Imperial College London

Faculty of Natural Sciences

Department of Theoretical Physics

Shaping a Spacetime from Causal Structure

Author:

Giovanni Guccione

Supervisor:

Helen Fay Dowker

Submitted on September 21th 2011 in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London.

Acknowledgements

This dissertation would not have been possible without the supervision of Helen Fay Dowker and her countless advices. Working with her has been both an honour and a pleasure.
I am also grateful to: my new colleagues and friends, for moments of amusement and inspiration; my former colleagues, for their steady presence in the background of every scribble, scientific in nature or not; my oldest friends, for their complicitous tolerance and prismatic influence.
Last but not least I thank my family, for giving me their support and the opportunity to chase my ambitions.

Contents

Chapter 1

Introduction

1.1 The role of causal structure in physics

Since the introduction of relativity in the first years of the 20th century, physicists became interested in models of the universe that treated space and time on the same level, thus creating the concept of *spacetime*.
Since then, spacetime has been fully established as the framework for any physical event, although we need to remember that it is not uniquely defined and that there are many different models compatible with the theory of General Relativity. For example, an empty universe is very well described by a flat model, while the presence of a massive object induces a perturbation in spacetime that can even cause the emergence of singularities.
There are specific conditions that a spacetime must satisfy in order to be compatible with our universe, usually accepted in an axiomatic form.

Causal structure is determined by the set of events in spacetime and the knowledge of the relations between them, namely whether an event can influence another or not.
The way such a relatively simple structure encodes most of the information of a spacetime by itself has been rediscovered several times in the past, and with a rigorous approach it is possible to specify the minimal set of axioms necessary to its definition while preserving the interactions between different events.

The question may be asked of what is possible to get back about the spacetime from the knowledge of causal structure alone. We will address this problem trying to use the minimal amount of assumptions necessary for each different level of the architecture underlying the spacetime.
There have also been suggestions that the fundamental nature of spacetime might be discrete; in this case the study of causal structure turns out to be particularly functional and convenient through the introduction of *causal sets*.

We will start with the establishment of the most basic definitions used in the mathematical approach to the topic, specifying all the main characteristics wished for a spacetime without the requirement that they must be fundamental and cannot be obtained from more basic elements. Different degrees of restrictions on the causality of spacetime are then introduced, together with properties that are always satisfied and can be deduced from the essential principles.
In Chapter 2 we will try to understand how much structure is deducible from the simple set of points constituting the spacetime and all causal relations between. Starting from the relations other than causal, the main results about topology, metric and differential structure will be presented with an eye on the weakest conditions required in each case.
Chapter 3 will deal with the application of the content from the previous chapter to recreate the main structure of Minkowski spacetime. A constructive method is used to obtain the metric and the tangent space from nothing more than causality.
Finally, in Chapter 4 we will present alternative topologies more closely related to causal structure and with a deeper connection to the physical essence of a spacetime.

In the following dissertation we assume that the reader is familiar with the definitions of *topology*, *manifold* and other basic notions from Differential Geometry.

1.2 Definitions

A *spacetime* is a real, connected, smooth, Hausdorff manifold M.
The geometrical structure of spacetime is determined by the *metric* g, a globally defined non-degenerate $(0, 2)$ tensor field. This is taken to be Lorentzian, i.e. for any point x of the manifold the tangent space $T_x(M)$ admits a basis, called *normal coordinate basis*, such that the metric has signature $(-1, 1, \ldots, 1)$.
The metric acts on elements of $T_x(M)$ as a bilinear function. A vector $X \in T_x(M)$ is either *spacelike*, *timelike*, or *null*, according to whether its square norm $g(X, X)$ is positive, negative, or zero.
The spacetime is *time-orientable* if the timelike tangent vectors at each point can be divided in two distinct classes, allowing the distinction between past- or future-pointing. The choice for past or future is arbitrary, as long as it is consistent within all points in spacetime. The value of the metric applied to two timelike vectors of the same class is strictly negative, while it can assume any value if one vector is past-pointing and the other is future-pointing. We will always assume a time-orientable spacetime.

A *path* is a smooth map $\gamma\colon I \to M$, where I is a connected (non-singular) interval in $\mathbb{R}$. When no confusion is likely, the same symbol will be used for its image $\gamma[I]$, known as the *curve* on M.
A curve is said to be *chronological* when its tangent at any point is timelike, and *causal* when its tangent is anywhere timelike or null. If the tangent vectors have positive time component (respectively, negative) at all points, then the curve is *future-oriented* (*past-oriented*).

Given two points $x, y \in M$, we say that y *follows* x if there is a future-oriented chronological curve connecting x to y, and we write $x \ll y$. Analogously, we say that y *follows* x *causally* if there is a future-oriented causal curve from x to y or if $x = y$, and use the notation $x \prec y$. The *horismotic relation* $x \to y$ indicates that $x \prec y$ but $x \not\ll y$. By $x \natural y$ we denote two points that are not causally related, i.e. spacelike separated.
The set of points following a fixed $x \in M$ constitutes its *chronological future*, $I^+(x) := \{y \in M : \ x \ll y\}$; the *causal future* is the set of points fol-

lowing x causally: $J^+(x) := \{y \in M : x \prec y\}$; the points in the causal future but not in the chronological future of x constitute the *future horismos* $E^+(x) := \{y \in M : x \to y\}$. Substituting "following" with "preceding" we get similar definitions for the *chronological past* $I^-(x)$, the *causal past* $J^-(x)$ and the *past horismos* $E^-(x)$. The union of the past and future horismos is the *null cone*, $C(x) := E^+(x) \cup E^-(x)$.
The intersection of the past of a point y with the future of a point $x \prec y$ is called *chronological interval*, or simply *interval*: $I(x, y) := I^+(x) \cap I^-(y)$. The *causal interval* $J(x, y)$ and *horismos interval* $E(x, y)$ are defined correspondingly.

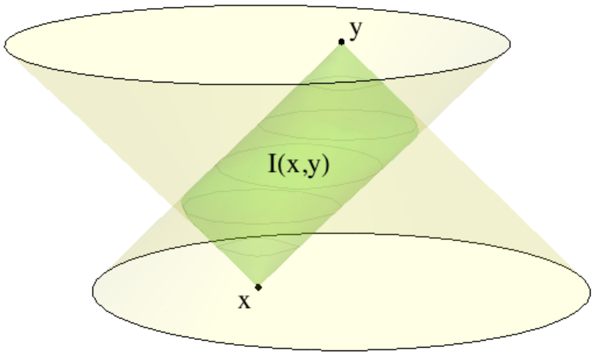

Figure 1.1: Example of interval in flat spacetime, given by the intersection of the interior of the light cones.

A spacetime M is *chronological* if admits no closed chronological curve, and similarly it is *causal* if there is no closed causal curve. It is reasonable to require the spacetime M to be at least causal.
Given a point $x \in M$, if $I^+(x) \neq I^+(y)$ for all $y \in M$ different from x we say that the spacetime is *future distinguishing* at x, and similarly it is *past distinguishing* at x if $x \neq y$ implies $I^-(x) \neq I^-(y)$. A spacetime is *future/past distinguishing* (F/P) if it is future or past distinguishing at any point, and a spacetime that is both future and past distinguishing (FPD) may be simply referred to as *distinguishing*.
A neighbourhood U of x is *causally convex* if it intersects any causal curve γ at most once, i.e. $\gamma \cap U$ is connected. A spacetime is *strongly causal* at

x if such a neighbourhood exists, and it *strongly causal* if the condition is satisfied at any point.
Every global causality condition stated implies all the previous ones, but the converse is not true:
Strongly causal $\Rightarrow$ FPD $\Rightarrow$ F/P distinguishing $\Rightarrow$ Causal $\Rightarrow$ Chronological .

For a causal spacetime the causal relation $\prec$ is a *partial order*, satisfying:

- Reflexivity: $x \prec x$
- Antisymmetry: $x \prec y,\ y \prec x \implies x = y$
- Transitivity: $x \prec y,\ y \prec z \implies x \prec z$

If causality was not required, we might find some points $x, y \in M$ such that $x \prec y,\ y \prec x$ but $x \neq y$, breaking the antisymmetry property of the partial order.
Any subset of a chronological curve is called a *chronological chain*, whereas the same for a causal curve is a *causal chain* or simply *chain*. Points in a chain are either timelike or null separated, but never spacelike separated; they satisfy the additional condition required for a *total order*:

- Totality: either $x \prec y$ or $y \prec x$

1.3 Basic properties

We will now introduce the main features of the causal structure of spacetime. For a full approach to the topic we suggest Penrose's "*Techniques of Differential Topology in Relativity*" [4].

The following propositions are true for any spacetime M, and most of them have a dual counterpart obtained by the substitution of every future with a past set and the reflection of any relation. Here x, y and z are distinct events of M.

Proposition 1.1 $I^+(x)$ *is open.*

Proposition 1.2 $y \in I^+(x) \implies x \in I^-(y)$.

Proposition 1.3 $J^+(x) \subseteq \overline{I^+(x)}$ *(closure of* $I^+(x)$*).*

Proposition 1.4 $x \ll y \implies x \prec y$.

Proposition 1.5 *Either* $\begin{cases} x \ll y,\ y \prec z \\ x \prec y,\ y \ll z \end{cases} \implies x \ll z.$

Proposition 1.6 $x \prec y \prec z,\ x \to z \implies x \to y \to z.$

Global causality conditions on M are also very important in the study of causal structure, and it is worth mentioning a few properties for them as well. As a reference on causality conditions we reccomend the consultation of "*The causal hierarchy of spacetimes*" [13] by Minguzzi and Sánchez.
We are already familiar with the following proposition:

Proposition 1.7 *M is strongly causal* $\implies$ *M is distinguishing.*

This may not seem immediate from the definitions, apparently different from each other. There is however an equivalent definition for distinguishing spacetimes based on causal curves and neighbourhoods: M is distinguishing if and only if for all $x \in M$ there exists a neighbourhood U of x such that the intersection with any causal curve through x is connected. Compared with the strong causality condition, where the connectedness is required for any causal curve intersecting U and not just for those passing through x, it is clear that the requirement of future and past distinction is weaker.

There is also an alternative definition for strong causality. Let U be an open subset of the manifold, and consider the problem of finding the causal relations between its points. We can either consider the relations between the events of U under the causal structure of the entire manifold, or we could treat U as a spacetime by itself and have all causal curves lying in the subset define a new causal structure.
These two methods usually produce different results, as the second case is more restrictive; this however does not apply when the spacetime is strongly causal, thanks to the property of causally convex subsets that allow no causal curve between two points of U to go "outside" of the open set. Under strong

causality the local and the global behaviour of causal structure is the same, and the two methods above coincide.

Finally, the following proposition characterizes the behaviour of events in proximity of a point where strong causality fails.

Proposition 1.8 *M is not strongly causal at x if and only if there exists $y \prec x$ in M, $y \neq x$, such that $a \ll b$ for any a, b satisfying $a \ll x$ and $y \ll b$.*

If the spacetime is future distinguishing, this attribute of achronality extends to an entire neighbourhood in the future of the point where strong causality is violated. Analogous conditions hold, with the opportune adaptations, for the past of the point.

Chapter 2

Structure encoded by causality

We will now take in examination a set of points together with all causal relations, $(M, \prec)$, and try to determine all the information that can be obtained on the structure of the manifold starting from these two basic objects.

2.1 Chronological and horismotic relations

The first thing to observe is that, of all the definitions introduced in Chapter 1, we are assuming the causal relation but not the chronological and horismotic relations.
It is actually possible to extract the missing two starting from the one we already have, and this problem has been analysed in details by Kronheimer and Penrose in their article "*On the structure of causal spaces*" [3] under the most general conditions.
We will however restrict ourselves to strongly causal spacetimes, an assumption that we will carry to the following stages. Note that weakest choices on the causality condition of spacetime can yield the same results, however the proofs might get unnecessarily complicated or uninteresting.

Instead of chronological and causal relations, we will deal with future sets, and of course dual arguments will hold for the pasts. This approach is completely equivalent to the consideration of relations in the first place, since the latter can be easily established starting from the sets.

Definition A *least upper bound* of a set U is a point y such that:

- $\forall\, x \in U,\; x \prec y$
- $\exists\, z :\; \forall\, x \in U,\; x \prec z \implies y \prec z$

Define now the set of points $dJ^+(x)$ as all points $y \in J^+(x)$ such that there exists a finite causally increasing succession $\{z_n\}$ of N points in M i.e. $z_{n-1} \prec z_n$ holds for all $n \le N$, and for which y acts as a least upper bound. Let $\partial J^+(x)$ be the topological boundary of the causal future at point x.

Proposition 2.1 $dJ^+(x) \equiv \partial J^+(x)$

Proof The fact that every point $y \in \partial J^+(x)$ is also in $dJ^+(x)$ is straightforward: by definition as an element of the boundary y belongs to $J^+(x)$, and it is clearly an upper bound for any succession $\{z_n\} \in J^-(y)$ such that any $z_n \notin J^+(x)$.
Now suppose that $y \in dJ^+(x)$ but $\notin \partial J^+(x)$. Since by assumption it is a point of $J^+(x)$ but not of its boundary, it must be in the interior, open by definition. Let $\{z_n\}$ be a finite increasing succession having y as least upper bound, and consider an open neighbourhood U of y, small enough to leave the last element of the succession z_N outside. Because U is open, we can find a point p in the intersection of U with the causal curve connecting z_N to y such that $p \prec y$. Strong causality makes sure that this condition holds globally and not only inside the neighbourhood. However, this means that we can find this point p satisfying both $p \prec y$ and $z_n \prec p$ for all elements of the succession, contradicting the fact that y is the *least* upper bound. Hence, y must belong to $\partial J^+(x)$. □

Having determined the boundary of $J^+(x)$ without the necessity of a topology or any structure other than causality, we can define other sets of interests.
It is quite intuitive to show that the future horismos of x is given by all points of $J^+(x)$ that also belong to the boundary $\partial J^+(x)$.

Proposition 2.2 $E^+(x) \equiv J^+(x) \cap \partial J^+(x)$

From this we can define the chronological future as all points of $J^+(x)$ which are not horismotically related.

Proposition 2.3 $I^+(x) \equiv J^+(x) \setminus \partial J^+(x)$

Note that, while Proposition 1.3 states that the causal future is *included* in the closure of $I^+(x)$, it is always true that the chronological future is *equal* to the interior of $J^+(x)$.

An alternative procedure more commonly found in literature uses the causal relation to define first the horismotic relation and then the chronological relation, although one again needs to be careful that the spacetime is sufficiently well-behaved and is at least future and past distinguishing.
Two points x and y are defined to be horismotically related if the causal interval $J(x, y)$ is a *null chain*, i.e. it is a chain such that any subset is also a chain. Then the chronological relation is simply obtained by requiring that two points are causally but not horismotically related.
The notion of null geodesics as null chains is investigated again and in more detail in Section 3.1.

2.2 Topology

Next, we want to specify the topology of the manifold, again only in terms of causal structure. We can also use our previous results and assume the knowledge of chronological relations, and we will indeed use this knowledge to define a new topology for the spacetime that coincides with the original topology of the manifold for strongly causal spacetimes.
Again, the condition of strong causality is not the weakest choice: in his paper "*The class of continuous timelike curves determines the topology of spacetime*" [8], Malament shows that a causal bijection between distinguishing spacetimes is also a causal homeomorphism, proving in fact that the topology is inferred from causality even if the spacetime is not strongly causal. He makes, however, the silent assumption that the spacetimes involved have the same dimension, although this is not trivially guaranteed by

the simple preservation of the causal structure; Parrikar and Surya analyze this matter in "*Causal Topology in Future and Past Distinguishing Spacetimes*" [14], getting to the conclusion that for distinguishing spacetimes the dimension of the manifold is indeed attainable from causality, confirming that a causal bijection is all that is necessary to preserve the dimension of the manifold and assure that the topology is completely determined from causal structure. Malament also shows that the same result is not true if the spacetime is just future *or* past distinguishing.

It is better to specify exactly what is meant by manifold topology, as this notion will be recurrent here and in Chapter 4.

Definition The *manifold topology* $\mathcal{M}$ is the topology generated by the set of n-dimensional open balls $B_\varepsilon(x) := \{y \in M : \ \|x - y\| < \varepsilon\}$, where the norm is intended to be of Euclidian type.

The definition above has no apparent connection with causality. A new topology defined in terms of causal relations is necessary, and it must be showed that it can be associated with $\mathcal{M}$ under the proper conditions.

Definition The *Alexandrov topology* $\mathcal{A}$ is the coarsest topology on the manifold such that the chronological intervals are open sets.

Open sets for $\mathcal{A}$ are then given by the union of chronological intervals $I(x, y)$, that form a basis for this topology.
It is not easy to determine which topology is more suitable for a physical theory. It would seem natural to adopt the Alexandrov topology as the one that we can observe and measure, as the events of spacetime are constricted by the relativistic limits and therefore evolve along timelike curves. In any case, the answer seems to depend on the behaviour of spacetime itself.

According to the definition, the Alexandrov topology is in general coarser than the original manifold topology:

Proposition 2.4 *$\mathcal{A} \subseteq \mathcal{M}$.*

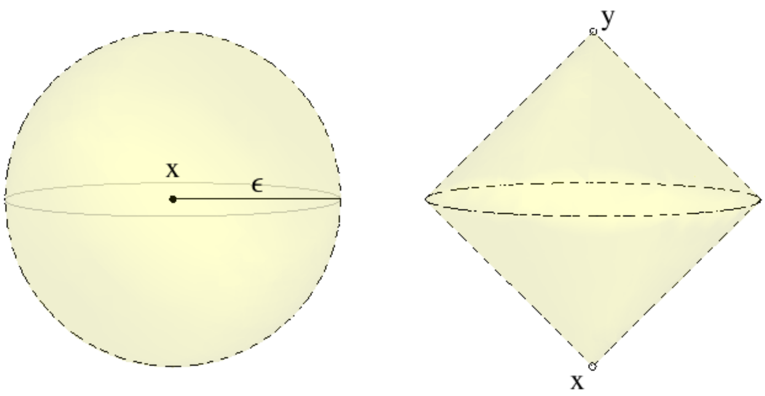

Figure 2.1: To the left, the $\mathcal{M}$-open ball $B_\varepsilon(x)$; to the right, the $\mathcal{A}$-open interval $I(x,\, y)$. The sets are only examples on a plain 3-dimensional spacetime; dashed lines indicate removed points.

If we had defined $\mathcal{A}$ directly as the topology generated by the chronological intervals $I(x,\, y)$, the coarseness with respect to $\mathcal{M}$ would have followed automatically, as from Proposition 1.1 we know that the intervals are always open in the original topology.

The question may be asked of whether the Alexandrov topology ever coincides with the manifold topology instead of being coarser. We will now show that strong causality is the necessary and sufficient condition for this equivalence to hold.

Theorem 2.5 *M is strongly causal $\iff$ $\mathcal{A}$ coincides with $\mathcal{M}$.*

Proof Let's first prove that strong causality implies agreement between the two topologies. We already know that any open set in $\mathcal{A}$ is also open in $\mathcal{M}$, and need to prove the converse. Let O be an $\mathcal{M}$-open neighbourhood of a point $x \in M$. Strong causality tells us that there is a causally convex neighbourhood $V \in O$ of x. Let $a, b \in V$ be the past and future endpoints of the intersection of V with a causal curve passing through x, so that $x \in I(a,\, b)$. We know that such subset and points exist for any $x \in O$, meaning that O is $\mathcal{A}$-open.
To prove the inverse implication, assume that $\mathcal{A}$ and $\mathcal{M}$ agree but strong

causality for M is violated at a point x. It follows that, thanks to Proposition 1.8 we can find points a, b of an open neighbourhood O of x such that $a \prec b$, $a \neq b$, and any point in the future of a follows chronologically any point in the past of b; because O is also $\mathcal{A}$-open we can find intervals $I(p_a,\, q_a) \ni a$ and $I(p_b,\, q_b) \ni b$. Because $a \prec b$ and $b \ll q_b$, it follows from Proposition 1.5 that $a \ll q_b$. As the past of q_b is open, choose $c \in I(p_a,\, q_a)$ just to the future of a so that $c \ll q_b$: by construction, it will also automatically be in the future of p_b and belong to $I(p_b,\, q_b)$. We have just proved that $I(p_a,\, q_a) \cup I(p_b,\, q_b) \neq \emptyset$, and this must hold for any intervals containing a and b, meaning that M is not Hausdorff. This however is in contradiction with the assumption that $\mathcal{A}$ coincides with the standard topology on M. □

2.3 Metric

For a strongly causal spacetime we already know how to derive the chronological relation from causal structure, and use it for the definition of a topology equivalent to the original.
Thanks to an important result by Hawking and Ellis in "*The large scale structure of space-time*" [5], the process can be pushed further with the recovery of the metric, or at least most of it, from the knowledge of null geodesics naturally included in that of causal structure.

Theorem 2.6 *The metric at a point x is determined up to a conformal factor by the knowledge of all null vectors at x.*

In Chapter 3 we present the proof by construction to this theorem on the example of Minkowski spacetime, although the procedure itself is not restricted to this case and can be used on a general curved spacetime.

At this point it is convenient to introduce a class of functions useful for a theory on causal structure: a *causal map* is an invertible map $f\colon M \to \tilde{M}$ preserving the causal relation between two events, i.e. for any $x, y \in M$, if $x \prec y$ then $f(x) \prec f(y)$, where of course the causal relation is to be intended on the appropriate spacetime.
The concept is easily extended to *causal bijections* by requiring that f is

invertible and f^{-1} also preserves causal relations; we can also generalize to *causal homeomorphisms* and *causal diffeomorphisms* by adding the condition that f is respectively continuous and smooth, and so is f^{-1}.
Note that, under the strong causality condition (or weaker, as already discussed before), a map preserving causal relations automatically preserves the topology, and consequently is a causal homeomorphism.

Because causal bijections naturally preserve null geodesics, Theorem 2.6 is conveniently rephrased as following:

Theorem 2.6 (2) *Let f be a causal bijection $f\colon M \to \tilde{M}$ and g, $\tilde{g}$ be the metrics for M and $\tilde{M}$. Then, for some conformal factor Ω, we have that $f_* g|_x = \Omega^2\, \tilde{g}|_x$.*

2.4 Differential structure

The role of causality is so fundamental that it can determine almost everything about the spacetime. Even the differential structure can be recovered, and together with metric and topology this means that the complete conformal geometry of the spacetime is determined.
The theorem we are about to present was firstly suggested by Hawking in his unpublished essay "*Singularities and the Geometry of Space-Time*" [1]. and later reintroduced in the paper "*A new topology for curved space-time which incorporates the causal, differential, and conformal structures*" [7] with the collaboration of King and McCarthy. The assumptions of the original theorem required a chronological homeomorphism, or equivalently a chronological bijection on a strongly causal spacetime; Malament's result [8] can be applied also in this case to offer a generalization to future and past distinguishing spacetimes. Finally, Levichev showed in "*Prescribing the conformal geometry of a Lorentz manifold by means of its causal structure*" [9] that a chronological bijection between distinguishing spacetimes is also a causal bijection, and vice versa; this permits to use a class of functions more closely related to causality in the assumptions.
Thanks to the contribution from all these authors, this is known as the "Malament-Hawking-King-McCarthy-Levichev" (MHKML) theorem.

For its proof we will continue with the assumption of strong causality.
Before going further we will also have to appeal to the fact that causal bijections are order-preserving, not only locally but globally.

Proposition 2.7 *Let f be a causal bijection $f\colon M \to \tilde{M}$. Then, both f and f^{-1} either preserve or reverse the future/past orientation of timelike curves.*

Proof Pick two timelike curves ϕ, π through a point $x \in M$ with the same orientation, e.g. future-pointing, and suppose that $f[\phi]$ is future-pointing while $f[\pi]$ is past-pointing, so that the orientation of one curve is preserved while the other is reversed. Consider the segments $f[\phi \cap J^{-}(x)]$ and $f[\pi \cap J^{+}(x)]$; their union is not a timelike curve, because the tangent at $f(x)$ is not defined is not defined. But by assumption f is a causal bijection and should map timelike curves to timelike curves, so this is impossible.
To verify that the behaviour of f is systematic and not restricted to the point x, we need to verify that the set of point where the orientation is preserved/reversed by f is open. Let Φ be the set of all point where orientation is preserved, and Π the set where it is reversed, and consider a point $x \in \Phi$ and a point y in the future of x. Since $I^{-}(y)$ is open, we can find an open neighbourhood O of x. Suppose there exists a point $z \in O \cap \Pi$, then the union of a timelike curve ϕ from x to y and another timelike curve π from z to y is not timelike, but its image $f[\phi \cup \pi]$ could be, in view of the fact that the segment lying in Π is reversed. This however means that we reach again a contradiction, since f^{-1} should also preserve timelike curves by assumption. Then the open set O does not intersect Π, and is included in Φ, meaning that Π is open. The same argument can be repeated for a point $x \in \Phi$, showing that Φ is also open. □

Theorem 2.8 *Let f be a causal bijection $f\colon M \to \tilde{M}$ between two future and past distinguishing spacetimes of dimension greater than 2. Then, f is a C^{∞} diffeomorphism.*

Proof The proof is based on a construction using null geodesics. Because f is a causal bijection, it naturally maps null vectors into null vectors and

preserves the nature of the geodesics; we only need to prove that the parameter is mapped smoothly and that it preserves a C^∞ atlas.
Consider a convex subset U of M and four C^∞ null geodesic paths $\gamma_i\colon I_i \to U$ satisfying:
(1) For all $t_1 \in I_1$ there exists a unique null geodesic curve λ in U joining $\gamma_1(t_1)$ to the null geodesic curve γ_2;
(2) For all $t_3 \in I_3$ there exists a unique point $q \in \lambda$ joined to $\gamma_3(t_3)$ by some null geodesic curve in U;
(3) For all $q \in \lambda$ there exists a unique $t_4 \in I_4$ such that there is a null geodesic curve joining q to $\gamma_4(t_4)$;
(4) The map $\psi\colon I_1 \times I_3 \to I_4$ defined by $\psi(t_1, t_3) = t_4$ is C^∞ and has non-vanishing derivatives $\dfrac{\partial\psi}{\partial t_1}, \dfrac{\partial\psi}{\partial t_3} \neq 0$.
Note that, if all conditions including (4) must be met, the dimension of the manifold n should not be less than 3. Refer to Figure 2.2 for an example of null geodesics satisfying these criteria on a flat spacetime; if the spacetime is curved, we only need to consider that metric can be chosen arbitrarily close to Minkowski if the neighbourhood U is small enough.
The geodesics on M induce four curves $\tilde{\gamma}_i = f[\gamma_i]$ on $\tilde{M}$. Thanks to the properties of f, and with an opportune parametrization of the paths $\tilde{\gamma}_i\colon \tilde{I}_i \to f(U)$ they will also be null geodesics and can satisfy conditions analogous to (1)-(4); in particular there is a C^∞ map $\tilde{\psi}\colon \tilde{I}_1 \times \tilde{I}_3 \to \tilde{I}_4$ with $\dfrac{\partial\tilde{\psi}}{\partial\tilde{t}_1}, \dfrac{\partial\tilde{\psi}}{\partial\tilde{t}_3} \neq 0$.
To prove that f maps the parametrization smoothly, consider the four maps $f_i\colon I_i \to \tilde{I}_i$ defined as $f_i := \tilde{\gamma}_i^{-1} \circ f \circ \gamma_i$; they are continuous, as both γ_i and $\tilde{\gamma}_i$ are continuous for all i, and the causal bijection f is a homeomorphism because of strong causality; they are also monotonic, since f satisfies the conditions of Proposition 2.7 and either preserves or reverses the orientation of light cones. It follows that, thanks to Lebesgue's theorem, they are differentiable almost anywhere. Consider the equality

$$\tilde{\psi}(f_1(t_1),\, f_3(t_3)) = \tilde{\psi}(\tilde{t}_1,\, \tilde{t}_3) = \tilde{t}_4 = f_4(t_4) = f_4(\psi(t_1,\, t_3)),$$

and differentiate with respect to t_3,

$$\frac{\partial\tilde{\psi}}{\partial\tilde{t}_3}\, f_3' = f_4'\, \frac{\partial\psi}{\partial t_3}.$$

We obtain an expression on the continuousness of f_4': the first derivatives of

ψ and $\tilde{\psi}$ are non-vanishing by assumption, f_3 is differentiable almost anywhere, and as the argument of f_4 has a double parametrization we can elude singularities by opportune variations on t_1 and t_3. Repeat the procedure on different null geodesics to show that each f_i is at least C^1. An additional differentiation, now with respect to t_1,

$$\frac{\partial^2 \tilde{\psi}}{\partial \tilde{t}_1 \, \partial \tilde{t}_3} f_1' \, f_3' = f_4'' \frac{\partial \psi}{\partial t_1} \frac{\partial \psi}{\partial t_3} + f_4' \frac{\partial^2 \psi}{\partial t_1 \, \partial t_3},$$

returns a similar expression for the second derivative, f_4''. Again, by repeating the process we can show that each f_i is at least C^2, and consecutive iterations can prove that they are in fact C^∞. This proves that f maps the parameter on each null geodesics smoothly.

Now consider n null geodesics $\gamma_j \subset U$, that will act as a kind of "basis" for our charts. The map $\Gamma \colon O \to \mathbb{R}^n$ defined by the intersection of the null cone of a point x with the null geodesics, $\Gamma(x) := \{\gamma_j \cap C(x)\}_{j=1,\dots,n}$, can be chosen to be smooth on a neighbourhood $W \subseteq U$, and we could repeat the procedure above on each geodesic to learn that the smoothness is preserved by f. Then, the set of charts $\{(W, \Gamma)\}$ is a C^∞ atlas for M preserved by the causal bijection, confirming that f is a diffeomorphism. $\square$

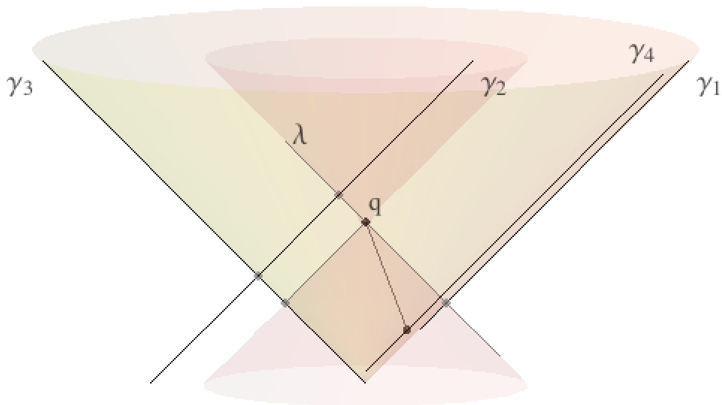

Figure 2.2: Example of the construction required for the theorem on 1+2 dimensions in Minkowski spacetime. γ_1 is parallel to γ_2, and both lie on the same plane as γ_3; γ_4 is parallel to γ_1 but is located behind the plane. The function $\psi(t_1, t_3)$ is given by $\gamma_4^{-1}(t_3 - \frac{1}{1+t_1}, t_3 - \frac{1}{1+t_1}, 2)$, and is clearly C^∞ for $t_1, t_3 \in [0, 1]$ and smooth γ_4.

Finally, the results of Theorem 2.6 and Theorem 2.8 can be combined together:

Theorem 2.9 *Let f be a causal bijection $f\colon M \to \tilde{M}$ between two future and past distinguishing spacetimes of dimension greater than 2. Then, f is a smooth conformal diffeomorphism.*

The conformal geometry of future and past distinguishing spacetimes is completely determined by the causal structure.

Chapter 3

Minkowski metric from causal structure

Causality is determined entirely by the knowledge of the metric, which regulates whether two points in spacetimes are related or not by the definition of a distance. However, the role of the metric is not fundamental, as in turn it can be obtained directly from causal structure up to a conformal factor (cf. Theorem 2.6).

Following the suggestion in [11], we will show in this chapter how this works in the practical case of the 4-dimensional Minkowski spacetime, $\mathbb{M}^4$, with a step by step construction of a framework compatible with its usual geometric structure. We will need the causal structure analogous of light rays, null hypersurfaces, spacelike plane, spacelike lines, arbitrary planes, arbitrary lines, parallel lines, parallelograms and finally vectors to obtain the familiar metric $g = \mathrm{diag}(-1, 1, 1, 1)$.

3.1 Light rays

Definition An *infinite light ray* is a straight line that lies on the null cone of some point $x \in \mathbb{M}^4$, and its points are all null separated from each other.

As it is, this definition depends on information given by the metric, and we need to find the equivalent from causal structure.

Definition A *null chain* is a chain such that the causal interval $J(x,\, y)$ between any two of its elements x, y is also a chain.

For our purposes we will need them to be maximal, i.e. such that the addition of any external point breaks the defining property.
Note that the definition works equally well for curved spacetimes, and so will the following proposition; even if the spacetime is not $\mathbb{M}^4$, light rays can be obtained directly from causality.

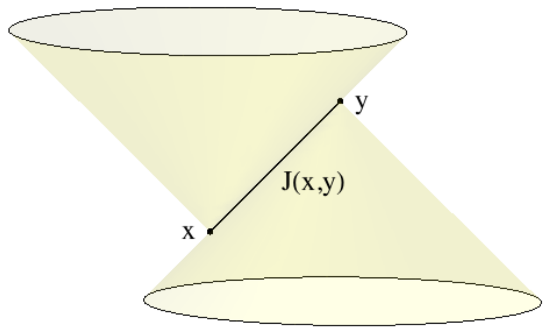

Figure 3.1: Example of points on a null chain. The intersection of the future and past light cones determines the causal interval as a straight line, hence a totally ordered set.

Proposition 3.1 *Infinite light rays correspond to maximal null chains.*

Proof Denote with $\mathcal{L}$ the set of infinite light rays ℓ, and with L the set of maximal null chains λ. The claim is that $\mathcal{L} \equiv L$.
Let's first prove that $\mathcal{L} \subseteq L$. If ℓ is an infinite light ray, by definition all of its elements are null separated, hence it does not contain spacelike points and it is a chain. It is also a null chain, since the causal interval between any two points of ℓ is a segment of light ray and therefore is also a chain. To prove that it is maximal, add a point $z \notin \ell$ to the light ray and consider the new set $\ell \cup \{z\}$; the light ray is infinite and it already includes all points null separated from each other, meaning that there will be some points on ℓ spacelike to z and thus the new set is not a chain. This proves that an infinite light ray is a maximal null chain.

Consider now a maximal null chain λ. Being a chain, it has no spacelike separated elements, and we only need to prove that elements are not timelike separated and that it is infinite. Suppose two points $x, y \in \lambda$ are timelike to each other: in this case there exist two points in the causal interval $J(x,\, y)$ that are spacelike to each other, so that $J(x,\, y)$ is not a chain and λ is not a null chain, getting a contradiction with the assumption. Hence all points in λ are null separated and it is a light ray. λ is also maximal, implying that it contains all points null separated from each other and as a result is infinite. This proves the inclusion $\mathcal{L} \subseteq L$, and consequently $\mathcal{L} \equiv L$. $\square$

3.2 Hypersurfaces

Once we have established the definition of light rays from causal structure, we are now allowed to use them for our purposes and can proceed with the construction of less fundamental objects.
From this point on, however, the flatness of $\mathbb{M}^4$ will start to play a preeminent role and a generalization to curved spacetimes will result non-trivial.

Definition A *null hyperplane* $\mathcal{N}$ is the 3-dimensional flat surface orthogonal to a null vector a_μ. It corresponds to the set of points x satisfying the linear condition $a_\mu x^\mu = c$, where c is a real constant determining the offset from the origin.

Proposition 3.2 *A null hyperplane in* $\mathbb{M}^4$ *corresponds to a light ray and all points spacelike to it.*

Proof Consider coordinates (x^0, x^1, x^2, x^3) for a point $x \in \mathbb{M}^4$.
Thanks to Lorentz invariance, without loss of generality we can assume the light ray $\ell = \big\{x \in \mathbb{M}^4 :\ x = (x^0, x^0, 0, 0)\big\}$ propagating along the x^1 direction. $\mathcal{N} = \big\{x \in \mathbb{M}^4 :\ x = (x^0, x^0, x^2, x^3)\big\}$ is the null hyperplane relative to ℓ. We want to prove that $\mathcal{N}$ is equivalent to $\ell \cup \ell^\natural$, a set given by a light ray ℓ and all points spacelike to it determined using the knowledge of causal relations.
Take a point $y \in \ell^\natural$, spacelike separated to any $x \in \ell$. This translates into

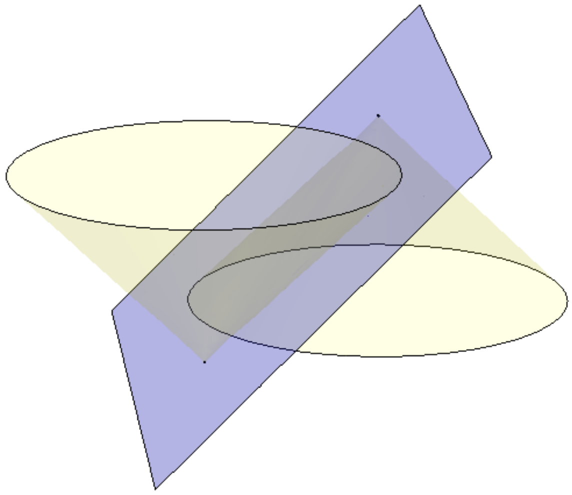

Figure 3.2: Null hyperplane in 1+2 dimensions. Null hyperplanes are tangent to the light cones propagating from each of its points.

the condition $g_{\mu\nu}(x-y)^\mu(x-y)^\nu = y_\mu y^\mu - 2x^0(y^0 - y^1) > 0$, an inequality that holds for all points x if and only if $y^0 = y^1$, so that $y \in \mathcal{N}$. Because $\ell \subset \mathcal{N}$ by assumption, we have that $\ell \cup \ell^\natural \subseteq \mathcal{N}$.

Now take a point $y \in \mathcal{N}$ but $y \notin \ell$, hence of the form (y^0, y^0, y^2, y^3) with either $y^2, y^3 \neq 0$. It is straightforward to verify that it is spacelike to any $x \in \ell$: $g_{\mu\nu}(x-y)^\mu(x-y)^\nu = (y^2)^2 + (y^3)^2 > 0$. Such y is thus an element of $\ell^\natural$, proving that $\mathcal{N} \subseteq \ell \cup \ell^\natural$. □

Null hyperplanes can be used to define planes of lower dimension. However, the flat surfaces determined by the intersection of two null hyperplanes is not arbitrary, since all of its point happen to be spacelike separated.

Proposition 3.3 *The non-empty intersection of two distinct null hyperplanes is a 2-dimensional spacelike plane.*

Proof Choose two distinct null hyperplanes $\mathcal{N}_1 = \{x \in \mathbb{M}^4 : \; a_\mu x^\mu = c\}$ and $\mathcal{N}_2 = \{x \in \mathbb{M}^4 : \; b_\mu x^\mu = d\}$. If they have a non-empty intersection it will satisfy both linear conditions and have two independent normals, resulting in a 2-dimensional plane.

Consider a light ray $\ell \subset \mathcal{N}_1$, and remember that by definition all other points in $\mathcal{N}_1$ will be spacelike separated to it. If all elements of ℓ were also included in $\mathcal{N}_2$, the two null hyperplanes would coincide, but since the premise is that they are different we have that $\ell \cap N_2$ can only be a point. All other elements of $\mathcal{N}_1 \cap \mathcal{N}_2$ will not be elements of ℓ, but they are still elements of $\mathcal{N}_1$, implying that they must be spacelike to the intersection point $\ell \cap N_2$. The 2-dimensional plane determined by the two normals is therefore spacelike. □

In $\mathbb{M}^4$ spacelike planes can be obtained from two null normal vectors, as seen before. The same plane can also be generated by normals that are one null and one spacelike, or one null and one timelike, or even one spacelike and one timelike. As an example, consider the spacelike plane $\Sigma = \left\{x \in \mathbb{M}^4 : \; x = (0,0,x^2,x^3)\right\}$; it can be generated by the null vector $(1,1,0,0)$ and the spacelike vector $(0,1,0,0)$, or similarly from the same null vector and the timelike vector $(1,0,0,0)$.
Linear combinations of two normal vectors is still a normal vector, and we can use this to get both of them to be null and establish the spacelike plane as the intersection of two null hyperplanes.
The intersection of two distinct spacelike planes is a *spacelike line.*

There is a combination of normal vectors that does not allow to choose the two to be null, and that is when the both normals are spacelike.
The non-spacelike plane so determined cannot be obtained from two null hypersurfaces and its elements can indifferently be spacelike, null or timelike separated. For example the plane corresponding to the two spacelike normals $(0,0,1,0)$ and $(0,0,0,1)$ is $\Xi = \left\{x \in \mathbb{M}^4 : \; x = (x^0,x^1,0,0)\right\}$, and points such that $\left|x^0\right| > \left|x^3\right|$ are clearly timelike to $(0,0,0,0)$.

Proposition 3.4 *A non-spacelike plane is given by the union of a light ray, a spacelike line and the set of all other spacelike lines intersecting both.*

Proof Denote the plane with Ξ and choose a point $x \in \Xi$. There will always exist a spacelike line in the plane passing through x. Also, since the plane is non-spacelike, it intersects the light cone of x in at least one light ray,

ℓ, clearly independent from the spacelike line. The rigid translation of the spacelike line along the light ray will give any other point on the plane, so that Ξ is included in the set claimed by the proposition.
The opposite inclusion is trivial: because the requirement is that all spacelike lines intersect both generating lines, they will all lie on the same plane as the original spacelike line and the light ray. This plane is non-spacelike, because at least the points on the light ray are null separated. $\square$

An *arbitrary plane* will either be a spacelike or a non-spacelike plane. The non-empty intersection of two arbitrary planes is an *arbitrary line*.

3.3 Affine space

The concept of parallelism is essential in the definition of an affine space where the vectors will live. This is straightforward in the flat space $\mathbb{M}^4$, but far from trivial if the spacetime is curved.

Definition Two lines are *parallel* if they lie on the same plane and they have empty intersection. A *parallelogram* is the figure determined by the intersection points of two pairs of parallel lines on the same plane.

Define an equivalence relation on the product space $\mathbb{M}^4 \times \mathbb{M}^4$ such that the ordered pair of points (p, q) is equivalent to another pair (p', q') if and only if there exists a parallelogram having p, q, p' and q' as vertices.

Definition The *affine space* $\mathbb{A}$ is the quotient space $\mathbb{M}^4 \times \mathbb{M}^4 / \sim$. The equivalence classes $[(p, q)]$ are denoted with $\overrightarrow{pq}$, and this elements are the *vectors* of the affine space. They will be timelike, spacelike or null according to how p and q are related.

This picture can be consistent with the notion of vector space only after we define the common operations on vectors.
For the sum of two vectors we can follow the familiar parallelogram law, for which all necessary tools are already available. We should only make sure that the vectors are applied at the same point, something that is always

possible in an affine space.
Multiplication by a scalar is a bit more tricky. Define at first multiplication by a positive integer n as the repeated sum of the vector to itself n times. Take the same vector with inverted order to define multiplication by (-1) and consequently by all negative integers. To multiply by a rational number n/m take the vector such that if we multiply it by m we get n times the original one. Finally, establish multiplication by a real number using one of the standard construction from the rationals, such as the Dedekind cuts method, and the fact that any set of points is already endowed with a partial order.

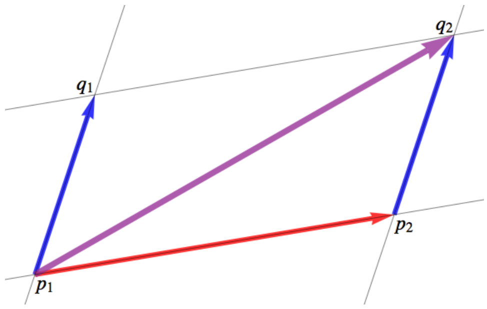

Figure 3.3: Sum of affine vectors. The vector along the diagonal $\overrightarrow{p_1q_2}$ is the sum of $\overrightarrow{p_1q_1} = \overrightarrow{p_2q_2}$ and $\overrightarrow{p_1p_2}$.

Thanks to these properties, $\mathbb{A}$ effectively acts as the tangent space of $\mathbb{M}^4$, although there is no preference for a point as origin and it is not attached to any specific element. This behaviour is peculiar to Minkowski spacetime, and we can make use of it to deduce the metric at any point.

3.4 Metric

We will now demonstrate that the knowledge of all null vectors determines the metric, as stated in Theorem 2.6. Note that the theorem refers to the metric at a fixed point, but in our case it extends automatically to any point thanks to the affine structure.

Definition The *metric* g is defined as a bilinear symmetric function of two vectors, X and Y. Require the norm of a vector $g(X,X)$ to be negative, positive or zero if the vector is respectively timelike, spacelike or null.

Let T, S be a timelike and a spacelike vector and consider the vector given by the linear combination $T+\lambda S$.
We can find two values of λ such that this vector is null: these values correspond to the roots of the quadratic equation $g(T+\lambda S, T+\lambda S) = g(T,T) + 2\lambda\, g(T,S) + \lambda^2\, g(S,S) = 0$, and are known to exist and to be real because $g(T,T) < 0$ and $g(S,S) > 0$. We can use the knowledge of all null vectors to determine the values of the roots λ_1 and λ_2. Note that $g(S,S) = \lambda_1\lambda_2\, g(T,T)$ and $g(T,S) = -\frac{1}{2}(\lambda_1+\lambda_2)\, g(T,T)$.
Now fix the timelike vector to be a specific T_0, and use the described procedure on three different spacelike vectors S_1, S_2 and S_3 to get the roots $\{\xi_1,\xi_2\}$, $\{\eta_1,\eta_2\}$ and $\{\zeta_1,\zeta_2\}$; do the same for the sums S_1+S_2, S_2+S_3 and S_3+S_1 to get roots $\{\rho_1,\rho_2\}$, $\{\sigma_1,\sigma_2\}$ and $\{\tau_1,\tau_2\}$. The aim is to find the value of the metric on any possible combination of these vectors, that act as a basis.
Writing this task in matrix form,

$$g = \begin{pmatrix} g(T_0,T_0) & g(T_0,S_1) & g(T_0,S_2) & g(T_0,S_3) \\ g(S_1,T_0) & g(S_1,S_1) & g(S_1,S_2) & g(S_1,S_3) \\ g(S_2,T_0) & g(S_2,S_1) & g(S_2,S_2) & g(S_2,S_3) \\ g(S_3,T_0) & g(S_3,S_1) & g(S_3,S_2) & g(S_3,S_3) \end{pmatrix},$$

we see that every term can be expressed in function of the known roots and the factor $g_{00} = g(T_0,T_0)$. The result is:

$$g = g_{00} \begin{pmatrix} 1 & A(\xi) & A(\eta) & A(\zeta) \\ A(\xi) & B(\lambda) & C(\rho,\xi,\eta) & C(\tau,\xi,\zeta) \\ A(\eta) & C(\rho,\xi,\eta) & B(\eta) & C(\sigma,\eta,\zeta) \\ A(\zeta) & C(\tau,\xi,\zeta) & C(\sigma,\eta,\zeta) & B(\zeta) \end{pmatrix},$$

where $A(\lambda) = -\frac{1}{2}(\lambda_1+\lambda_2)$, $B(\lambda) = \lambda_1\lambda_2$, $C(\lambda,\mu,\nu) = \frac{1}{2}(\lambda_1\lambda_2 - \mu_1\mu_2 - \nu_1\nu_2)$ are all known thanks to the information given by the null vectors. All of the values are determined but for g_{00}, that represents the conformal factor.

Technically all the work is done, and keeping track of the conditions on the roots imposed by the fact that S_1, S_2 and S_3 are linearly independent we could diagonalize the matrix to get the usual form for the metric, up to the conformal factor.
A quicker way to get the same result is given by the use of normal coordinates. Recall that in the definition of spacetime we require the existence of a set of coordinates such that locally the metric has form $\mathrm{diag}(-1, 1, 1, 1)$. Taking also into account that in $\mathbb{M}^4$ the tangent space at any point is equal and that what holds locally is also true at a global level, we know that the metric will be anywhere:

$$g = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that there is no additional factor in front of the metric with the introduction of normal coordinates. This is because, in this case, the use of the knowledge that spacetime is anywhere the same is equivalent to the contribution given by volume information for the determination of the conformal factor.

3.5 Curved spacetimes

The possibility to extend the results obtained to general, curved spacetime is intriguing. However, even though previous theorems were not specific to Minkowski spacetime and we already know that objects like the metric can be retrieved from causality, generalization is not always easy and might require more than a few considerations.

The definition of a tangent space arising only from causal structure would in particular be of great interest, since it would suggest what the corresponding entity would be when dealing with discrete spacetimes where causal sets are the underlying base for everything, and would therefore allow a not indifferent amount of theories to be transferred onto them, as it is well known that

vectors play a primary role in physics.
The task of generalization in this case may seem almost banal, as the tangent space itself is a local feature and we already know and have been using the fact that, locally, any reasonable spacetime is close to Minkowski. This line of thought leads to wrong results for two main reasons: first, even if arbitrarily small neighbourhoods are taken into account, there will always be minute but non-negligible differences from Minkowski spacetime, otherwise at a local level curved spacetimes would not only be close but agree exactly, with a one-to-one correspondence, to flat spacetimes; second, the definition of tangent space obtained before descends from the affine space, that is a global property of Minkowski spacetime and not a local one as would be required, making it incompatible with our needs.
A partial solution would be to find a new, local construction of the tangent space already on Minkowski. Then, the metric can be used to univocally determine the Levi-Civita connection, making it possible to parallel transport the tangent space thus obtained to any point of spacetime.
It might be argued that none of this work is necessary. Remembering that information on the differential structure is already encoded by causality, we can simply use the standard definition of the tangent space from equivalence classes of tangent differentiable curves. Even if all this is possible, a more constructive method like the one described before would be preferable, both for its direct uses and a more open dependance on causal relations.

Chapter 4

Alternative topologies

The realization of a topology directly from causal structure is a perfect example of the amount of information encoded about the spacetime. For this purpose we already introduced the Alexandrov topology $\mathcal{A}$ (cf. Section 2.2), which happens to coincide exactly with the original manifold topology $\mathcal{M}$ when the spacetime is strongly causal.

Many other kinds of topologies have been introduced in the past, each with its own advantages and disadvantages. We will present some of them, starting from Zeeman-type topologies and proceeding with other topologies based on convergence criterions that join the properties of the former to the desired derivation from causality and other characteristics observed in $\mathcal{A}$.

We will be mostly interested on the defining properties and how they differ from each other, trying to clarify the motivation behind each new topology and what advantages they bring over the previous ones; the main properties will be presented, with lesser or greater depth according to the relevance of the discussion, and the homeomorphism group will be mentioned when known.

For a complete analysis of any of the following topologies we recommend a consultation of the respective original articles.

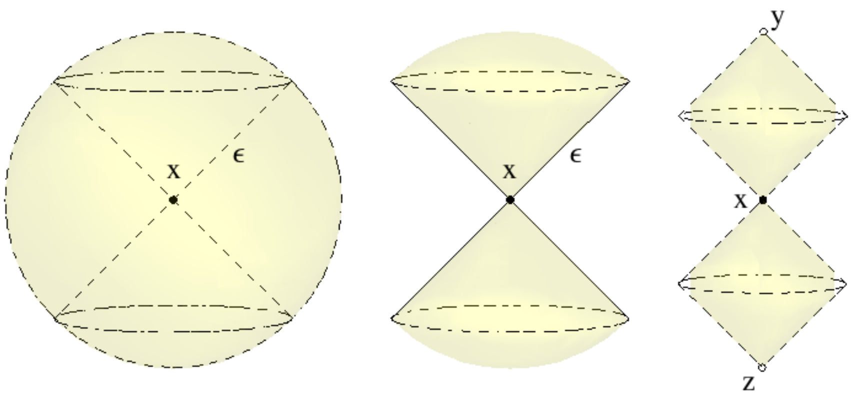

Figure 4.1: From left to right, the $\mathcal{Z}$-open set $Z_\varepsilon(x)$, the $\mathcal{P}$-open set $L_\varepsilon(x)$, and the $\mathcal{F}$-open interval $I(y, x, z)$. The sets are only examples on a plain 3-dimensional spacetime; dashed lines indicate removed points.

4.1 $\mathcal{Z}$ - Zeeman topology

Zeeman suggested a topology for Minkowski spacetime in 1966, described in his work "*The Topology of Minkowski Space*" [2].
Although designed on a very specific case, it is worth of mention because of the ideas proposed and to better comprehend the origin and motivation of later works. An extension to general-relativistic spacetimes was proposed in "*Zeeman Topologies on Space-Times of General Relativity Theory*" [6] by Göbel, replacing in the subsequent discussion timelike lines with timelike geodesics and spacelike hyperplanes with spacelike hypersurfaces.

The main motivation driving Zeeman is the fact that the original topology $\mathcal{M}$ does not reflect the physics underlying the spacetime, and in addressing this problem he gets to the definition of the new topology, $\mathcal{Z}$.
The undesired quality of $\mathcal{M}$ is the fact that it is homogeneous and locally homogeneous, ignoring any difference between space and time and preventing a natural emergence of light cones; its group of homeomorphisms is also too abundant, as some even allow swaps between spatial and temporal axes, and clearly not all of them can possibly be physical.

To overcome the homogeneity of the original topology, Zeeman treats time and space dimensions separately, although still without a proper distinction between them.

Definition The *Zeeman topology* $\mathcal{Z}$ is the finest topology on the manifold to induce 1-dimensional Euclidian topology on any timelike line and 3-dimensional Euclidian topology on any spacelike hyperplane.

It seems that to determine $\mathcal{Z}$ we still need the notion of the original topology. This dependency is made more evident if we consider the condition for a set to be open in $\mathcal{Z}$, following directly from the definition.

Proposition 4.1 *$O \subseteq M$ is $\mathcal{Z}$-open $\iff$ $O \cap \tau$, $O \cap \Sigma$ are $\mathcal{M}$-open, for any timelike line τ and spacelike hyperplane Σ.*

$\mathcal{M}$-openness on lines or hyperplanes is intended relative to the induced topology on the subspaces, with open balls of corresponding dimension.
The reliance on $\mathcal{M}$ is one of the weak points of $\mathcal{Z}$, when the aspiration is to construct a topology only from causal structure.

Since $\mathcal{M}$ is clearly among the topologies satisfying the defining condition, it follows that $\mathcal{Z}$ is finer. There are also coarser topologies than $\mathcal{M}$ that still meet the requirement, implying that its choice as the topology for M seems to be unjustified if this physical constraint is accounted for.
$\mathcal{Z}$ is actually strictly finer than $\mathcal{M}$:

Proposition 4.2 *$\mathcal{Z} \supsetneq \mathcal{M}$.*

An example of sets open in $\mathcal{Z}$ but not in $\mathcal{M}$ is given by $Z_\varepsilon(x) := (B_\varepsilon(x) \setminus C(x)) \cup \{x\}$, i.e. open balls with all points on the light cone of x but x itself subtracted (Figure 4.1, left). They are $\mathcal{Z}$-open because any timelike line or spacelike hyperplane is untouched and has the original topology, but the point x itself does not admit an open ball as neighbourhood and therefore they are not $\mathcal{M}$-open. These sets however are not enough to form a basis for the entire topology (for example the same set with a sequence of points converging in $\mathcal{M}$ but not in $\mathcal{Z}$ removed from it is also $\mathcal{Z}$-open), nor does $\mathcal{Z}$ admit a countable basis of open neighbourhoods.

A consequence of Proposition 4.2 is that $\mathcal{Z}$ inherits some of the properties of $\mathcal{M}$:

Proposition 4.3 *$\mathcal{Z}$ is Hausdorff.*

Proof The topology is Hausdorff if for any two distinct points $x, y \in M$ there exist open neighbourhoods $O_x \ni x$ and $O_y \ni y$ such that $O_x \cap O_y = \emptyset$. Since $\mathcal{M}$ is Hausdorff, these two sets exist and are $\mathcal{M}$-open. Since $\mathcal{Z}$ is finer than $\mathcal{M}$, they are also $\mathcal{Z}$-open, implying that $\mathcal{Z}$ is also Hausdorff. $\square$

Following are other properties of $\mathcal{Z}$ not necessarily descending from $\mathcal{M}$.

Proposition 4.4 *$\mathcal{Z}$ is connected and locally connected. $\mathcal{Z}$ is not normal, compact or locally compact.*

Proposition 4.5 *The topology induced on light rays by $\mathcal{Z}$ is discrete.*

Proof For any point x of a light ray ℓ, to induce the topology on the light ray we must intersect $\mathcal{Z}$-open neighbourhoods of x with ℓ itself. If the set $Z_\varepsilon(x)$ is chosen, then $Z_\varepsilon(x) \cap \ell \equiv \{x\}$ is the open set induced on ℓ. Hence every point of ℓ is an open subset, and its topology is discrete. $\square$

A motive for a discrete topology on light rays is provided by the evidence that photons can be observed only during discrete events of emission and absorption.

Another significant quality of $\mathcal{Z}$ is that its group of homeomorphisms is smaller and more reasonable than the one relative to $\mathcal{M}$, adding to the reasons why this topology is more physical.

Proposition 4.6 *f is a $\mathcal{Z}$-homeomorphism $\iff$ f is a dilation, a translation or a Lorentz transformation.*

The statement is intuitively evident, as dilations, translations, and Lorentz transformations all maintain the straightness of lines and hyperplanes, thus preserving the topology.
The implication is that the invariants of physics given by Lorentz representation can be directly deduced from $\mathcal{Z}$ itself, instead of being imposed on it separately.

Finally, Zeeman suggests a few alternatives based on the same principle.
If a discrete topology is desired not only on light rays but on spacelike axes as

well, a topology $\mathcal{Z}^\natural$ could be defined correspondingly as the finest topology to induce the 1-dimensional Euclidian topology on timelike lines only. This allows an extension of Proposition 4.5 to the spacelike dimensions if the $\mathcal{Z}^\natural$-open neighbourhoods used are the sets obtained by the intersection of open balls with the chronological past and future of the point, plus the point itself.
Applications to this topology might be found when dealing with lattices or more generally with the discreteness of matter, although it is not qualified for a discrete spacetime as only the space coordinates are discretized.

In summary, the Zeeman topology represents a first step in the definition of a topology for spacetime that is more suitable to reflect its physical properties than the original topology. It still presents a few inadequacies, like its dependance on the original topology and the lack of a derivation from pure causality, but has a good predisposition for generalizations and improvements.

4.2 $\mathcal{P}$ - Path topology

Proceeding from Zeeman's idea, Hawking, King and McCarthy introduce a new topology based on arbitrary timelike lines and not limited to inertial observer evolving along straight lines. This topology is described in "A new topology for curved space-time which incorporates the causal, differential, and conformal structures" [7] and is suitable for general spacetimes.

The authors argue that, even if Zeeman's topology is appealing for its physical significance, it incorporates aspects with no concrete application. For example, $\mathcal{Z}$-homeomorphisms include dilations, transformations that could never preserve the volume of the total spacetime. Also, there is no manifest reason to take into account spacelike hyperplanes (or hypersurfaces), as spacelike separated events are incomparable.

Differently from what done by Zeeman, the approach to the path topology aims not just to separate time and space, but to give them the appropriate physical significance. For this reason attention is drawn away from spacelike entities to focus more on the timelike evolution of events, verifiable by observation.

Definition The *path topology* $\mathcal{P}$ is the finest topology on the manifold to induce 1-dimensional Euclidian topology on arbitrary timelike curves.

$\mathcal{P}$ is still a Zeeman-type topology and is dependent on the original topology, although this time the defining requirement is extended to incorporate all timelike paths. We will show that, differently from $\mathcal{Z}$, this allows for the existence of a basis for the topology as pathological cases occurring because of sequences converging in other topologies other than $\mathcal{P}$ are precluded. The condition for a set to be $\mathcal{P}$-open is descended from the definition:

Proposition 4.7 *$O \subseteq M$ is $\mathcal{P}$-open $\iff$ $O \cap \tau$ is $\mathcal{M}$-open, for any timelike line τ.*

An argument analogous to the one for $\mathcal{Z}$ leads to the conclusion that $\mathcal{P}$ is (strictly) finer than $\mathcal{M}$.

Proposition 4.8 *$\mathcal{P} \supsetneq \mathcal{M}$.*

As sets open in $\mathcal{P}$ but not in $\mathcal{M}$ choose the intersection of open balls with the chronological past and future of a point, including the point itself: $L_\varepsilon(x) := B_\varepsilon(x) \cap (I^+(x) \cup I^-(x) \cup \{x\})$ (Figure 4.1, middle).
Sets of this kind are not even $\mathcal{Z}$-open, since on any spacelike hypersurface they induce a discrete topology and not the Euclidean one as already pointed out in the consideration of topologies alternative to $\mathcal{Z}$.
As a matter of fact, the two topologies happen to be incomparable.

Proposition 4.9 *$\mathcal{P}$ is incomparable to $\mathcal{Z}$.*

Proof The evidence of $\mathcal{P}$-open sets that are not open in $\mathcal{Z}$ has already been provided by the sets L_ε.
To obtain a set open in $\mathcal{Z}$ but not in $\mathcal{P}$ consider $Z_\varepsilon(x)$ and subtract from it all points lying on a non-straight timelike curve τ passing through x but x itself. The curve can be chosen so that it does not intersect any timelike line,

meaning that the topology of any timelike line and hyperplane intersecting the set is Euclidean and the set is $\mathcal{Z}$-open. On the other hand it is not $\mathcal{P}$-open, as there exist a timelike curve, τ, such that no Euclidean topology is induced on their intersection, $\{x\}$. □

$\mathcal{P}$ further differs from $\mathcal{Z}$ in the fact that it admits a countable basis of open sets, making it more functional for practical purposes.

Proposition 4.10 *The sets $\{L_\varepsilon\}$ form a basis for the topology $\mathcal{P}$.*

Like $\mathcal{Z}$, the path topology benefits of various properties, some derived from its being finer than $\mathcal{M}$.

Proposition 4.11 *$\mathcal{P}$ is Hausdorff, connected and locally connected. $\mathcal{P}$ is not regular, normal, compact or locally compact.*

While the homeomorphism group of $\mathcal{Z}$ includes dilations, preserving the metric up to a constant factor, the homeomorphism group of $\mathcal{P}$ coincides with the group of conformal transformations, preserving the metric up to a variable factor. This allows to protect the total volume of spacetime and makes the new topology more desirable for a physical theory.

Proposition 4.12 *f is a $\mathcal{P}$-homeomorphism $\iff$ f is a conformal transformation, including translations and Lorentz transformations.*

$\mathcal{P}$-homeomorphisms have the interesting feature that they coincide with homeomorphisms on the Alexandrov topology under very weak conditions.

Proposition 4.13 *If M is chronological, then a $\mathcal{P}$-homeomorphism f is an $\mathcal{A}$-homeomorphism.*

This property comes from the fact that a continuous map on $\mathcal{P}$ maps light cones into light cones and chronological pasts/futures to chronological pasts/futures.
A further consequence is that if the spacetime is strongly causal, so that the Alexandrov and the manifold topology coincide, then the $\mathcal{P}$-homeomorphism is also an $\mathcal{M}$-homeomorphism.

To sum up, the path topology answers to most of the disadvantages of $\mathcal{Z}$, like the inappropriateness for direct applications due to the lack of a countable basis, and has a more explicit physical interpretation of timelike evolution. However, there are some issues common to Zeeman-type topologies still unresolved, and the connection with causality is not sufficiently explicit.

4.3 $\mathcal{F}$ - Fullwood topology

As a result of private talks with McCarthy, Fullwood was inspired to take the connection between $\mathcal{P}$ and the Alexandrov topology to a further level. In doing so he defines a new topology, $\mathcal{F}$, presented in "*A new topology on space-time*" [10].

The path topology presents a variety of advantages, but has a problem at the core of its definition: it is not independent of the original topology and is not fit to replace it unconditionally. On the other hand the Alexandrov topology does not rely on $\mathcal{M}$ for its definition, but it still presents the same problems of the original topology that motivated Zeeman in the first place. A solution to this is provided by $\mathcal{F}$, that has its origin in causal structure but still possesses most of the characteristics of $\mathcal{P}$. Several results suggest that in distinguishing spacetimes $\mathcal{F}$ relates to $\mathcal{P}$ like in strongly causal spacetimes $\mathcal{A}$ relates to $\mathcal{M}$.

Definition The *Fullwood topology* $\mathcal{F}$ is the topology generated by the set of double chronological intervals $I(z,\, x,\, y) := I(z,\, x) \cup \{x\} \cup I(x,\, y)$.

Note how the double intervals $I(z,\, x,\, y)$ (Figure 4.1, right) resemble a combination of the light-cone neighbourhoods $L_\varepsilon(x)$ for the path topology and the chronological intervals of the Alexandrov topology.
The same topology can be obtained with a notion of converging sequences.

Definition A sequence of events $\{x_n\}$ is said to *causally converge* to x if and only if either $I^+(x) = \bigcup_n I^+(x_n)$ or $I^-(x) = \bigcup_n I^-(x_n)$, i.e. the future/past of x is given by the union of all futures/pasts of the sequence.

Let $\tilde{\mathcal{F}}$ be the collection of sets U with the property that if $\{x_n\} \subset M \setminus U$ is a monotonic chronological sequence in the complement causally converging to x, then $x \in M \setminus U$.

Proposition 4.14 $\mathcal{F} \equiv \tilde{\mathcal{F}}$

Proof Let's prove first that $\mathcal{F} \supseteq \tilde{\mathcal{F}}$, i.e. that given a set $U \in \tilde{\mathcal{F}}$ there exists a subset $I(z, x, y) \subseteq U$ about any point $x \in U$. Without loss of generality, let's prove the existence of an interval $I(x, y) \subseteq U$; then, reflecting the argument for a z to the past obtain $I(z, x)$ and join the sets to get the double interval. Suppose the thesis statement is wrong, and for any interval $I(x, y)$ there exists some point $p \in I(x, y)$ but $\notin U$. If we choose $y_1 \in I^+(x)$, then there is a $p_1 \in I(x, y_1) \setminus U$; choose then $y_2 \in I(x, p_1)$, and again a $p_2 \in I(x, y_2) \setminus U$, and proceed until we get two sequences $\{y_n\}$ and $\{p_n\}$. If $\{y_n\}$ is chosen to causally converge to x, then $\{p_n\}$ also causally converges to x. But $\{p_n\} \subset M \setminus U$, hence by definition of $\tilde{\mathcal{F}}$ its limit is also in $M \setminus U$, and $x \notin U$ in contradiction with the original assumption.
For the inverse inclusion, $\mathcal{F} \subseteq \tilde{\mathcal{F}}$, we need to prove that any double interval $I(z, x, y)$ is an element of $\tilde{\mathcal{F}}$. Let $\{x_n\} \subset M \setminus I(z, x, y)$ be a monotonic chronological sequence causally converging to x. Without loss of generality consider the case $I^+(x) = \bigcup_n I^+(x_n)$. Then $x \ll x_n$ for all n, and also there exists some $\bar{n}$ such that $x_{\bar{n}} \ll y$, i.e. the sequence eventually precedes z. Consequently $x_{\bar{n}} \in I(x, y) \subseteq I(z, x, y)$. But $x_n \notin I(z, x, y)$ for all n by assumption, hence we reach again a contradiction and any sequence causally converging to $x \in I(z, x, y)$ is not contained in the complement. On the other hand, we could repeat the argument for any sequence converging to any other point p of $I(z, x, y)$ by choosing subsets $I(x, p, y)$ or $I(z, p, x)$, according to whether $x \ll p$ or $p \ll x$. This means that any sequence causally converging to a point in $I(z, x, y)$ is not contained in the complement, or that any causally converging sequence in the complement has its limit in the complement itself, and proves the inclusion. □

Thanks to their being equivalent, we can use the notation $x_n \xrightarrow{\mathcal{F}} x$ to denote that the sequence $\{x_n\}$ causally converges to x.

In view of the analogy between $\mathcal{F}$ to $\mathcal{A}$, we see that the Fullwood topology is generally coarser than the path topology, as the Alexandrov topology is generally coarser than the manifold topology.

Proposition 4.15 $\mathcal{F} \subseteq \mathcal{P}$.

Proof We have to prove that for any $p \in I(z, x, y)$ there exists a local neighbourhood $L_\varepsilon(p) \subseteq I(z, x, y)$.
If $p \in I(z, x)$ or $\in I(x, y)$, there exists an open ball $B_\varepsilon(p) \subseteq I(z, x, y)$ since the intervals are $\mathcal{M}$-open sets, and therefore there also is a $L_\varepsilon(p) \subseteq B_\varepsilon(p)$.
If $p = x$, again we have that $I^+(x)$ and $I^-(x)$ are $\mathcal{M}$-open, and we can designate $L_\varepsilon(x) = \{q \in I^+(x) \cup I^-(x) \; : \; \|x - q\| < \varepsilon\} \cup \{x\}$. □

The parallelism of $\mathcal{F}$ and $\mathcal{A}$ goes further with the following proposition, that echoes Theorem 2.5.

Proposition 4.16 *M is future and past distinguishing $\iff$ $\mathcal{F}$ coincides with $\mathcal{P}$.*

Proof For the implication that a distinguishing spacetime makes the two topologies agree we will give only the idea. We already know that $\mathcal{F}$ is coarser than $\mathcal{P}$, and need two prove the opposite inclusion, $\mathcal{F} \supseteq \mathcal{P}$. To do this, we need to find a double interval $I(z, x, y)$ for any point x of a $\mathcal{P}$-open set $L_\varepsilon(p)$. There are two different cases. If $x = p$, both past and future are $\mathcal{M}$-open and we can select two points $y \in I^+(x) \cap L_\varepsilon(p)$ and $z \in I^-(x) \cap L_\varepsilon(p)$ to form the double interval $I(z, x, y) \subseteq L_\varepsilon(p)$. If instead $x \in I^+(p)$ (or equivalently $I^-(p)$), we need the future and past distinguishing condition to make sure of the existence of a local neighbourhood inside $L_\varepsilon(p)$ where the strong causality condition holds and $\mathcal{A} \equiv \mathcal{M}$, so that intervals contained in $L_\varepsilon(p)$ and having x as limit for future and past monotonic chronological sequences respectively can be found; two of these intervals can then be used to form the double interval.
The opposite implication proceeds similarly to the Alexandrov case. Suppose the two topologies coincide but there exist two distinct points x, y where either the future or past distinguishing condition fails (we will consider the former without loss of generality). Let $I(c, a, b) \ni x$ and $I(r, p, q) \ni y$ be

two $\mathcal{F}$-open neighbourhoods of the points, and find two subsets $I(d, x, e) \subseteq I(c, a, b)$ and $I(s, y, t) \subseteq I(r, p, q)$. Since $I^+(x) = I^+(y)$, we have that $x \ll e$ implies $y \ll e$, meaning that $I(y, e) \cap I(y, t) \neq \emptyset$, or equivalently $I(x, e) \cap I(y, t) \neq \emptyset$, so that $I(d, x, e) \cap I(s, y, t) \neq \emptyset$. This chain of implications applies to all $\mathcal{F}$-neighbourhoods of x and y, violating the Hausdorff condition that $\mathcal{F}$ would satisfy if it agreed with $\mathcal{P}$. □

Even if the definition of the path topology might seem arbitrary and dependent on external elements, this result confirms its causal nature when the spacetime is reasonable. It is also implied that $\mathcal{F}$ is not Hausdorff unless the distinguishing condition is satisfied.
Among other properties of $\mathcal{F}$ we find the following:

Proposition 4.17 *The set of points where the future and past distinguishing condition holds is open in $\mathcal{F}$.*

Equivalently, we could say that the set of points where the distinguishing condition fails is $\mathcal{F}$-closed, and since $\mathcal{P}$ is finer, also $\mathcal{P}$-closed.

The group of homeomorphisms for $\mathcal{F}$ is known for a chronological spacetime.

Proposition 4.18 *f is a $\mathcal{F}$-homeomorphism $\iff$ f is a chronological or anti-chronological bijection.*

This makes perfect sense in the global picture. If the spacetime is future and past distinguishing, Theorem 2.9 implies that causal bijections (or equivalently chronological bijections) are smooth conformal transformations, and the homeomorphism group of $\mathcal{F}$ coincides with the homeomorphism group of $\mathcal{P}$, confirming the agreement of the two topologies under such condition.

The Fullwood topology acts as a bridge between the path topology and the Alexandrov topology, representing an alternative of causal origin to the former in spacetimes that are at least future and past distinguishing. The inspiration derived from a topology based on intervals accounts for a number of properties involving $\mathcal{F}$ and $\mathcal{P}$ matching analogous properties for $\mathcal{A}$ and $\mathcal{M}$, like the parallelism between the open sets and the agreement under a precise causality condition. This last feature is particularly useful if a topology of great physical interest and arising only from causality is desired for spacetimes satisfying the relatively weak condition of future and past distinction, although is not helpful if the original manifold topology is wished instead.

4.4 $\mathcal{N}$ - Null chain topology

In "*Causal Topology in Future and Past Distinguishing Spacetimes*" [14] by Parrikar and Surya a new topology is suggested, based on the definition of a convergence criterion on null chains and inspired by a construction using null geodesics in the main theorem of Malament's publication [8].

The new topology, based on a convergence of null chains, tries to fill the gap between strongly causal and future and past distinguishing spacetimes, acting as an alternative to the Alexandrov topology in the latter case. Unlike the Fullwood topology, that in distinguishing spacetimes coincides with the path topology, the new topology under the same condition is shown to be equivalent to the manifold topology. Moreover, the fact that the topology descends from something as basic as a convergence criterion makes it more available for an axiomatization of the theory.

There are several kinds of convergences that we should already be familiar with, starting from the Euclidian limit on the original topology:
$x_n \xrightarrow{\mathcal{M}} x \iff \forall O \in \mathcal{M},\ \exists \bar{n} :\ \forall n \geq \bar{n},\ x_n \in O.$
A similar definition applies to a sequence converging in the Alexandrov topology:
$x_n \xrightarrow{\mathcal{A}} x \iff \forall O \in \mathcal{A},\ \exists \bar{n} :\ \forall n \geq \bar{n},\ x_n \in O.$
Since $\mathcal{M}$ is finer than $\mathcal{A}$, manifold convergence implies Alexandrov convergence, although it should be specified that the origin of the latter is purely causal. Both of them can be extended from a sequence of points $\{x_n\}$ to a sequence of open sets $\{O_n\}$, as long as it is required that for all $n \geq \bar{n},\ O_n \cap O \neq \emptyset$ instead. Note that if $x \in O_n$ for all n, this convergence is trivial.
Convergence in the Fullwood topology has a different definition:
$x_n \xrightarrow{\mathcal{F}} x \iff$ either $I^+(x) = \bigcup_n I^+(x_n)$ or $I^-(x) = \bigcup_n I^-(x_n)$.

Before defining the new convergence criterion, recall that null chains are totally ordered set of points such that the causal interval between any two of its points is also totally ordered. Also, denote with Δ the set of points in M where the strong causality condition fails (the spacetime itself is required to be just causal).

Definition A sequence of events $\{x_n\}$ is said to $\mathcal{N}$*-converge* to x if and only if there exists a future (or past) directed null chain $\Omega = J(x,\, y)$ (or $J(z,\, x)$), with $\Omega \setminus \{x\} \subseteq M \setminus \Delta$, and a sequence of future (or past) non-intersecting null chains $\{\Omega_n = J(x_n,\, y_n)\}$ (or $J(z_n,\, x_n)$) such that any point of $\Omega \setminus \{x\}$ is an Alexandrov limit of $\{\Omega_n\}$ and any subsequence of $\{\Omega_n\}$ has no limit out of Ω.

Denote $\mathcal{N}$-convergence with $x_n \xrightarrow{\mathcal{N}} x$. We can then use this null chain convergence criterion to define a topology.
Let $\mathcal{N}$ be the collection of sets U with the property that if $x_n \xrightarrow{\mathcal{N}} x$, $\{x_n\} \subset M \backslash U$, then $x \in M \backslash U$, that is any $\mathcal{N}$-converging sequence in the complement has its limit in the complement.

Proposition 4.19 *$\mathcal{N}$ is a topology.*

Proof We need to verify that the three defining properties of a topology are satisfied. Let $\{U_i\} \subseteq \mathcal{N}$.
$M, \emptyset \in \mathcal{N}$. This is obvious, as the complement of M is $\emptyset$ and cannot possibly have any sequence, and the complement of $\emptyset$ is M and any converging sequence must converge to some point of the manifold itself.
$\bigcup_i U_i \in \mathcal{N}$. For any $\mathcal{N}$-converging sequence $\{x_n\} \subset M \setminus \bigcup_i U_i$, then $\{x_n\} \subset M \setminus U_i$ for any i. If the sequence converges to x, since every U_i is $\mathcal{N}$-open it means that $x \in M \setminus U_i$ for all i. Then $x \notin \bigcup_i U_i$, and $\bigcup_i U_i$ is $\mathcal{N}$-open.
$\bigcap_i U_j \in \mathcal{N}$ for finite $\{U_j\}$. For any sequence $\{x_n\} \subset M \backslash \bigcap_j U_j = \bigcup_j (M \backslash U_j)$, since the sets are in finite number there must exist some value k for which $M \setminus U_{\bar{k}}$ contains an infinite subsequence of $\{x_n\}$. Then, if $x_n \xrightarrow{\mathcal{N}} x$, any subsequence also $\mathcal{N}$-converges to the same point. As U_k is $\mathcal{N}$-open it follows that $x \in M \setminus U_k$, and consequently $x \in \bigcup_j (M \setminus U_j) = M \setminus \bigcap_j U_j$, and $\bigcap_j U_j$ is also open in $\mathcal{N}$. $\square$

By virtue of the concerns with strong causality handled in the definition, the requirement of $\mathcal{N}$-convergence seems to be weaker than that of Alexandrov convergence, as it is only necessary that the spacetime is future and past distinguishing for it to coincide with the original manifold convergence.

Proposition 4.20 *If M is future and past distinguishing, and the region of strong causality violation Δ is either empty or locally achronal, then $\mathcal{N}$-convergence is equivalent to $\mathcal{M}$-convergence.*

Immediate consequence is that, under the same conditions, the two topologies coincide.

Proposition 4.21 *If M is future and past distinguishing, and the region of strong causality violation Δ is either empty or locally achronal, then $\mathcal{N}$ coincides with $\mathcal{M}$.*

The result is significant, indicating that the topological structure of a spacetime is more complex than one could expect. According to the causality condition holding on M, the same topology $\mathcal{M}$ could be equivalent to an entire range of other topologies differently defined. The level of refinement required is not the only relevant parameter in the choice of a topology for spacetime, and the spacetime's own behaviour should be taken into account.

Even if $\mathcal{M}$ admits a countable basis and $\mathcal{N}$ coincides with it under the distinguishing condition, it is not easy to find a neighbourhood basis for the new topology. This is because causal intervals and other sets of causal nature are not necessarily local if the spacetime is not strongly causal.
Still it is not an impossible task; intersecting causal intervals with an $\mathcal{M}$-open ball still produces a local neighbourhood, and we could use the same method to construct local neighbourhoods in the $\mathcal{N}$-topology. The use of the manifold topology as aid in the construction should not be a concern for physical theories of purely causal origin as long as the spacetime is distinguishing.

The null chain topology is a causal topology that should be used to recover the original topology on spacetimes that are not robust enough to satisfy the strongly causal condition but still has the distinguishing property. Although its definition has no apparent relation to that of the Alexandrov topology, it is still emerges from causality and is thought as an alternative to $\mathcal{A}$. There are some issues that unless solved would make this topology rather unpractical. Among them are the fact that the homeomorphism group is unknown and the absence of a recognized basis of open sets.

Conclusions

As a consequence of the theorems presented throughout the dissertation, it would be unfair to treat causality as just a product of the relations emerging from the metric, and should occupy a fundamental position in the structure of spacetime. This becomes especially true for the most “causal” spacetimes, whose behaviour is non-pathological and prevents the manifestation of physical paradoxes.

Up until now the continuousness of spacetime was never doubted, even though one of the main uses of causal structure is its application to the problem of whether the continuous spacetime that we experience reflects only an approximation of a more fundamental configuration of discrete nature, and how can this approximation be obtained from it. This is done with the employment of *causal sets* theory, which relies purely on a discrete set of points, “atoms” of spacetime, and their causal relations.
The question is whether it is always possible to reconstruct the continuous case only from these elements, and even if all the results from Chapter 2 seem to point in that direction, the general answer is no. It is easy to think of pathological examples that cannot be recreated through the use of causal sets, but it’s not easy to demarcate the exact point beyond which it is always possible to get the whole spacetime.

One of the examples is given by the cylinder described by Hawking and Ellis [5]. Here the 2-dimensional spacetime is periodic along the spatial axis with cylindrical conditions; light cones are tilted of 45° at $t = 0$ so to create a unique null geodesic going round as a circle, and are straightened up when approaching the lower and top limits of the cylinder. The existence of a null geodesic going on a loop means that the spacetime is not causal. The main argument that explicits the impossibility of recovery of this case from discrete causal sets goes as following: consider two points, x and y, on opposite sides of the circular null geodesic, and consider another point z just to the future (or past) of x; from a causal point of view, there is no way to tell that z is much closer to x than y, as these two points have exactly the same past and future; all the points on the loop behave identically and could somehow be considered as the same point, and a retrieval of the topological properties from causality would be impossible. A point could be removed from the loop to make the spacetime causal, although still not future nor past distinguishing; this removal is not registered in the discrete case, as single points in the continuum are only a result of the approximation and have no direct correspondent, and the problem persists. Similar arguments can be used when only future but not past distinction holds, or vice versa, as described by Malament [8].

On the other hand, future and past distinguishing spacetimes already possess all the necessary requirements to make their reconstruction possible, from the topology to the differential structure.

The dividing line for the determination of success or failure of the reconstruction seems to lie in between these causality conditions, with future or past distinguishing spacetimes as a gray area. For these spacetimes it would still be possible to define a topology, like the one introduced by Parrikar and Surya based on null chain convergences, although it will not coincide with the original topology and leaves the complications unsolved.

Spacetimes where even future or past distinction fails could be even more problematic, and in some cases deny entirely the likelihood of success, as in the example above. It should be noted however that there are two different ways for future or past distinction to fail. In one case there exist points such that both future and past coincide, and it is closer to the causal condition in the hierarchy; the instance of the cylinder falls in this category. In the other case there are some points with the same past and some with the same future, but two distinct points never share both of them; this means that at least locally the spacetime is either future or past distinguishing, and in a way it echoes the local strongly causal behaviour of a distinguishing spacetime. The implications of this dissimilarity and the attempt to narrow the uncertainty of reconstruction on these spacetimes could be object of further investigations.

The complexity of the topic appears to be related to the variety of causality conditions that could hold on the spacetime. This is also cause for the abundance of alternative topologies, usually different from each other unless a specific condition is satisfied. For future and past distinguishing spacetimes for example we have more than one choice: if the original topology is required, it can be obtained from causality through the alternative based on null chains; otherwise, if the topology should reflect more the physical constitution of spacetime, another choice based on double interval and still arising from causality is possible, leading to the path topology. Causality is not the only reason for the diversification, and other physical requirements also play a central role for the choice of a topology.
It could be conjectured that the same is true for other fabrics, like the differential structure. It is already known that the one we are familiar with is not the only possibility, and a number of exotic differential structures is possible depending on the manifold. An interesting manifestation, possibly of physical origin, is represented by the n-dimensional Euclidian spaces, that admit only one differential structure unless the dimension is 4, in which case the exotic alternatives are uncountable. An account of the role of exotic differential structure in physics can be found in "*Exotic smoothness and physics: differential topology and spacetime models*" [12] by Asselmeyer-Maluga and Brans.

Nevertheless, the information encoded in causality has the potential to extend to each different aspect of spacetime, underlying and influencing its different structures to a lesser or greater extent. Once the causality condition is sufficiently strong, nothing should prevent the decodification of this information for the reconstruction of the spacetime.

Bibliography

[1] S. W. Hawking,
Singularities and the Geometry of Space-Time
Adams Prize Essay (1966) 116 - unpublished

[2] E. C. Zeeman,
The Topology of Minkowski Space
Topology 6 (1967) 161

[3] E. H. Kronheimer, R. Penrose,
On the structure of causal spaces
Proc. Camb. Phil. Soc. 63 (1967) 481

[4] R. Penrose,
Techniques of Differential Topology in Relativity
Society for industrial and applied mathematics, Philadelphia, PA (1972)

[5] S. W. Hawking, G. F. R. Ellis,
The large scale structure of space-time
Cambridge U. P., Cambridge (1973)

[6] R. Göbel,
Zeeman Topologies on Space-Times of General Relativity Theory
Commun. Math. Phys. 46 (1976) 289

[7] S. W. Hawking, A. R. King, P. J. McCarthy,
A new topology for curved space-time which incorporates the causal, differential, and conformal structures
J. Math. Phys. 17 (1976) 174

[8] D. B. Malament,
The class of continuous timelike curves determines the topology of space-time
J. Math. Phys. 18 (1977) 1399

[9] A. V. Levichev,
Prescribing the conformal geometry of a Lorentz manifold by means of its causal structure
Soviet Math. Dokl. 35 (1987) 452

[10] D. T. Fullwood,
A new topology on space-time
J. Math. Phys. 33 (1992) 2232

[11] R. D. Sorkin,
Causal Sets: Discrete Gravity
arXiv:gr-qc/0309009v1 (2003)

[12] T. Asselmeyer-Maluga, C. H. Brans,
Exotic smoothness and physics: differential topology and spacetime models
World Scientific Publishing (2004)

[13] E. Minguzzi, M. Sanchéz,
The causal hierarchy of spacetimes
arXiv:gr-qc/0609119v3 (2008)

[14] O. Parrikar, S. Surya,
Causal Topology in Future and Past Distinguishing Spacetimes
arXiv:gr-qc/1102.0936v1 (2011)