O. Müller. “Lorentzian Gromov-Hausdorff theory and finiteness results”. *Gen. Relativity Gravitation* 54 (10) (2022), Paper No. 117, 17.

[Mun00] J. R. Munkres. *Topology*. Prentice Hall, 2000, xvi+537.

[Nac65] L. Nachbin. *Topology and order*. Van Nostrand Mathematical Studies, No. 4. Translated from the Portuguese by Lulu Bechtolsheim. D. Van Nostrand Co., 1965, vi+122.

[NP07] H. Nicolai and K. Peeters. “Loop and Spin Foam Quantum Gravity: A Brief Guide for Beginners”. *Approaches to Fundamental Physics: An Assessment of Current Theoretical Ideas*. Ed. by I.-O. Stamatescu and E. Seiler. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, 151–184.

[Nol04a] J. Noldus. “A Lorentzian Lipschitz, Gromov-Hausdoff notion of distance”. *Classical Quantum Gravity* 21 (2004), 839–850.

[Nol04b] J. Noldus. “The Limit space of a Cauchy sequence of globally hyperbolic space-times”. *Classical Quantum Gravity* 21 (2004), 851–874.

[NO61] K. Nomizu and H. Ozeki. “The existence of complete Riemannian metrics”. *Proc. Amer. Math. Soc.* 12 (1961), 889–891.

[Pen79] R. Penrose. “Singularities and time asymmetry”. *General Relativity. An Einstein Centenary Survey*. Cambridge University Press, 1979, 581–638.

[Pen65] R. Penrose. “Gravitational collapse and space-time singularities”. *Phys. Rev. Lett.* 14 (1965), 57–59.

[RS05] M.-K. von Renesse and K.-T. Sturm. “Transport inequalities, gradient estimates, entropy, and Ricci curvature”. *Comm. Pure Appl. Math.* 58 (7) (2005), 923–940.

[Rin09] H. Ringström. *The Cauchy problem in general relativity.* ESI Lectures in Mathematics and Physics. European Mathematical Society (EMS), Zürich, 2009, xiv+294.

[Rot22] F. Rott. “Gluing of Lorentzian length spaces and the causal ladder” (2022). arXiv: `2209.06894 [math.DG]`.

[SW77] R. K. Sachs and H. H. Wu. *General relativity for mathematicians.* Graduate Texts in Mathematics, Vol. 48. Springer-Verlag, New York-Heidelberg, 1977, xii+291.

[SS23] A. Sakovich and C. Sormani. “The null distance encodes causality”. *J. Math. Phys.* 64 (1) (2023), 012502.

[Säm16] C. Sämann. “Global hyperbolicity for spacetimes with continuous metrics”. *Ann. Henri Poincaré* 17 (6) (2016), 1429–1455.

[Sán22] M. Sánchez. “Globally hyperbolic spacetimes: slicings, boundaries and counterexamples”. *Gen. Relativity Gravitation* 54 (10) (2022), Paper No. 124, 52.

[Sán23] M. Sánchez. “A class of cosmological models with spatially constant sign-changing curvature”. *Port. Math.* (2023). DOI 10.4171/PM/2099.

[Sbi18] J. Sbierski. “The C^0-inextendibility of the Schwarzschild spacetime and the spacelike diameter in Lorentzian geometry”. *J. Diff. Geom.* 108 (2) (2018), 319–378.

[Sbi22] J. Sbierski. “On holonomy singularities in general relativity and the $C^{0,1}_{\mathrm{loc}}$-inextendibility of space-times”. *Duke Math. J.* 171 (14) (2022), 2881–2942.

[Sei77] H.-J. Seifert. “Smoothing and extending cosmic time functions”. *Gen. Relativity Gravitation* 8 (10) (1977), 815–831.

[SW96] R. D. Sorkin and E. Woolgar. “A causal order for spacetimes with C^0 Lorentzian metrics: proof of compactness of the space of causal curves”. *Classical Quantum Gravity* 13 (7) (1996), 1971–1993.

[Sor89] R. D. Sorkin. “Consequences of spacetime topology”. *Proceedings of the Third Canadian Conference on General Relativity and Relativistic Astrophysics.* 1989, 137—163.

[Sor97] R. D. Sorkin. “Forks in the road, on the way to quantum gravity”. *Int. J. Theor. Phys.* 36 (1997), 2759–2781.

[Sor18] C. Sormani. “Spacetime Intrinsic Flat Convergence”. *Oberwolfach Rep. 15.* 2018, 2226–2230.

[SV16] C. Sormani and C. Vega. “Null distance on a spacetime”. *Classical Quantum Gravity* 33 (8) (2016), 085001, 29.

[SW11] C. Sormani and S. Wenger. “The intrinsic flat distance between Riemannian manifolds and other integral current spaces”. *J. Differential Geom.* 87 (1) (2011), 117–199.

[Sou22] E. Soultanis. “Globally hyperbolic Lorentzian length spaces with orthogonally split metric”. In preparation. 2022.

[Stu06] K.-T. Sturm. “On the geometry of metric measure spaces I & II”. *Acta Math.* 196 (1) (2006), 65–177.

[Sur19] S. Surya. “The causal set approach to quantum gravity”. *Living Rev. Rel.* 22 (1) (2019), 5.

[Veg21] C. Vega. “Spacetime distances: an exploration” (2021). arXiv: `2103.01191 [gr-qc]`.

[WY81] R. M. Wald and P. Yip. “On the existence of simultaneous synchronous coordinates in spacetimes with spacelike singularities”. *J. Math. Phys.* 22 (11) (1981), 2659–2665.

[Wei82] D. Weingarten. “Euclidean quantum gravity on a lattice”. *Nucl. Phys. B* 210 (1982), 229–245.

[Wey88] H. Weyl. *Raum. Zeit. Materie.* Seventh. Vol. 251. Springer-Verlag, Berlin, 1988, xvi+349.

[Whe57] J. A. Wheeler. “On the nature of quantum geometrodynamics”. English. *Ann. Phys.* 2 (1957), 604–614.

[Wit22] E. Witten. “A Note On Complex Spacetime Metrics”. *Frank Wilczek. 50 Years of Theoretical Physics.* 2022, 245–280.

[Woo73] N. M. J. Woodhouse. “The differentiable and causal structures of space-time”. *J. Math. Phys.* 14 (1973), 495–501.

[Yod72] P. Yodzis. “Lorentz cobordism”. *Comm. Math. Phys.* 26 (1972), 39–52.

[Yod73] P. Yodzis. “Lorentz cobordism. II”. *Gen. Relativity Gravitation* 4 (4) (1973), 299–307.