Analysing topological aspects of QFT within the locally covariant algebraic framework

Michael Kiss

PhD

University of York
Mathematics

October 2020

Abstract

The locally covariant approach to quantum field theory (LCQFT) is a manifestly covariant functorial approach to quantum field theory (QFT) that applies to curved spacetimes and which builds on the local algebraic approach. In this thesis we investigate applications of LCQFT to topological aspects of QFT.

We analyse extensions of quantum field theories defined on contractible globally hyperbolic regions of spacetime, using Fredenhagen's *universal algebra* construction. This construction involves covering a spacetime by open contractible causally convex subregions, and applying the functor that defines the theory to each of them to get a net of local algebras. The universal algebra is then obtained by taking the colimit of this net. Morphisms between universal algebras can be defined with the result that the mapping between spacetimes and their corresponding universal algebras defines a functor. We prove two main results about this universal construction, which both require considerable geometric apparatus.

First we prove that for a broad class of theories modelled on the free scalar/Dirac field, the functor assigning universal algebras satisfies the Einstein causality axiom. We then restrict attention to Fermionic theories in this class, and analyse the universal theories obtained from the subtheories that assign even subalgebras. We show that for each spacetime $\mathcal{M}$, the universal theory assigns an algebra which decomposes into a product (in the categorical sense) of subalgebras, that are in bijective correspondence with the set $H^1(\mathcal{M}, \mathbb{Z}_2)$. The latter set counts the number of distinct spin structures the spacetime manifold $\mathcal{M}$ permits. The universal algebra for a Fermionic theory therefore has the geometric information encoded in it necessary to define half integer spin fields.

Contents

List of Figures

Acknowledgements

I'd first like to thank my supervisors Prof. Chris Fewster and Dr. Kasia Rejzner, for their exceptional guidance, patience and support throughout my PhD. I have learnt an incredible amount from them, including a deep appreciation for mathematical rigor which has fundamentally changed the way I view physics. With their insightful comments and feedback, I have become much clearer in the exposition of my ideas, and have picked up many tools useful for research.

I'd like to thank the other fellow PhD students at York for creating a fun and welcoming atmosphere. In particular I'd like to thank Allan, Brennen, Callum, Christoph, Francis, Matt, Nicola, Pete, Sky, Tom C, Tom QG and Vicky. My housemates Paul and Killian, for all the great times we had and for many interesting conversations. And of course the waterfowl on York campus, whose morning antics provided me with much peace of mind during my time at York.

Many thanks go to thank my wonderful girlfriend Lorna, who has been incredibly supportive and a great source of joy for me. She is incredibly kind and compassionate, and I am incredibly lucky to have her in my life. My greatest gratitude goes to my Mum, Neale, and my brother, who have been always there when I needed them, and whom I know I can always rely on.

Finally I'd like to dedicate this thesis to the memory of my late father Josef Kiss, who inspired me at an early age to pursue physics. I would not be where I am today without him, his passion for physics and the deep questions of the universe were undoubtedly a huge positive force in my life, for which I am eternally grateful.

Author's declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. Chapters 4 and 5 contain work that has been done partly in collaboration with Professor Chris Fewster and Dr Kasia Rejzner. The proofs of lemma 3.2.4 and lemma 3.3.3, which are minor technical results, were supplied by Chris Fewster. All sources are acknowledged as references.

Introduction

There are currently two main pillars of modern physics: general relativity, which describes the universe on large scales and quantum field theory (abbreviated QFT from now on), which describes the universe on small scales. These theories have been remarkably accurate in their domains of application, however they are incompatible with each other. There have been many attempts to find a theory of quantum gravity which can accommodate the predictions of QFT and general relativity in their respective regimes. One of the main issues is that the framework of QFT requires a causal structure, however causal structure is given by the spacetime metric, which would be quantised in a quantum theory of gravity. This means that for quantum states which are not eigenstates of the metric operator, there isn't a well defined causal structure.

Given this fact, it seems reasonable as a first approximation to a theory of quantum gravity, to consider the behaviour of quantum fields propagating on a fixed background spacetime. This allows for a concrete notion of causality which is an important ingredient of QFT. In this approach (often referred to as QFT on curved spacetimes) the gravitational back-reaction of the quantum fields is neglected, since it will generally have a negligible effect in most regimes and adds unnecessary complexity. Two of the main achievements of QFT in curved spacetimes, are the predictions of Hawking radiation [Haw74] and the Unruh effect [Unr76], which yield insights into features of a full theory of quantum gravity, for instance black hole entropy.

The usual framework for QFT is highly dependent on Poincaré covariance which is unique to flat spacetime, therefore certain principles have to be discarded in order to generalise to curved spacetime. For instance, the existence of a minimal energy state known as the vacuum state will in general not be possible, since general curved spacetimes will not have global time translation invariance. This means that in general, the notions of particles and S-matrices which the standard formulation of QFT relies on are ill-defined.

We will be analysing QFT on curved spacetimes using a mathematically rigorous framework which builds upon the local algebraic approach to QFT in flat spacetimes (abbreviated as AQFT) due to Haag and Kastler [HK64,Haa96]. In AQFT, the focus is shifted from the Hilbert space to the algebra of operators that model observable quantities. For linear theories these algebras will be canonical (anti-)commutation (abbreviated CCR/CAR) algebras, which are obtained by canonical quantisation of a classical theory, which consists of a symplectic vector space. The symplectic vector space models the phase space of the theory and the symplectic form models the dynamics. When the symplectic vector space is finite dimensional, there is a theorem due to Stone and Von Neumann which shows that there is a unique (up to unitary equivalence) irreducible Hilbert space representation of the corresponding CCR algebra. For infinite dimensional symplectic vector spaces, which are used to describe fields, there is no such result and the algebra has inequivalent irreducible Hilbert space representations. In order to capture the full information of the theory, it is then necessary to take the algebra of observables as the primary mathematical structure of QFT, rather than a particular Hilbert space representation. This will be discussed further in chapter 2.

The Locally Covariant approach to QFT (abbreviated as LCQFT), originally developed in [BFV03], is a mathematical framework for describing QFT, which builds on AQFT. The LCQFT framework is a more general framework that can be applied to to curved spacetimes and has the advantage of placing all spacetimes on the same footing. In LCQFT the principles of locality and covariance are manifest, since they are embedded in the building blocks of the theory, which are local algebras. In LCQFT, a theory is specified by a choice of functor that assigns an algebra to each object of a suitable background category (more on categories and functors in the next chapter). Usually this background structure consists of a spacetime equipped with some additional bundle structure. For two background objects related by a morphism, the theory gives a morphism between their associated algebras.

This thesis aims to investigate applications of LCQFT to topological aspects of QFT. We consider the difficulty of defining QFTs on spacetimes which have non-trivial global topology. Examples of this include Yang-Mills theories and theories with fields that transform in a half-integer rep of the Lorentz group. It is therefore useful to consider how a theory is defined on a topologically simple class of spacetimes, which intuitively correspond to small regions of spacetime that are topologically trivial, and to then use this information to extend the theory to a broader class of spacetimes. There is a known method of extending theories defined on a subclass of spacetimes called universal algebra construction, which was first introduced by Fredenhagen [Fre90] and was later realised to be a categorical construction known as a left-Kan extension [Lan12]. This method of extending theories is in some sense the simplest possible method[1] of extending a theory in a way that is consistent with how the theory acts on the subclass of spacetimes that it is defined on. This method of extension has the advantage of being relatively simple to describe, but notoriously difficult to do explicit calculations with. Even the question of non-triviality of an extended theory when applied to certain spacetimes is difficult to address; this problem was discussed in [Lan12] for instance.

We will examine this method of extension on a class of linear theories defined on contractible spacetimes, which consist of Bosonic (commutation relations) and Fermionic (anti-commutation relations) theories that can be considered generalisations of the theory of the complex scalar field and the theory of the Dirac field respectively. We first investigate whether the extended theories obtained by these methods satisfy the property of Einstein causality (defined in detail in chapter 2), which implements the principle of causality by asserting that quantum operators localised in spacelike separated regions of spacetime commute with each other. Fermionic theories satisfy a graded form of Einstein causality; field operators localised in spacelike separated regions of spacetime instead anti-commute with each other. We therefore restrict to the even subtheories of Fermionic theories, whose algebras are generated by pairs of Fermionic fields, since these theories do satisfy Einstein

[1] It is an example of a universal construction in category theory. More on this in chapter 1.

causality. There is a comparable construction in the Bosonic case; in both cases the assigned algebras are invariant under a $\mathbb{Z}_2$ transformation that flips the signs of the fields.

We find that the linear Bosonic and Fermionic theories, and their even subtheories, have extensions which do indeed satisfy Einstein causality. In order to prove this result, we use results in differential topology to develop geometrical tools which have applications to other local-to-global constructions in QFT. We then further investigate the extensions of even Fermionic theories, and find that the extended theories assign an algebra to each spacetime $\mathcal{M}$ that encodes information about the set of all spin structures that the spacetime admits. Each possible choice of spin structure corresponds to a choice of cohomology class in $H^1(\mathcal{M}, \mathbb{Z}_2)$, and the latter assigns a $\mathbb{Z}_2$ value to each loop in $\mathcal{M}$. We find that the universal even Fermionic theory assigns an algebra to each spacetime $\mathcal{M}$ that decomposes into a product of subalgebras, with each subalgebra corresponding to a cohomology class in $H^1(\mathcal{M}, \mathbb{Z}_2)$. This shows that global topological information is encoded in the structure of local algebras and how they relate to each other, therefore indicating that knowledge of local physics, which is the only type of physics that we can reliably probe, is enough to infer global information about the topology of spacetime.

Thesis layout

In the first chapter we will review some of the mathematical concepts that will be used in the rest of the thesis. In particular we will cover topology, differential topology, Lorentzian geometry and category theory. In the second chapter we will outline the LCQFT framework that we will be using in this thesis, together with some of the approaches to axiomatising QFT that the LCQFT framework builds on. We finish the second chapter with two example theories in the LCQFT framework, which form the basis of a larger class of theories that we will study in chapters 3 and 4. In chapter 3 we introduce this larger class of theories, together with the universal algebra construction which we use to extend theories defined on contractible regions of spacetime. We develop geometrical tools which are used to show that the extended theories obtained by these methods satisfy Einstein causality, which is a key property for theories in the LCQFT framework. These geometric tools also have wider applications, and are used in the subsequent chapter. In chapter 4 we restrict our attention to the even subtheories of Fermionic theories that are part of the general class of theories introduced in the previous chapter. We show that when we extend these even Fermionic theories from contractible regions of spacetime, the resulting theory encodes topological information in the algebra that it assigns to a given spacetime. In particular, information about the different possible choices of spin structure (defined in chapter 2 when defining the theory of the Dirac field) is encoded in the algebra by means of a decomposition into subalgebras. We finish with a chapter on conclusions that can be drawn from the results of the thesis, and possible avenues of further research.

Chapter 1

Mathematical preliminaries

In this chapter we will outline the mathematical techniques that will be used in the rest of the thesis. The topics covered in this section are: topology, differential topology, Lorentzian geometry and category theory.

1.1 Topology

A topology on a set is an additional structure that specifies which subsets are *open*, and roughly speaking two points can be considered "nearby" if they are both contained in many common open sets. This notion can then be used to give rigorous definitions to concepts such as convergence, continuity and connectedness. We now give the definition of a topology on a set [Kos80, Definitions 2.1 and 2.4].

Definition 1.1.1. *A topology τ on a set X is a collection of subsets of X that satisfy the following properties:*

- *Both the empty set $\emptyset$ and X are elements of τ.*
- *Any union of elements of τ is an element of τ.*
- *Any intersection of finitely many elements of τ is an element of τ.*

The pair (X, τ) is referred to as a topological space, although we will often just refer to X as the topological space and leave τ implicit. An element of τ is referred to as an open subset of X, and the complement of an open set U, denoted as $X \backslash U$, is referred to as a closed set.

An important object of study in topology is an open cover of a topological space [Kos80, Definitions 7.1, 7.2 and 7.3].

Definition 1.1.2. *An open cover of a subset A of a topological space X is a collection $\{U_n \,|\, n \in I\}$ of open sets U_n indexed by a set I such that $\cup_{n \in I} U_n \supset A$. If $I' \subset I$, then $\{U_n \,|\, n \in I'\}$ is called a subcover of $\{U_n \,|\, n \in I\}$, and is called a finite open subcover if I' is also finite.*

Various intuitive properties of a space can be defined as properties of its underlying topology, as we shall see in the following definitions ([Kos80, Definition 9.1] and [Kos80, Definition 7.4] respectively).

Definition 1.1.3. *A topological space X is connected if the only subsets of X that are both open and closed are X and $\emptyset$.*

Definition 1.1.4. *A subset A of a topological space X is compact if any open cover of A has a finite open subcover.*

A subset A of a topological space X can be equipped with a natural topology, which we will now introduce [Kos80, Definition 4.1].

Definition 1.1.5. *The subset topology of $A \subset X$ has as its open sets the intersections of A with open sets of X.*

Various properties of a topological space are framed in terms of continuous maps on them, so we now define what it means for a map to be continuous [Kos80, Definition 3.1].

Definition 1.1.6. *A mapping f between topological spaces X and Y is continuous if for any open subset U of Y, the set of elements of X that f maps to an element of U, which we denote as $f^{-1}(U)$, is an open subset of X.*

This notion of continuity coincides with the notion used in real analysis when the real numbers $\mathbb{R}$ are equipped with the standard topology. We now introduce a notion of equivalence for topological spaces [Kos80, Definition 3.6].

Definition 1.1.7. *Two topological spaces X and Y are homeomorphic, if there exists a continuous map $f : X \to Y$ which is bijective and has a continuous inverse. The map f is referred to as a homeomorphism.*

If two spaces are homeomorphic, then there is a bijection between the families of subsets that form their topologies. This means that all the topological properties of a topological space are preserved by homeomorphisms. Homeomorphisms are therefore the most natural form of equivalence between topological spaces. We now introduce a weaker notion of equivalence which will be useful for studying topological spaces [Kos80, Definitions 13.2 and 13.5].

Definition 1.1.8. *Two continuous maps $f, g : X \to Y$ are homotopic if there exists a continuous map $H : [0,1] \times X \to Y$ such that $H(0,\cdot) = f(\cdot)$ and $H(1,\cdot) = g(\cdot)$.*

Two topological spaces X and Y are homotopic, if there exist continuous maps $f : X \to Y$ and $g : Y \to X$ such that $f \circ g$ is homotopic to $\mathbb{1}_Y$ and $g \circ f$ is homotopic to $\mathbb{1}_X$.

With this notion of equivalence, the following groups can be constructed [Kos80, pp. 133].

Definition 1.1.9. *For topological space X and a base point $p \in X$, let $C^0([0,1]^n, X, p)$ denote the set of continuous maps $f : [0,1]^n \to X$ which map the boundary of $[0,1]^n$ to p.*

The set $\pi_n(X,p)$ is defined to be the set of equivalence classes of maps in $C^0([0,1]^n, X, p)$, where two maps are equivalent if there exists a homotopy H between them such that $H(t,\cdot) \in C^0([0,1]^n, X, p)$ for all $t \in [0,1]$. For $n \geq 1$, the set $\pi_n(X,p)$ can be equipped with a group structure given by

$$[f] \circ [g] = [f * g]$$

where

$$(f * g)(t_1, \ldots, t_n) = \begin{cases} f(2t_1, t_2, \ldots, t_n) & t_1 \in [0, 1/2] \\ g(2t_1 - 1, t_2, \ldots, t_n) & t_1 \in [1/2, 1] \end{cases} \quad . \tag{1.1}$$

The group $\pi_1(X, p)$ is commonly referred to as the fundamental group of X, and the base-point is usually omitted in the notation since the groups resulting from different choices of base-point are isomorphic if there is a path connecting the base-points [Kos80, Theorem 15.4] (although there is no canonical choice of isomorphism).

These groups can be used to classify spaces, since two homotopic spaces have isomorphic homotopy groups [Kos80, Theorem 15.13], although the converse does not necessarily hold. Various properties of a topological space can also be framed in terms of conditions on its associated homotopy groups. Another useful set of groups for exploring the properties of a topological space is obtained by a choice of homology theory [Kos80, §29]. We will only be concerned with singular homology in this thesis.

Definition 1.1.10. *Let X be a topological space. An n-simplex is an ordered tuple $(v_1, \ldots, v_n)$ of vectors in $\mathbb{R}^n$, which are called vertices of the simplex, with an associated topological space $\sigma \subset \mathbb{R}^n$ which is the convex hull of the vertices. For each $n \in \mathbb{N}$, let $C_n(X)$ denote the free group generated by continuous maps $f : \sigma \to X$ where σ is the topological space associated to an n-simplex. For each $n \in \mathbb{N}$, we define the map $\partial_n : C_n(X) \to C_{n-1}(X)$ by its action on generators $f \in C_n(X)$ as follows*

$$\partial_n f = \sum_{i=1}^{n} (-1)^i f|_{\sigma_i} \quad ,$$

where σ_i is the topological space associated to the simplex obtained by removing the i^{th} vertex from the n-simplex associated to σ. The collection of maps $\{\partial_n | n \in \mathbb{N}\}$ have the property that $\text{imag}(\partial_{n+1}) \subset \ker(\partial_n)$.

For each $n \in \mathbb{N}$, we define the n^{th} homology group $H_n(X)$ of X to be the quotient group

$$\ker(\partial_n)/\text{imag}(\partial_{n+1})$$

where $\text{imag}(\partial_{n+1})$ *and* $\ker(\partial_n)$ *are subgroups of* $C_n(X)$. *This quotient group is well defined since* $\text{imag}(\partial_{n+1})$ *is a normal subgroup of* $\ker(\partial_n)$.

The singular homology groups of two topological spaces are isomorphic if the spaces are homotopic [Rot88, Corollary 4.24]. Given the singular homology of a topological space X, we can also define its singular cohomology with coefficients in an abelian group A, which yields another set of groups that can be used to classify the properties of X.

Definition 1.1.11. *For an abelian group A, the singular cohomology of X with coefficients in A is defined by the cochain complex formed by groups $C^n(X; A) = \text{Hom}(C_n(X), A)$ together with coboundary maps δ_n defined by their action on generators $h \in C^n(X; A)$ as follows*

$$(\delta_n h)(f) = h(\partial_n f) \quad .$$

From this cochain complex we can form the cohomology groups of X with coefficients in A

as follows

$$H^n(X;A) = \ker(\delta_n)/\mathrm{imag}(\delta_{n-1}) \quad .$$

Just as was the case for singular homology, the singular cohomology groups of two topological spaces are isomorphic if the spaces are homotopic [Rot88, Theorem 12.4]. There are various relations between the homotopy, homology, and cohomology groups of a topological space X, but in order to describe the relations we first need to introduce the following definition [Rot88, pp. 383].

Definition 1.1.12. *Let A and B be Abelian groups, and $0 \to R \xrightarrow{i} F \to A \to 0$ be an exact sequence where R and F are free Abelian groups*[1]*. This induces an exact sequence $0 \to \mathrm{Hom}(A,B) \to \mathrm{Hom}(F,B) \xrightarrow{i^*} \mathrm{Hom}(R,B)$, and we define $\mathrm{Ext}(A,B)$ to be the cokernel*[2] *of i^* which is independent of the choice of the exact sequence (see comments below the definition in [Rot88]).*

We now state some results on the relations between homotopy, homology, and cohomology groups. Our interest will be in counting the elements of various of these groups, hence the precise nature of the isomorphisms stated below will not be needed.

Theorem 1.1.13. *For any topological space X and abelian group A, one has the following isomorphism of groups: $H^n(X;A) \cong \mathrm{Hom}\left(H_n(X), A\right) \oplus \mathrm{Ext}\left(H_{n-1}(X), A\right)$.*

Proof. See [Rot88, Theorem 12.11]. □

Theorem 1.1.14. *For each path-connected topological space X, there exists a set of homomorphisms $h_n : \pi_n(X,p) \to H_n(X)$ known as the Hurewicz maps, such that h_1 restricted to the Abelianization*[3] *of $\pi_1(X,p)$ is an isomorphism, and h_n is an isomorphism for $n \geq 2$ if the first $n-1$ homotopy groups are trivial.*

Proof. See [Spa82, pp. 387-400]. □

Corollary 1.1.15. *For each path-connected topological space X, $\mathrm{Hom}(\pi_1(X,p), \mathbb{Z}_2) \cong H^1(X;\mathbb{Z}_2)$.*

Proof. Since $\mathbb{Z}_2$ is Abelian, the group $\mathrm{Hom}(\pi_1(X,p), \mathbb{Z}_2)$ is isomorphic to $\mathrm{Hom}(\pi_1(X,p)_{\mathrm{ab}}, \mathbb{Z}_2)$ where $\pi_1(X,p)_{\mathrm{ab}}$ is the Abelianization of $\pi_1(X,p)$. Combining this with theorem 1.1.14, we find $\mathrm{Hom}(\pi_1(X,p), \mathbb{Z}_2) \cong \mathrm{Hom}\left(H_1(X), \mathbb{Z}_2\right)$. As $H_0(X) \cong \mathbb{Z}$ [Rot88, Theorem 4.14], it follows that $\mathrm{Ext}(H_0(X), G)$ is trivial for all Abelian G [Rot88, pp. 384]. Theorem 1.1.13 therefore yields

$$H^1(X;\mathbb{Z}_2) \cong \mathrm{Hom}(H_1(X), \mathbb{Z}_2) \oplus \mathrm{Ext}(H_0(X), \mathbb{Z}_2) \cong \mathrm{Hom}(\pi_1(X,p), \mathbb{Z}_2) \quad .$$

□

[1] A free Abelian group is an Abelian group with a basis i.e, a subset such that every element of the group can be **uniquely** expressed as a linear combination of elements of the subset with integer coefficients.

[2] The cokernel of a map is the quotient of the target space by the image of the map.

[3] Obtained by quotienting out the subgroup generated by elements of the form $g^{-1}h^{-1}gh$.

We now switch our focus to topological spaces defined in terms of pairs of topological spaces. For instance, the Cartesian product of sets can be generalised to topological spaces by the following definition [Kos80, Definition 6.1].

Definition 1.1.16. *Given topological spaces X and Y, the Cartesian product $X \times Y$ comes with a natural topology: the coarsest topology (in the sense of being the smallest subset of the powerset of $X \times Y$) such that the projection maps $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ are continuous.*

The notion of product spaces is generalised by the definition of coordinate bundles [Ste51, Definition 1.2.3], which we now give.

Definition 1.1.17. *A coordinate bundle consists of a tuple $(B, \pi, b, F, G, \{U_i, \psi_i\})$ where B, b and F are topological spaces referred to as the bundle space, the base space, and the fibre space F respectively. The map π is a projection $\pi : B \to b$ from the bundle space to the base space. The structure group G is a topological group which is an effective topological transformation group of F meaning; there is a homeomorphism T_g for each $g \in G$ such that $\forall f \in F$ $T_{\mathrm{Id}} \circ f = f$, $(T_{g_1 g_2}) \circ f = T_{g_1} \circ (T_{g_2} \circ f)$ and $T_g \circ f = f \ \forall f \in F \ \Rightarrow \ g = \mathrm{Id}$.*

The set $\{U_i, \psi_i\}$ which we call local trivialisations, consists of open regions U_i that cover b which we call coordinate regions, and homeomorphisms $\psi_i : \pi^{-1}(U_i) \to U_i \times F$ which "preserve fibres" in the sense that the following diagram commutes

$$\begin{tikzcd} \pi^{-1}(U_i) \subset B \arrow[rr, leftrightarrow, "\psi_i"] \arrow[dr, "\pi"'] & & U_i \times F \arrow[dl, "pr_1"] \\ & U_i \subset b & \end{tikzcd}.$$

For two overlapping coordinate regions U_i and U_j, there is a transition function $t_{ij} := \psi_j^{-1}\psi_i : U_i \cap U_j \times F \to U_i \cap U_j \times F$. These transition functions satisfy the cocycle condition $t_{ij} \circ t_{jk} = t_{ik}$ and each transition function factors through $G \hookrightarrow \mathrm{Aut}(F)$.

From this we can define fibre bundles [Ste51, Definition 1.2.4].

Definition 1.1.18. *Two coordinate bundles $(B, \pi, b, F, G, \{U_i, \psi_i\})$ and $(B, \pi, b, F, G, \{V_j, \phi_j\})$ are considered equivalent if all the homeomorphisms $\phi_j^{-1}\psi_i$ correspond to the group action of the structure group G on the fibres. A Fibre bundle is defined as an equivalence class of coordinate bundles with respect to the previously defined equivalence relation, and we simply denote it by (B, π, b, F, G).*

A section of a fibre bundle is a continuous map $f : b \to B$ such that $(f \circ \pi)(p) = p$ for all $p \in b$.

This definition has the advantage of being coordinate independent. As can be seen from definition 1.1.17, the bundle space of a fibre bundle locally looks like a product space. We now introduce a couple of types of fibre bundle that will be used in the thesis.

Definition 1.1.19. *A fibre bundle where the fibre space is a vector space and the trivialisations are fibrewise linear, is called a vector bundle.*

A fibre bundle where the fibre space is a group G, which is additionally equipped with a continuous right action $R : B \times G \to B$ such that for any $p_1, p_2 \in B$, there is precisely one $g \in G$ such that $p_1 = R(p_2, g)$, is referred to as a principal bundle, and we simply denote it as (B, π, b, G).

Each fibre bundle has an associated principle bundle, which can be obtained from knowledge of its transition functions [Ste51, Theorem 1.3.3].

1.2 Differential topology

Differential topology is a branch of topology concerned with a topological spaces equipped with additional structure. We begin by defining this additional structure, together with some additional conditions on topological spaces which we will then use to define the central objects of study in differential topology. Most of the definitions in this section can be found in chapter 1 of [Hir76].

Definition 1.2.1. *An n dimensional chart on a topological space X is a pair (U, ψ) where $U \subset X$ is open and ψ is a homeomorphism $\psi : U \to \mathbb{R}^n$. Two charts (U_i, ψ_i) and (U_j, ψ_j) are said to be C^r compatible if $U_i \cap U_j = \emptyset$ or the map $\psi_j \psi_i^{-1} : \psi_i(U_i \cap U_j) \subset \mathbb{R}^n \to \psi_j(U_i \cap U_j) \subset \mathbb{R}^n$ is a homeomorphism which is r times continuously differentiable and whose inverse is also r times continuously differentiable.*

A C^r n dimensional atlas on X is a collection of C^r compatible n dimensional charts $\{(U_i, \psi_i)\}$ such that the coordinate regions $\{U_i\}$ cover X. Two atlases are C^r compatible if every chart of one atlas is C^r compatible with every chart of the other atlas. A maximal C^r atlas is an atlas that contains every atlas that is C^r compatible with it, and the existence of a maximal atlas is guaranteed by Zorn's lemma[4].

Definition 1.2.2. *A topological space X is Hausdorff if any two distinct points of X can be contained in disjoint open regions, and paracompact if every open cover of X has a refinement[5] which only has finitely many open sets that intersect any given point of X.*

Definition 1.2.3. *An n dimensional manifold is a Hausdorff paracompact topological space X equipped with a maximal C^∞ n dimensional atlas.*

Differential topology is concerned with the global properties of manifolds, and uses tools which leverage the known local structure of manifolds. There is an important bundle associated to manifolds which we now define.

Definition 1.2.4. *The tangent space TX of a manifold X is defined to be the set of equivalence classes $[p, i, v]$ where $p \in X$, i belongs to the index set listing the charts (U_i, ψ_i) with $p \in U_i$ of the atlas of X, $v \in \mathbb{R}^n$, and equivalence is generated by the relation $(p, i, v) \sim (q, j, u)$ if $p = q$ and*

$$\left.\frac{\partial(\psi_j \circ \psi_i^{-1})}{\partial x^\mu}\right|_{\psi_i(p)} v^\mu = u$$

where x^μ are coordinates in the chart ψ_i, v^μ is the x^μ component of v, and we have used the convention that a repeated Greek index is summed over. We refer to p as the base

[4] Zorn's lemma requires the axiom of choice.

[5] The cover $\{U_i \mid i \in I\}$ is a refinement of $\{V_j \mid j \in J\}$ if $\forall i \in I$, $\exists j \in J$ such that $U_i \subset V_j$.

point of an element $[p,i,v] \in TX$, and the set of all elements of TX with base point p is denoted by T_pX. The elements of T_pX are referred to as tangent vectors at p. For a subset $U \subset X$, we define $TU := \bigsqcup_{p \in U_i} T_pX$.

The topology of TX is generated by requiring that for each chart (U_i, ψ_i) of the atlas of X, the map $\widetilde{\psi}_i : TU_i \to \mathbb{R}^n \times \mathbb{R}^n$ given by

$$\widetilde{\psi}_i([p,i,v]) = (\psi_i(p), v)$$

is continuous. The charts $(TU_i, \widetilde{\psi}_i)$ form a smooth atlas of TX, thus giving it the structure of a smooth manifold.

We now define an important class of maps between manifolds.

Definition 1.2.5. *Let X be an n dimensional manifold and Y a m dimensional manifold. A C^r map is a map $f : X \to Y$ such that for any chart (U, ψ) of X and any chart (V, ϕ) of Y such that $f(U) \subset V$, the map $\phi \circ f \circ \psi^{-1} : \psi(U) \subset \mathbb{R}^n \to \mathbb{R}^m$ is r times differentiable. We use the notation $C^r(X,Y)$ to denote the set of C^r maps from X to Y.*

A C^r diffeomorphism is a bijective C^r map with C^r inverse.

Unless stated otherwise, we will use the term diffeomorphism to mean a C^∞ diffeomorphism. Given a C^r map with $r \geq 1$, we can define an induced map between the corresponding tangent bundles.

Definition 1.2.6. *For a C^r map $f : X \to Y$ with $r \geq 1$, we define $df : TX \to TY$ by its action on a generic element $(p, [i,v]) \in TX$ by*

$$df(p,[i,v]) = \left(f(p), \left[j, \left. \frac{\partial(\phi_j \circ f \circ \psi_i^{-1})}{\partial x^\mu} \right|_{\psi_i(p)} v^\mu \right] \right)$$

where x^μ are coordinates in the chart ψ_i, and v^μ is the x^μ component of v. This does not depend on the choice of i and j, as can be seen by a simple application of the chain rule, hence the map is well defined.

With tangent maps defined, we now introduce two important subsets of $C^r(X,Y)$.

Definition 1.2.7. *A C^r map $f : X \to Y$ with $r \geq 1$ is immersive at $p \in X$ if $df_p = df|_{T_pX} : T_pX \to T_{f(p)}Y$ is injective. The map f is an immersion if it is immersive at all $p \in X$, and it is an embedding if it is an immersion which is a homeomorphism onto its image. We use $\mathrm{Imm}^r(X,Y)$ and $\mathrm{Emb}^r(X,Y)$ to denote the set of C^r immersions and embeddings respectively.*

Throughout this thesis we will be dealing with maps in $C^r(X,Y)$, and approximating them by maps which have additional properties. In order to make this notion of approximating maps precise, we must introduce a topology on the set $C^r(X,Y)$. We take the following definition from [Hir76, §2.1].

Definition 1.2.8. *Consider $f \in C^r(X,Y)$ together with charts (U,ψ) and (V,ϕ) of X and Y respectively. For compact $K \subset U$ and $\epsilon > 0$, the set $N(f; (U,\psi), (V,\phi), K, \epsilon)$ denotes the subset of $C^r(X,Y)$ consisting of maps $\widetilde{f}$ such that $\widetilde{f}(U) \subset V$ and*

$$\|D^n(\phi \circ \widetilde{f} \circ \psi^{-1})(p) - D^n(\phi \circ f \circ \psi^{-1})(p)\| < \epsilon \quad \forall p \in \psi(K) \quad \forall n \leq r$$

where D^n denotes the n^{th} order derivative, and $\|\cdot\|$ denotes the usual Euclidean norm. The weak topology on $C^r(X,Y)$ is generated by the $N(f;(U,\psi),(V,\phi),K,\epsilon)$ sets, meaning any union or finite intersection of these sets defines an open subset of the topology.

There is another topology that can be defined on $C^r(X,Y)$ called the strong topology, however we will be using approximation results in the case that X is compact (usually a circle or closed line interval), and in this case the strong and weak topologies coincide (see remarks after the definition of the strong topology in [Hir76, §2]).

We will be using the fact that various properties of functions are generic, which means that a general function can be perturbed by an arbitrarily small amount to a function which satisfies one or many of these properties. This concept is formalised by saying that the subset of functions satisfying these properties is dense, where dense subsets are defined as follows.

Definition 1.2.9. *A subset A of a topological space X is dense if A has non-empty intersection with every non-empty open subset of X.*

We now introduce a couple more subsets of $C^r(X,Y)$ that have key properties which we will make use of throughout the thesis (see definitions in [Hir76, §2.2] and [Kos93, Definition 4.1.1] respectively).

Definition 1.2.10. *Given a submanifold A of Y, the set $\pitchfork^r_L(X,Y;A)$ consists of C^r maps $f: X \to Y$ that are transverse to A along a subset $L \subset X$ i.e, if $x \in L$ and $f(x) = y \in A$ then we have $df(T_xX) + T_yA = T_yY$. When $L = X$ we omit the subscript on $\pitchfork^r_L(X,Y;A)$.*

Definition 1.2.11. *For $f, g \in C^\infty(X,Y)$, f is transverse to g if $f(p) = g(q)$ implies $df_p(T_pX) + dg_q(T_qX) = T_{f(p)}Y$. If $f(p) = f(q)$ and $p \neq q$, we refer to the pair (p,q) as double points of f. For $L \subset X$, let $ST^\infty_L(X,Y)$ (when $L = X$ we omit the subscript) denote the set of functions h such that for $p, q \in L$ with $p \neq q$ and $h(p) = h(q)$, $dh_p(T_pX) + dh_q(T_qX) = T_{h(p)}Y$. We say that the functions in $ST^\infty_L(X,Y)$ are self-transverse on L.*

Functions in $\pitchfork^r_L(X,Y;A)$ have the following nice property.

Lemma 1.2.12. *$f \in \pitchfork^r_L(X,Y;A)$ implies $f^{-1}(A) \cap L$ is a submanifold of X and* $\dim(X) - \dim\big(f^{-1}(A) \cap L\big) = \dim(Y) - \dim(A)$.

Proof. See [Hir76, Theorem 1.3.3]. □

We now state a definition which will allow us to then state a folklore theorem.

Definition 1.2.13. *Let X and Y be manifolds and $L \subset X$. For $f \in C^r(X,Y)$, we use $C^r(X,Y)_{f|_L}$ to denote the subset of $C^r(X,Y)$ such that for each $g \in C^r(X,Y)_{f|_L}$, $g|_L = f|_L$.*

We also use $\mathrm{Imm}^r(X,Y)_{f|_L}$ and $\mathrm{Emb}^r(X,Y)_{f|_L}$ to denote $\mathrm{Imm}^r(X,Y) \cap C^r(X,Y)_{f|_L}$ and $\mathrm{Emb}^r(X,Y) \cap C^r(X,Y)_{f|_L}$ respectively.

Folklore theorem 1.2.14. *Let X be a compact manifold with boundary ∂X, Y be a manifold with boundary ∂Y such that $\dim(Y) \geq 2\dim(X)$, L be a closed subset of X that contains ∂X, and $f \in C^0(X,Y)$ a function such that*

- *f is smooth and immersive at all points of L.*
- *$f \in ST^{\infty}_{L}(X, Y)$.*
- *$f|_{\partial X} \in ST^{\infty}(\partial X, Y)$.*
- *$f \in \pitchfork^{\infty}_{L \setminus \partial X}(X, Y, f(\partial X))$.*
- *$f(\partial X) \subset \partial Y$ if $\partial Y \neq \emptyset$.*

Then the set of $g \in \mathrm{Imm}^{\infty}(X, Y)_{f|_L}$ such that $g|_{\mathrm{int}(X)}$ is self-transverse and transverse to $g|_{\partial X}$, is dense in $C^0(X, Y)_{f|_L}$.

We were not able to obtain a proof for this theorem, however the theorem can be split into three parts which are folk theorems within the differential topology literature. The first part is that $C^{\infty}(X, Y)_{f|_L}$ is dense in $C^0(X, Y)_{f|_L}$, and this is stated to be true in [Hir76, Exercise 2.2.4].

The second part is that $\mathrm{Imm}^{\infty}(X, Y)_{f|_L}$ is dense in $C^0(X, Y)_{f|_L}$, and can be obtained by the relative jet transversality result [Vok, Theorem 9.14] combined with the proof of density of immersions using jet transversality in [Hir76, pp. 82]. The use of this result requires the $\dim(Y) \geq 2\dim(X)$ condition. The condition that $f(\partial X) \subset \partial Y$ if $\partial Y \neq \emptyset$ is required so that $f|_{X \setminus L}$ maps into the interior of Y, which allows for the use of jet transversality results for $f|_{X \setminus L}$ since these results don't necessary hold if the target and domain spaces both have boundary.

The third part is that the set of functions g such that $g|_{\mathrm{int}(X)}$ is self-transverse and transverse to $g|_{\partial X}$ is dense in $\mathrm{Imm}^{\infty}(X, Y)_{f|_L}$. This part has a sketch proof that can be extracted from the proof of [Vok, Theorem 9.19]. Although [Vok, Theorem 9.19] proves the stronger result that any neighbourhood of f in $C^{\infty}(X, Y)_{f|_L}$ contains an injective function, it also includes the stronger assumption that $\dim(Y) \geq 2\dim(X) + 1$. Self-transversality and injectivity are actually equivalent when $\dim(Y) \geq 2\dim(X) + 1$, hence the dimension assumption can be relaxed to $\dim(Y) \geq 2\dim(X)$ and the proof still holds but proves relative approximation to self-transverse functions rather than injective functions. The set of g such that $g|_{\mathrm{int}(X)}$ is transverse to the closure of $g|_{\partial X}$ is open and dense in $C^{\infty}(X, Y)$ by [Hir76, Theorem 3.2.1 (b)], we can therefore modify $g|_{\mathrm{int}(X)}$ to satisfy this condition (without modifying $g|_L$ since $f \in \pitchfork^{\infty}_{L \setminus \partial X}(X, Y, f(\partial X))$ by supposition), and further modification to $g|_{\mathrm{int}(X)}$ so that it is self-transverse will preserve this condition.

This folklore theorem is important since self-transverse functions have the following important property.

Proposition 1.2.15. *Let X be a compact manifold with boundary ∂X, Y be a manifold with boundary ∂Y, and $f : X \to Y$ be a smooth immersion such that $f|_{\mathrm{int}(X)}$ and $f|_{\partial X}$ are both self-transverse and transverse to each other. Then the set of double points of f is finite if $\dim(Y) = 2\dim(X)$ and empty if $\dim(Y) \geq 2\dim(X) + 1$.*

Proof. For a general topological space T, let $T^{(2)}$ denote $T \times T$ with the diagonal removed. Let $f \times f : X^{(2)} \to Y \times Y$ be given by $(f \times f)(p, q) = (f(p), f(q))$. We wouldl like to use lemma 1.2.12 to show that the double points form a submanifold, however if the boundaries of X and Y are non-empty then there is the issue that $X \times X$ and $Y \times Y$ may have corners, and therefore fail to be manifolds. We deal with this by splitting up the double points of f into parts

$$D(\widetilde{X^{(2)}}, \widetilde{Y}) := (f \times f)|^{-1}_{\widetilde{X^{(2)}}}(\Delta\widetilde{Y}) \subset X^{(2)}$$

where $\widetilde{X^{(2)}}$ denotes one of the following subsets of $X^{(2)}$: $\mathrm{int}(X) \times \mathrm{int}(X)$, $\mathrm{int}(X) \times \partial X$, $\partial X \times \mathrm{int}(X)$ or $\partial X \times \partial X$, and $\Delta\widetilde{Y}$ denotes the diagonal of $\widetilde{Y} \times \widetilde{Y}$ where $\widetilde{Y}$ is either $\mathrm{int}(Y)$ or ∂Y. By supposition $f|_{\mathrm{int}(X)}$ and $f|_{\partial X}$ are both self-transverse and transverse to each other, hence $f \times f$ restricted to $\widetilde{X^{(2)}}$ is transverse to $\Delta\widetilde{Y}$. We can therefore use lemma 1.2.12 to show that $D(\widetilde{X^{(2)}}, \widetilde{Y})$ is a submanifold of $X^{(2)}$ of dimension

$$\dim(\widetilde{X^{(2)}}) - \dim(\widetilde{Y}) = \begin{cases} 2\dim(X) - \dim(Y) & \text{if } \widetilde{X^{(2)}} = \mathrm{int}(X) \times \mathrm{int}(X) \text{ and } \widetilde{Y} = \mathrm{int}(Y) \\ 2\dim(X) - \dim(Y) - 1 & \text{if } \widetilde{X^{(2)}} = \mathrm{int}(X) \times \partial X \text{ and } \widetilde{Y} = \mathrm{int}(Y) \\ 2\dim(X) - \dim(Y) - 2 & \text{if } \widetilde{X^{(2)}} = \partial X \times \partial X \text{ and } \widetilde{Y} = \mathrm{int}(Y) \\ 2\dim(X) - \dim(Y) - 1 & \text{if } \widetilde{X^{(2)}} = \mathrm{int}(X) \times \mathrm{int}(X) \text{ and } \widetilde{Y} = \partial Y \\ 2\dim(X) - \dim(Y) - 2 & \text{if } \widetilde{X^{(2)}} = \mathrm{int}(X) \times \partial X \text{ and } \widetilde{Y} = \partial Y \\ 2\dim(X) - \dim(Y) - 3 & \text{if } \widetilde{X^{(2)}} = \partial X \times \partial X \text{ and } \widetilde{Y} = \partial Y \end{cases}.$$

The full set of double points is given by the union of the submanifolds $D(\widetilde{X^{(2)}}, \widetilde{Y})$ for each of the different cases above, and we see that for each case the submanifold has dimension ≤ 0 if $\dim(Y) = 2\dim(X)$ and dimension < 0 if $\dim(Y) \geq 2\dim(X) + 1$. By supposition X is compact, and since a 0 dimensional submanifold of a compact manifold has finitely many points, this implies the set of double points is finite if $\dim(Y) = 2\dim(X)$ and is empty if $\dim(Y) \geq 2\dim(X) + 1$.

□

From this we get a relative approximation to embedding result.

Corollary 1.2.16. *Let X, Y, L and f satisfy the conditions of theorem 1.2.14 and the additional condition that* $\dim(Y) \geq 2\dim(X) + 1$. *Then* $\mathrm{Emb}^\infty(X,Y)_{f|_L}$ *is dense in* $C^0(X,Y)_{f|_L}$.

Proof. Since f and L satisfy the conditions of theorem 1.2.14, any neighbourhood of f in $C^0(X,Y)_{f|_L}$ contains a function g that satisfies the conditions of proposition 1.2.15. This implies that the set of double points of g are empty since $\dim(Y) \geq 2\dim(X) + 1$, hence g is an injective immersion. An injective immersion on a compact domain space is an embedding, hence g is an embedding and therefore $\mathrm{Emb}^\infty(X,Y)_{f|_L}$ is dense in $C^0(X,Y)_{f|_L}$. □

1.3 Lorentzian geometry

In general relativity, the central object of study is spacetime which is modelled as a Lorentzian manifold which we now define. Most of the definitions in this section can be found in [Wal84, §8].

Definition 1.3.1. *A metric tensor g on a manifold $\mathcal{M}$ gives a map $g_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$ for each $p \in \mathcal{M}$ which is bilinear, symmetric, and non-degenerate. Moreover, given any* *pair of smooth vector fields* X *and* Y*, the map* $p \mapsto g_p(X(p), Y(p))$ *is smooth. The metric* g *and its inverse metric* g^{-1} *are given in local coordinates* $\{x^\mu\}$ *by* $g_{\mu\nu}dx^\mu dx^\nu$ *and* $g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}$ *respectively, such that the coefficients* $g_{\mu\nu}$ *and* $g^{\mu\nu}$ *satisfy* $g^{\mu\nu}g_{\nu\lambda} = \delta^\mu_\lambda$ *where* δ *is the Kronecker delta.*

The signature of g *is a pair* (p, q) *where* p *is the number of vectors in an psuedo-orthogonal basis of* $T_p\mathcal{M}$ *such that* $g(v, v) > 0$ *and* q *is the number such that* $g(v, v) < 0$*. These numbers are independent of the choice of the point* $p \in \mathcal{M}$*, and the choice of orthogonal basis for* $T_p\mathcal{M}$*.*

A Riemannian manifold consists of a manifold equipped with a metric tensor of signature $(n, 0)$ *called a Riemannian metric, and a Lorentzian manifold consists of a manifold equipped with a metric tensor of signature* $(1, n-1)$ *called a Lorentzian metric.*

It should be noted that not every manifold can be made Lorentzian, since the existence of a Lorentzian metric imposes topological restrictions on the manifold. If $\mathcal{M}$ is not compact then it will always admit a Lorentzian metric, however if $\mathcal{M}$ is compact then it admits a Lorentzian metric if and only if its Euler characteristic[6] also vanishes [O'N83, Proposition 5.37]. This means for instance the 4-sphere S^4 does not admit a Lorentzian metric, due to its non-vanishing Euler characteristic. Lorentzian manifolds have causal structure, which plays an important role in physics, and we introduce this structure in the following definitions.

Definition 1.3.2. *A timelike curve on a Lorentzian manifold* $(\mathcal{M}, g)$*, is a* C^1 *map* $\gamma : I \to \mathcal{M}$ *where* I *is a connected subset of* $\mathbb{R}$*, such that* $g(\dot\gamma(t), \dot\gamma(t)) > 0$ *for all* $t \in I$*. A causal curve on* $(\mathcal{M}, g)$ *is a* C^1 *map* $\gamma : I \to \mathcal{M}$ *such that* $\dot\gamma(t) \neq 0$ *and* $g(\dot\gamma(t), \dot\gamma(t)) \geq 0$ *for all* $t \in I$*.*

Definition 1.3.3. *A time-orientation on* $(\mathcal{M}, g)$ *is a nowhere vanishing timelike covector field* τ*, which means* $g_p^{-1}(\tau(p), \tau(p)) > 0$ *at each point* $p \in \mathcal{M}$*. Given a time-orientation* τ*, a causal/timelike curve* γ *is future directed if* $\tau(\gamma(t))\,[\,\dot\gamma(t)\,] > 0$ *for all* $t \in \mathbb{R}$ *and past directed if* $\tau(\gamma(t))\,[\,\dot\gamma(t)\,] < 0$ *for all* $t \in \mathbb{R}$*.*

Definition 1.3.4. *The timelike future/past of a region* $\mathcal{O}$*, denoted as* $I^+(\mathcal{O})$ / $I^-(\mathcal{O})$ *respectively, consists of all points* p *that have a future/past directed timelike curve (defined on a compact subset of* $\mathbb{R}$*) starting at a point of* $\mathcal{O}$ *and ending at* p*.*

The causal future/past of a region $\mathcal{O}$*, denoted as* $J^+(\mathcal{O})$ / $J^-(\mathcal{O})$ *respectively, consists of all points* p *that have a future/past directed causal curve starting at a point of* $\mathcal{O}$ *and ending at* p*. We will use the notation* $J(\mathcal{O})$ *to denote the union* $J^+(\mathcal{O}) \cup J^-(\mathcal{O})$*.*

We now introduce a type of compactness property specific to Lorentzian manifolds.

Definition 1.3.5. *A region* $\mathcal{O}$ *is time compact if for all* $p \in \mathcal{M}$*,* $J^+(p) \cap \mathcal{O}$ *and* $J^-(p) \cap \mathcal{O}$ *are compact.*

One of the central tenets of relativity is that given two events, which are modelled as points in the Lorentzian manifold, the past event can only affect the future event if there is a causal curve connecting them. From this emerges a useful concept in relativity known as the domain of dependence, which we now define.

[6] The Euler characteristic is the alternating sum of the ranks of the homology groups [Rot88, pp. 145]

Definition 1.3.6. *A future directed curve γ has a future endpoint $p \in \mathcal{M}$ if for any neighbourhood $N \ni p$, there exists $t_0 \in \mathbb{R}$ such that $\gamma(t) \in N$ for all $t > t_0$. We can define a past endpoint of γ similarly. The curve γ is inextendible if it has no past or future endpoints.*

Definition 1.3.7. *The future domain of dependence $D^+(\Sigma)$ of an achronal[7] hypersurface Σ, is given by all points p such that any inextendible past-directed causal curve passing through p must intersect Σ. We define the past domain of dependence $D^-(\Sigma)$ similarly. We call the union $D^+(\Sigma) \cup D^-(\Sigma)$ the Cauchy development of Σ, and denote it as $D(\Sigma)$.*

Many fundamental physical theories are described by fields satisfying hyperbolic PDEs with principal symbol given by the spacetime metric. This ensures that the fields satisfy causality, which means that if a field configuration is specified on an achronal hypersurface Σ, there is at most one solution to the field equations within the domain of dependence $D(\Sigma)$. We now introduce an important class of Lorentzian manifolds.

Definition 1.3.8. *A Lorentzian manifold $(\mathcal{M}, g)$ is globally hyperbolic, if it has a Cauchy surface, which is a closed achronal hypersurface Σ such that $D(\Sigma) = \mathcal{M}$.*

This implies that complete knowledge of the physical fields on $\mathcal{M}$ is given once Cauchy data on a Cauchy surface $\Sigma \subset \mathcal{M}$ has been specified. It was proven in [BS05], that the class of globally hyperbolic manifolds has the following property.

Theorem 1.3.9. *A globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$ is diffeomorphic to $\mathbb{R} \times \Sigma$ where for each $t \in \mathbb{R}$, $\{t\} \times \Sigma$ is a smooth spacelike Cauchy surface of $\mathcal{M}$, and the diffeomorphism ψ maps the metric g (via pullback) to a metric of the form*

$$\psi^*(g) = \beta \, dT \otimes dT - h_\Sigma(T) \tag{1.2}$$

where β is a smooth positive function on $\mathbb{R} \times \Sigma$, and $h_\Sigma(T)$ is a smooth 1-parameter family of Riemannian metrics on Σ.

We now define an important class of functions on Lorentzian manifolds.

Definition 1.3.10. *A Cauchy temporal function $\mathcal{T} : \mathcal{M} \to \mathbb{R}$ is a smooth function with timelike gradient everywhere, and Cauchy surfaces as level sets i.e, sets of the form $\mathcal{T}^{-1}(t)$.*

It is clear from theorem 1.3.9 that every globally hyperbolic manifold has an onto Cauchy temporal function, given by the time coordinate T in equation (1.2).

Two metrics g and $\widetilde{g}$ on a manifold $\mathcal{M}$ are said to be conformally related if there exists some positive function $f \in C^\infty(\mathcal{M})$ such that $\widetilde{g} = fg$. Metrics that are conformally related have the property that they define the same causal structure. This is due to the fact that causality is defined by causal curves, which are in turn defined by the condition that $g(v, v) \geq 0$ for each tangent vector v of the curve. We see that if $g(v, v) \geq 0$ and $\widetilde{g} = fg$ for any positive function $f \in C^\infty(\mathcal{M})$, then $\widetilde{g}(v, v) \geq 0$, hence curves that are causal in a metric g are also causal in any conformally related metric $\widetilde{g}$. We can use a conformal transformation to define the instantaneous optical metric on Cauchy surfaces as follows.

[7] No two points of the hypersurface can be connected by a timelike curve.

Definition 1.3.11. *For a globally hyperbolic manifold $(\mathcal{M}, g)$ where g is of the form in equation (1.2) , we obtain a metric $\widetilde{g} := g/\beta$ conformally related to g which is of the form*

$$\widetilde{g} = d\mathcal{T} \otimes d\mathcal{T} - k_\Sigma(\mathcal{T}) \quad , \tag{1.3}$$

where $\mathcal{T}(p) = T(p)/\sqrt{\beta(p)}$ and $k_\Sigma(\mathcal{T})$ is referred to as the instantaneous optical metric on $\{t\} \times \Sigma$.

Globally hyperbolic Lorentzian manifolds have the following useful properties.

Theorem 1.3.12. *For a globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$, the following statements are true*

- *$J^\pm(p) \cap J^\mp(q)$ is compact for all $p, q \in \mathcal{M}$.*
- *$J^\pm(K)$ is closed for all compact $K \subset \mathcal{M}$.*
- *$J^\pm(p) \cap \Sigma$ is compact for any Cauchy surface Σ and $p \in \mathcal{M}$.*

Proof. See theorems 8.3.10, 8.3.11 and 8.3.12 in [Wal84]. □

The first of these properties can be used to characterise globally hyperbolic Lorentzian manifolds in an alternative way.

Theorem 1.3.13. *A Lorentzian manifold $(\mathcal{M}, g)$ is globally hyperbolic if and only if it has no closed timelike curves and for each $p, q \in \mathcal{M}$, $J^+(p) \cap J^-(q)$ is compact.*

Proof. See [BS07]. □

We now introduce an important type of subset of a Lorentzian manifold.

Definition 1.3.14. *A subset $\mathcal{O} \in \mathcal{M}$ is causally convex if any causal curve connecting two points in $\mathcal{O}$ is entirely contained in $\mathcal{O}$.*

Using theorem 1.3.13, the following result can be proven.

Proposition 1.3.15. *An open causally convex subset $\mathcal{O}$ of a globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$ is itself a globally hyperbolic Lorentzian manifold in the relative topology when equipped with the metric $g|_\mathcal{O}$.*

Proof. Since $\mathcal{O}$ is an open subset of a manifold, it is a manifold in the relative topology. To see that $\mathcal{O}$ is globally hyperbolic with the metric $g|_\mathcal{O}$, let $p, q \in \mathcal{O}$ and consider the region $J^+_\mathcal{M}(p) \cap J^-_\mathcal{M}(q)$. This region is compact since $\mathcal{M}$ is globally hyperbolic, and entirely contained in $\mathcal{O}$ since $\mathcal{O}$ is causally convex. This implies that $J^+_\mathcal{O}(p) \cap J^-_\mathcal{O}(q)$ is also compact. Combining this with the fact that $\mathcal{O}$ has no closed timelike curves since $\mathcal{M}$ does not have any, theorem 1.3.13 implies $\mathcal{O}$ is globally hyperbolic. □

We now establish some standard results in Lorentzian geometry.

Lemma 1.3.16. *For fixed $\mathcal{O} \subset \mathcal{M}$, $J^\pm(p) \cap \mathcal{O}$ is compact for all p, if and only if $J^\pm(K) \cap \mathcal{O}$ is compact for all compact K.*

Proof. The only if part simply follows from the fact that the single point set $\{p\}$ is compact, so we focus on the if part. We can cover K by finitely many sets of the form $I^{\pm}(p_n) \cap K$ with $p_n \in \mathcal{M}$, since these sets form an open cover of K due to the fact that $I^{\pm}(p_n)$ is open, and any open cover of K has a finite open subcover of K, since K is compact. We therefore see that

$$J^{\pm}(K) \cap \mathcal{O} = \bigcup_n J^{\pm}\left(I^{\pm}(p_n) \cap K\right) \cap \mathcal{O} = \bigcup_n J^{\pm}(p_n) \cap \mathcal{O}$$

which implies $J^{\pm}(K) \cap \mathcal{O}$ is compact since it is the union of finitely many compact regions. □

Corollary 1.3.17. *In a globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$, given any Cauchy surface Σ of $\mathcal{M}$ and compact K, the region $J^{\pm}(K) \cap \Sigma$ is compact.*

Proof. Combine lemma 1.3.16 in the case $\mathcal{O} = \Sigma$ with the third property in theorem 1.3.12. □

Lemma 1.3.18. *In a globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$ and fixed $\mathcal{O} \subset \mathcal{M}$, if $J^{\pm}(K) \cap \mathcal{O}$ is compact for all compact K, then $J^{\pm}(K) \cap J^{\mp}(\mathcal{O})$ is also compact for all compact K.*

Proof. For any $p \in \mathcal{M}$ we have $J^{\pm}(p) \cap J^{\mp}(\mathcal{O}) \subset J^{\pm}(p) \cap J^{\mp}\left(J^{\pm}(p) \cap \mathcal{O}\right)$, since any future/past directed causal curve starting at p and ending in $J^{\mp}(\mathcal{O})$ can be extended so that it ends in $J^{\pm}(p) \cap \mathcal{O}$. We also have $J^{\pm}(p) \cap J^{\mp}(\mathcal{O}) \supset J^{\pm}(p) \cap J^{\mp}\left(J^{\pm}(p) \cap \mathcal{O}\right)$ because $\mathcal{O} \supset J^{\pm}(p) \cap \mathcal{O}$, hence $J^{\pm}(p) \cap J^{\mp}(\mathcal{O}) = J^{\pm}(p) \cap J^{\mp}\left(J^{\pm}(p) \cap \mathcal{O}\right)$.

By supposition $J^{\pm}(K) \cap \mathcal{O}$ is compact for all compact K, hence $J^{\pm}(p) \cap \mathcal{O}$ is compact for all $p \in \mathcal{M}$. The first property of theorem 1.3.12 together with lemma 1.3.16 implies $J^{\pm}(p) \cap J^{\mp}(K)$ is compact for all compact K and $\forall p \in \mathcal{M}$, and since $J^{\pm}(K) \cap \mathcal{O}$ is compact this implies $J^{\pm}(p) \cap J^{\mp}(\mathcal{O})$ is compact $\forall p \in \mathcal{M}$. We can then use 1.3.16 again to see that $J^{\pm}(K) \cap J^{\mp}(\mathcal{O})$ is compact for all compact K, thus proving the lemma. □

Corollary 1.3.19. *In a globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$, given any Cauchy surface Σ of $\mathcal{M}$ and compact K, the region $J^{\pm}(K) \cap J^{\mp}(\Sigma)$ is compact.*

Proof. Combine corollary 1.3.17 together with lemma 1.3.18 in the case $\mathcal{O} = \Sigma$. □

Corollary 1.3.20. *For a time compact region $\mathcal{O}$ and compact region K, $J^{\pm}(\mathcal{O}) \cap J^{\mp}(K)$ is compact.*

Proof. Since $\mathcal{O}$ is time compact, $J^{\pm}(p) \cap \mathcal{O}$ is compact for all $p \in \mathcal{M}$, which by lemma 1.3.16 implies $J^{\pm}(K) \cap \mathcal{O}$ is compact for all compact K. Combining this with lemma 1.3.18 proves the result. □

1.4 Category theory

Category theory is a useful way of abstracting various constructions that often appear in different branches of mathematics, and finding properties of these abstract constructions which can then be applied in a much wider range of contexts. It is akin to "not reinventing the wheel", a result can be proven in category theory rather than having to prove the same result over and over again in different contexts. The basic objects of category theory are categories, which we will now define. The first five definitions of this section can be found in [Mac78, §1].

Definition 1.4.1. *A category consists of a collection of objects, and a collection of morphisms each with a domain and codomain which are objects of the category. These collections may or may not be sets in the set-theoretic sense. The morphisms have the following structure:*

- *Any morphisms f and g, where the domain of f is the codomain of g, can be composed to get a morphism $f \circ g$. These compositions must be associative, so that $(f \circ g) \circ h = f \circ (g \circ h)$.*
- *Each object A has an identity morphism $\mathbb{1}_A$. For any morphism $f : A \to B$ we have $f \circ \mathbb{1}_A = \mathbb{1}_B \circ f$.*

Categories are often concrete[8], which roughly means that they consist of objects which are sets equipped with extra structure, and have morphisms which are maps between sets that "preserve" that structure (what this means is context dependent). For instance, the category of groups consists of sets equipped with a binary operation and an identity element obeying certain conditions, and maps which preserve those structures.

Category theory formalises the notion of equivalence between mathematical structures, which is usually given by a context dependent definition of isomorphism. In category theory, the definition of an isomorphism between two objects depends on the category that contains them.

Definition 1.4.2. *Let C be a category with objects A and B. An isomorphism is a C-morphism $f : A \to B$ that has an inverse C-morphism $g : B \to A$, which means $g \circ f = \mathbb{1}_A$ and $f \circ g = \mathbb{1}_B$.*

Another key concept of category theory is that of subcategories, which formalises the idea of one category being contained in another.

Definition 1.4.3. *A subcategory S of a category C is a collection of some of the objects of C and some of the morphisms of C such that; the domain and codomain in C of each morphism of S are also objects of S, the identity morphism of each object in S is also a morphism of S, and if S contains two morphisms of C it also contains their composition.*

A full subcategory S of C is a subcategory such that for any objects O_1, O_2 of S, all of the morphisms between O_1 and O_2 in C are also morphisms of S.

We now define mappings between categories.

Definition 1.4.4. *A functor $F : C \to D$ between categories C and D is a mapping which associates to each object $X \in C$ an object $F(X) \in D$, and associates to each morphism f of C a morphism $F(f)$ of D such that $F(\mathbb{1}_X) = \mathbb{1}_{F(X)}$ and $F(f \circ g) = F(f) \circ F(g)$,*

A commutative diagram is a diagram where all compositions of morphisms that map between any two fixed objects of the diagram are equal. Functors have the important

[8] A concrete category has a faithful functor to the category of sets.

property that they map commutative diagrams to commutative diagrams, and map isomorphisms to isomorphisms. A functor therefore provides a useful means of relating properties of objects in one category to those of another. For example, the assignment of homotopy groups to their associated topological spaces defined in 1.1.9, can be extended to a functor from the category Top_* of topological spaces with base point which has continuous base-point preserving maps as morphisms, to the category Grp of groups with homomorphisms as morphisms [Rot88, pp. 334]. Properties of the object of Top_* can be inferred from the properties of its associated object in Grp. We now define mappings between functors.

Definition 1.4.5. *A natural transformation N from the functor $F : C \to D$ to the functor $G : C \to D$ is a mapping which associates to each object $X \in C$, a map $N_X : F(X) \to G(X)$ such that the following diagram commutes for any morphism $f : X \to Y$ of C*

$$\begin{tikzcd} F(X) \arrow[r, "F(f)"] \arrow[d, "N_X"'] & F(Y) \arrow[d, "N_Y"] \\ G(X) \arrow[r, "G(f)"] & G(Y) \end{tikzcd}.$$

A natural isomorphism is a natural transformation N such that each N_X is an isomorphism.

An example of a natural transformation is the assignment to a topological space X of its Hurewicz homomorphisms (see theorem 1.1.14) which relates the homotopy and homology groups of X.

Another important concept in category theory is that of universal constructions, which guarantee the existence of unique morphisms that make certain diagrams commute (what types of diagram depends on the universal construction). The remaining definitions in this section are taken from [Mac78, §3]. We first give a couple of examples of universal constructions.

Definition 1.4.6. *Given a pair of objects X and Y of a category C, the product of X and Y* ***(if it exists)*** *is defined to be an object $X \times Y$ together with a pair of morphisms $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ such that for any pair of morphisms $f_X : Z \to X$ and $f_Y : Z \to Y$, there exists a unique morphism $\langle f_X, f_Y \rangle : Z \to X \times Y$ which makes the following diagram commute*

$$\begin{tikzcd} & Z \arrow[dl, "f_X"'] \arrow[d, dashed, "\langle f_X, f_Y \rangle"'] \arrow[dr, "f_Y"] & \\ X & X \times Y \arrow[l, "\pi_X"'] \arrow[r, "\pi_Y"] & Y \end{tikzcd}.$$

The universal property shows that any two products are unique up to isomorphism, hence the product is uniquely defined up to isomorphism. This is also true of the subsequent universal constructions we will be defining, their definitions are unique up to isomorphism. The product of topological spaces defined in definition 1.1.16 is an example of this construction in the category Top. There is also a dual construction known as the co-product which we now define.

Definition 1.4.7. *Given a pair of objects X and Y of a category C, the coproduct of X and Y **(if it exists)** is defined to be an object $X \bigsqcup Y$ together with a pair of morphisms $\iota_X : X \to X \bigsqcup Y$ and $\iota_Y : Y \to X \bigsqcup Y$ such that for any pair of morphisms $f_X : X \to Z$ and $f_Y : Y \to Z$, there exists a unique morphism $f_X \bigsqcup f_Y : X \bigsqcup Y \to Z$ which makes the following diagram commute*

$$\begin{tikzcd} & Z & \\ & & \\ X \arrow[uur, "f_X"] \arrow[r, "\iota_X"] & X \bigsqcup Y \arrow[uu, dashed, "f_X \bigsqcup f_Y"] & Y \arrow[uul, "f_Y"'] \arrow[l, "\iota_Y"'] \end{tikzcd}.$$

An example of a coproduct is the disjoint union of sets in the category of sets. In this thesis we will be concerned with a particular universal construction called a colimit, which generalises many other universal constructions in category theory.

Definition 1.4.8. *Consider a functor $F : I \to C$. A cocone over F is a pair (X, ϕ) where $X \in C$ and ϕ assigns to each $i \in I$ a C morphism $\phi_i : F(i) \to X$, such that given any I morphism $f : i \to j$ we have $\phi_i = \phi_j \circ F(f)$.*

*The colimit of F **(if it exists)**, is a cocone (X, ϕ) over F such that given any other cocone (Y, θ) over F, there exists a unique C morphism $U : X \to Y$ such that the following diagram commutes for all $f : i \to j$*

$$\begin{tikzcd} F(i) \arrow[rr, "F(f)"] \arrow[dr, "\phi_i"] \arrow[ddr, "\theta_i"'] & & F(j) \arrow[dl, "\phi_j"'] \arrow[ddl, "\theta_j"] \\ & X \arrow[d, dashed, "U"'] & \\ & Y & \end{tikzcd}.$$

The coproduct of C-objects X and Y is a special case of this construction, as can be seen by choosing the index category I to consist of two objects with no morphisms between each other, and F to be the functor mapping these objects to X and Y respectively.

Chapter 2

Axiomatic quantum field theory

Quantum field theory is the most stunningly accurate theory ever created, with predictions in quantum electrodynamics verified up to two parts in a billion [DM04]. Despite this, the theory still lacks precise mathematical underpinnings. This may not seem like a problem, if one's goal was only to make experimental predictions, since attempts at making QFT mathematically rigorous haven't produced any experimentally verifiable predictions that haven't already been made by the standard heuristic approach to QFT. The power of the mathematical approach is rather in clearing up the foundations of QFT, building up the theory from a small set of axioms. A clearer understanding of the conceptual underpinnings of the theory and how the key properties relate to each other could possibly lead to generalisations of the axioms which allow for a quantised theory of gravity.

One of the first attempts to place QFT on a solid mathematical basis were the Gårding–Wightman axioms [WG65]. In this framework, a theory is specified by a tuple $(\mathcal{H}, U, \mathcal{A}, \mathcal{D})$. The axioms of the framework in the case of a theory of a single scalar field $\mathcal{A}$ on Minkowski spacetime are:

(1) $\mathcal{H}$ is separable[1] Hilbert space and $U : \mathcal{P}_r \to \mathcal{H}$ is a strongly continuous[2] representation of the restricted (orientation and time-orientation preserving) Poincaré group $\mathcal{P}_r$.

Stone's theorem [RS81, Theorem 8.8] establishes a one-to-one correspondence between self-adjoint operators and one-parameter families of unitary operators. The self-adjoint operators corresponding to time translation and space translation implemented by U, are energy P^0 and momentum P^i respectively.

(2) The domain of P^μ is in $\mathcal{D}$ and the simultaneous spectrum of P^μ is contained in the closed forward light cone, equivalently $P^0 \geq 0$ and $P^\mu P_\mu \geq 0$.

(3) There exists a unique vector $\Omega \in \mathcal{D}$ which is invariant under the unitary action of spacetime translations implemented by U. This vector represents the vacuum state.

(4) The field $\mathcal{A}$, is a map from Schwartz functions on Minkowski spacetime to unbounded self-adjoint operators defined (together with their adjoints) on the dense subset $\mathcal{D} \subset$

[1] A seperable Hilbert space has a countable dense subset.

[2] For all $\epsilon > 0$ and $g_1, g_2 \in \mathcal{P}_r$, there exists $\delta > 0$ such that: $\|g_1 v - g_2 v\| < \delta \quad \forall v \in \mathbb{R}^4$ implies $\|U(g_1)\psi - U(g_2)\psi\| < \epsilon \quad \forall \psi \in \mathcal{H}$.

$\mathcal{H}$. These operators leave the domain $\mathcal{D}$ invariant and for fixed $\psi \in \mathcal{D}$, the mapping $f \mapsto \mathcal{A}(f)\psi$ is linear.

(5) For any ψ_1 and ψ_2 in $\mathcal{D}$, the map $f \mapsto \langle \psi_1, \mathcal{A}(f)\psi_2 \rangle$ where $\langle \cdot, \cdot \rangle$ is the inner product on $\mathcal{H}$, is a tempered distribution.

(6) The vacuum is cyclic i.e, the set of finite linear combinations of vectors of the form $\mathcal{A}(f_1)\cdots\mathcal{A}(f_n)\Omega$ is dense in $\mathcal{H}$.

(7) The action of U on fields is given by $U(a,\Lambda)^{-1}\mathcal{A}(f)U(a,\Lambda) = \mathcal{A}(\tilde{f})$ where $\tilde{f}(x) = f\left(\Lambda^{-1}(x-a)\right)$.

(8) For f_1 and f_2 with spacelike separated support, the corresponding operators $\mathcal{A}(f_1)$ and $\mathcal{A}(f_2)$ commute.

These axioms place Hilbert space and fields as central concepts which the theory is built upon. There is however reason to doubt that these concepts are best suited for describing the mathematical content of QFT. This is because of the existence of inequivalent Hilbert space representations of the canonical commutation relations.

For a theory obtained by quantising a classical system with a finite dimensional symplectic vector space, as is the case for non-relativistic quantum mechanics, this does not pose an issue. The Stone-von Neumann theorem states that up to isomorphism, there is a unique irreducible representation of the finitely generated canonical commutation relations on a Hilbert space[3]. This theorem does not apply to infinite dimensional systems however. An example of inequivalent representations for an infinite dimensional system is the van Hove model [Hov52], which is the quantum theory obtained from the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\nabla^\mu\phi)\nabla_\mu\phi - \frac{1}{2}m^2\phi^2 - \rho(\underline{x})\phi$$

where ρ is a time-independent real valued function or distribution. When $\rho \equiv 0$ we get free fields, in which case we can represent the fields as operators on the usual Bosonic Fock space. The algebras defined by the CCRs for different potentials $\rho(\underline{x})$ turn out to be isomorphic, the fields are just the free fields shifted by a scalar multiple of the identity (see section 3 in [FR19] for details). If however, ρ has δ-singularities or if either $m = 0$ or $\rho \equiv 1$, there is no unitary transformation that maps the field operators to the free field operators (for UV and IR reasons respectively) [FR19]. This means we have inequivalent Hilbert space representations of the same CCR algebra.

Due to the existence of inequivalent representations it seems that emphasis should be placed on the canonical commutation relations, or more generally the algebraic relations between observables/fields, rather than placing a particular Hilbert space representation front and centre as is done in the Gårding–Wightman framework. This motivated the algebraic approach to QFT, which instead emphasises algebras above their Hilbert space representations.

The rest of the chapter is structured as follows: We give an introduction to algebraic QFT in the first subsection. We then introduce the locally covariant framework for QFT

[3]Due to domain issues with unbounded operators, this uniqueness result actually applies to the Weyl algebra which is a C^*-algebra.

(abbreviated LCQFT) in the second subsection, and we end with a subsection outlining two examples of theories of LCQFT.

2.1 Algebraic quantum field theory

Algebraic QFT, sometimes known as AQFT or local QFT, is based on the Haag-Kastler axioms [HK64]. Before going through the axioms of AQFT, we begin by defining what an algebra is, and the different types of algebras considered in AQFT.

Definition 2.1.1. *An algebra $\mathscr{A}$ is a vector space over $\mathbb{C}$ equipped with an associative*[4] *bilinear map $\circ : \mathscr{A} \times \mathscr{A} \to \mathscr{A}$. A unital algebra has a distinguished element $\mathbb{1}$, such that $\mathbb{1} \circ A = A = A \circ \mathbb{1}$ for all $A \in \mathscr{A}$.*

A $*$-algebra is an algebra equipped with an anti-linear map $* : \mathscr{A} \to \mathscr{A}$, called the involution map, which satisfies $(A^*)^* = A$, $(\lambda A)^* = \overline{\lambda} A^*$ (where $\overline{\lambda}$ is the complex conjugate of λ) and $(A \circ B)^* = B^* \circ A^*$ for all $A, B \in \mathscr{A}$ and $\lambda \in \mathbb{C}$.

A C^*-algebra is a $*$-algebra equipped with a norm $\| \cdot \|$ such that the algebra is complete with respect to $\| \cdot \|$, $\|A \circ A^*\| = \|A\| \, \|A^*\|$ and $\|A \circ B\| \leq \|A\| \, \|B\|$ for all $A, B \in \mathscr{A}$.

A C^*-algebra is equipped with a natural topology induced from its norm, and can be represented as an algebra of bounded operators on a Hilbert space. The convention adopted in this thesis is that unless otherwise explicitly stated, all $*$-algebras and C^*-algebras will be unital. We now introduce categories of $*$ and C^* algebras.

Definition 2.1.2. *The categories Alg and C^*-Alg have $*$-algebras and C^*-algebras as objects respectively, and injective unit preserving homomorphisms as morphisms. The categories $\mathsf{Alg}^{(h)}$ and C^*-$\mathsf{Alg}^{(h)}$ are defined in the same way except the condition that the morphisms be injective is relaxed.*

The categories $\mathsf{Alg}^{(h)}$ and C^*-$\mathsf{Alg}^{(h)}$ are more appropriate for modelling quantum gauge theories due to the presence of topological charges, see for instance [DL12, DHS14]. We now list the axioms of AQFT on Minkowski spacetime:

- **Local algebras:** Each open bounded causally convex region $\mathcal{O}$ of Minkowski spacetime (which we denote as $\mathscr{M}$) is assigned a C^*-algebra $\mathscr{A}(\mathcal{O})$ (sometimes weakened to the assignment of $*$-algebras instead).

- **Isotony:** If $\mathcal{O}_1 \subset \mathcal{O}_2$, there exists an injective unit preserving homomorphism $i_{\mathcal{O}_1\mathcal{O}_2} : \mathscr{A}(\mathcal{O}_1) \to \mathscr{A}(\mathcal{O}_2)$, moreover if $\mathcal{O}_2 \subset \mathcal{O}_3$ then $i_{\mathcal{O}_2\mathcal{O}_3} \circ i_{\mathcal{O}_1\mathcal{O}_2} = i_{\mathcal{O}_1\mathcal{O}_3}$.

- **Einstein causality:** If $\mathcal{O}_1$ and $\mathcal{O}_2$ are causally disjoint, then for any $\mathcal{O}_3 \supset \mathcal{O}_1 \cup \mathcal{O}_2$, we get the following relation in $\mathscr{A}(\mathcal{O}_3)$

$$[i_{\mathcal{O}_1\mathcal{O}_3}(A_1), i_{\mathcal{O}_2\mathcal{O}_3}(A_2)] = 0 \quad \forall A_1 \in \mathscr{A}(\mathcal{O}_1) \ , \ \forall A_2 \in \mathscr{A}(\mathcal{O}_2) \quad .$$

- **Poincaré covariance:** For each $g \in \mathcal{P}_r$ and each $\mathcal{O}$, there is an isomorphism $\alpha_{g,\mathcal{O}} : \mathscr{A}(\mathcal{O}) \to \mathscr{A}(g\mathcal{O})$ such that $\alpha_{g_1,g_2\mathcal{O}} \circ \alpha_{g_2,\mathcal{O}} = \alpha_{g_1g_2,\mathcal{O}}$ for all $g_1, g_2 \in \mathcal{P}_r$ and $i_{g\mathcal{O}_1,g\mathcal{O}_2} \circ \alpha_{g,\mathcal{O}_1} = \alpha_{g,\mathcal{O}_2} \circ i_{\mathcal{O}_1\mathcal{O}_2}$.

[4] Sometimes the associative property is dropped in the definition of algebras.

- **Timeslice axiom:** If $\mathcal{O}_1 \subset \mathcal{O}_2$ and $\mathcal{O}_1$ contains a Cauchy surface of $\mathcal{O}_2$, then the map $i_{\mathcal{O}_1\mathcal{O}_2}$ is an isomorphism.

The set of open bounded causally convex regions of Minkowski spacetime form an upward directed set with respect to inclusion, which implies the collection of local algebras is a net. Combining this with the isotony condition allows us to construct a global algebra, called the quasi-local algebra $\mathscr{A}(\mathscr{M})$, by taking the colimit of the net of local algebras (known as the inductive limit since in this case the net is up-directed).

In order for the theory to be able to make predictions that can be compared with experiment, we need states on the net of local algebras. A state is defined as follows.

Definition 2.1.3. *A state on an algebra $\mathscr{A}$, is a linear map $\omega : \mathscr{A} \to \mathbb{C}$ such that $\omega(\mathbb{1}) = 1$ and $\omega(A^*A) \geq 0$ for all $A \in \mathscr{A}$.*

Given a state on a (C)*-algebra, we can obtain a Hilbert space representation of the algebra. This is done by using the GNS construction, which takes a state ω and yields a tuple $(H_\omega, D_\omega, \Omega_\omega, \pi_\omega)$ consisting of a Hilbert space H_ω with a dense subspace D_ω, a cyclic vector Ω_ω and a representation $\pi_\omega : \mathscr{A}(\mathscr{M}) \to \mathrm{End}(D_\omega)$ such that

$$\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle \quad .$$

Moreover, this tuple is unique in the sense that for any other tuple $(H'_\omega, D'_\omega, \Omega'_\omega, \pi'_\omega)$ satisfying these conditions, there exists a unitary map $U_\omega : H_\omega \to H'_\omega$ such that $U_\omega \pi_\omega(A)\Omega_\omega = \pi'_\omega(A)\Omega'_\omega$ for all $A \in \mathscr{A}(\mathscr{M})$.

The vacuum state [Haa96, Definition 3.2.3] on the quasi-local algebra $\mathscr{A}(\mathscr{M})$, is a state such that $\omega(A^*A) = 0$ implies A is of the form

$$A = B \int d^4x \; f(x)\, \alpha_x(C)$$

where $B, C \in \mathscr{A}(\mathscr{M})$, α_x is the automorphism for translation by x and f has a Fourier transform with support outside of the closed forward light cone. The integral results in an algebra element corresponding to an operator that imparts an energy-momentum transfer within the support of the Fourier transform of f. Since f has support outside the closed forward light cone, A^*A imparts negative energy in some Lorentz frame and therefore registers the presence of matter, thus $\omega(A^*A) = 0$ implies ω is the vacuum state.

The representations induced by different states can be unitarily inequivalent, and some states may yield Hilbert spaces which lack certain useful properties. It is therefore common to consider subsets of states satisfying certain conditions. For instance, the DHR (Doplicher, Haag and Roberts) selection criterion requires the GNS representation of a state satisfying the condition to be unitarily equivalent to the vacuum state representation outside of a sufficiently large diamond[5]. The field algebra and gauge group can be reconstructed from the category of GNS representations of DHR states of the observable algebra [DHR69a, DHR69b, DHR71, DHR74]. This suggests that the structure of the category of representations is the really interesting theoretical content of AQFT.

Having briefly introduced AQFT, we are now ready to introduce locally covariant QFT.

[5] Cauchy development of a subset of a Cauchy surface.

2.2 Locally covariant quantum field theory

The techniques of AQFT from the previous subsection can be generalised to curved spacetimes, as was done in the seminal paper of Brunetti, Fredenhagen and Verch [BFV03]. To see this, we begin by noting that the net of local algebras, the central object of AQFT, defines a functor from the category of open bounded causally convex subsets of Minkowski spacetime with inclusion maps as morphisms, to the category Alg of (C)*-algebras with injective (due to the isotony condition) morphisms. It seems reasonable then to modify the domain category of the functor, so that it assigns algebras to spacetimes in their own right, rather than just certain subregions of Minkowski spacetime. For this purpose we introduce the following category.

Definition 2.2.1. *The category Loc has objects consisting of globally hyperbolic spacetimes $(\mathcal{M}, g)$ which are connected, with orientation $\mathfrak{o}$ and time-orientation τ. We will simply use $\mathcal{M}$ to denote the object and suppress the additional structure in the notation. The morphisms of Loc are smooth isometric open embeddings with causally convex image which preserve orientation and time-orientation.*

A theory in locally covariant QFT (LCQFT) is specified by a functor $\mathscr{A}$ from Loc to Alg. Aspects of LCQFT can be studied by placing physical assumptions on $\mathscr{A}$, such as Einstein causality and the timeslice axiom which we define in this framework in the following definitions. We note that the Einstein causality condition has other formulations, such as the monoidal formulation [BFIR14] and the operadic formulation [BSW17].

Definition 2.2.2. *A theory $\mathscr{A} : \mathsf{Loc} \to \mathsf{Alg}$ satisfies Einstein causality if for any Loc morphisms ψ_1 and ψ_2 with a common codomain and images which are causally disjoint in that codomain, the images of $\mathscr{A}(\psi_1)$ and $\mathscr{A}(\psi_2)$ commute in their common codomain.*

Definition 2.2.3. *A theory $\mathscr{A} : \mathsf{Loc} \to \mathsf{Alg}$ satisfies the timeslice axiom if for any Loc morphism ψ with image containing a Cauchy surface of its codomain, the map $\mathscr{A}(\psi)$ is an isomorphism.*

This gives us a model-independent way of studying QFT in curved spacetimes. The framework also has the advantage that it is manifestly covariant, since there is no particular spacetime that plays a central role. A quantum field can be defined in this framework as follows.

Definition 2.2.4. *A quantum field in a theory $\mathscr{A} : \mathsf{Loc} \to \mathsf{Alg}$ is defined as a natural transformation $\Phi : \mathscr{D} \to F \circ \mathscr{A}$ where $F : \mathsf{Alg} \to \mathsf{Set}$ is the forgetful functor and $\mathscr{D} : \mathsf{Loc} \to \mathsf{Set}$ is a general functor.*

In most cases the functor $\mathscr{D}$ assigns to each spacetime $\mathcal{M}$ the set of sections we "smear" our field against (in the sense of Wightman fields defined at the start). So for instance, if Φ was the real scalar field then $\mathscr{D}(\mathcal{M}) = C_0^\infty(\mathcal{M})$.

An important concept in LCQFT is *relative Cauchy evolution* (from now on referred to as RCE). RCE compares the dynamics of a system, with those of the system subject to a compact/time-compact perturbation to its spacetime metric. The functional derivative of the RCE operator can be interpreted as the stress-energy tensor [FV15], which indicates that the RCE operator can be considered as a proxy for the action of the theory. This shows that there is an action principle present in the framework. It also has applications to describing what it means for physics to be the same on all spacetimes [FV12].

For a given spacetime $\mathcal{M}$ and metric perturbation supported in a compact/time-compact region K, we define $\mathcal{M}^+$ to be $\mathcal{M}\backslash J^-(K)$ and $\mathcal{M}^-$ to be $\mathcal{M}\backslash J^+(K)$. These regions will contain Cauchy surfaces of $\mathcal{M}$ and the perturbed spacetime $\mathcal{M}'$ [BFV03,FV12]. Since $\mathcal{M}^+$ and $\mathcal{M}^-$ are both contained in $\mathcal{M}$ and $\mathcal{M}'$, there exist canonical inclusion maps, as shown in the diagram below

$$\begin{tikzcd}
& \mathcal{M}^+ \arrow[dl, "\iota^+"'] \arrow[dr, "\iota'^+"] & \\
\mathcal{M} & & \mathcal{M}' \\
& \mathcal{M}^- \arrow[ul, "\iota^-"] \arrow[ur, "\iota'^-"'] &
\end{tikzcd},$$

which define Cauchy morphisms. Hence by the timeslice axiom, when we apply our functor to the morphisms in this diagram we will get algebra isomorphisms. This means the arrows can be inverted and we get a diagram relating the algebras of $\mathcal{M}$ and $\mathcal{M}'$

$$\begin{tikzcd}
& \mathscr{A}(\mathcal{M}^+) \arrow[dr, "\mathscr{A}(\iota'^+)"] & \\
\mathscr{A}(\mathcal{M}) \arrow[ur, "\mathscr{A}(\iota^+)^{-1}"] & & \mathscr{A}(\mathcal{M}') \arrow[dl, "\mathscr{A}(\iota'^-)^{-1}"] \\
& \mathscr{A}(\mathcal{M}^-) \arrow[ul, "\mathscr{A}(\iota^-)"] &
\end{tikzcd}.$$

RCE is defined as the composition of the morphisms going clockwise around the diagram above. In general, this is a non-trivial operation since the diagram is not necessarily commutative.

2.3 Complex scalar field and the Dirac field in LCQFT

We end the chapter with a couple of examples of theories in the LCQFT framework. First is the complex scalar field which we will define after we have defined advanced/retarded Green operators for complex valued functions on $\mathcal{M}$.

Definition 2.3.1. *Let P be a linear differential operator on smooth functions $C^\infty(\mathcal{M})$. An advanced/retarded Green operator for P is a linear operator $E^\mp : C_0^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ (where $C_0^\infty(\mathcal{M})$ denotes compactly supported functions) which satisfies*

- $P \circ E^\mp = \mathcal{I}$, *where $\mathcal{I}$ is the inclusion map from $C_0^\infty(\mathcal{M})$ to $C^\infty(\mathcal{M})$.*
- $E^\mp \circ P|_{C_0^\infty(\mathcal{M})} = \mathcal{I}$.
- $\mathrm{supp}(E^\mp f) \subset J^\mp(\mathrm{supp} f)$.

We use the notation $E := E^- - E^+$ for the advanced minus retarded Green operator of P and use the notation $(f, g) := \int_\mathcal{M} \mathrm{dvol}_g\, f\,\overline{g}$ and $E(f, g) := \int_\mathcal{M} \mathrm{dvol}_g\, f\,\overline{Eg}$. The operator P is said to be Green hyperbolic if it admits advanced and retarded Green operators (in *which case the Green operators are unique). We say that P is formally self-adjoint if $(f, Pg) = (Pf, g)$ for all $f, g \in C^\infty(\mathcal{M})$ with compactly intersecting supports.*

Definition 2.3.2. *The theory of the complex scalar field is given by the functor $\mathscr{A} : \mathsf{Loc} \to \mathsf{Alg}$ which for a Loc object $\mathcal{M}$, assigns the algebra $\mathscr{A}(\mathcal{M})$ with generators $\boldsymbol{\Phi}_{\mathcal{M}}(f)$ indexed by smooth compactly supported test functions $f \in C_0^\infty(\mathcal{M})$ and an identity element $\mathbb{1}$ which satisfy the following relations:*

- *The mapping $f \mapsto \boldsymbol{\Phi}_{\mathcal{M}}(f)$ is linear.*
- *$\boldsymbol{\Phi}_{\mathcal{M}}(P_{\mathcal{M}} f) = 0$ where $P_{\mathcal{M}} = \nabla^\mu \nabla_\mu + m^2$.*
- *$[\boldsymbol{\Phi}_{\mathcal{M}}(f), \boldsymbol{\Phi}_{\mathcal{M}}(g)]_- := \boldsymbol{\Phi}_{\mathcal{M}}(f)\boldsymbol{\Phi}_{\mathcal{M}}(g) - \boldsymbol{\Phi}_{\mathcal{M}}(g)\boldsymbol{\Phi}_{\mathcal{M}}(f) = 0$.*
- *$[\boldsymbol{\Phi}_{\mathcal{M}}(f), \boldsymbol{\Phi}_{\mathcal{M}}(g)^*]_- = iE_{\mathcal{M}}(f, g)\mathbb{1}$ where $E_{\mathcal{M}}$ is the advanced minus retarded Green function for the operator $P_{\mathcal{M}}$.*

For a Loc morphism $\psi : \mathcal{M} \to \mathcal{N}$, the theory maps this morphism to $\mathscr{A}(\psi)$ whose action on generators is given by

$$\mathscr{A}(\psi)\, \boldsymbol{\Phi}_{\mathcal{M}}(f) = \boldsymbol{\Phi}_{\mathcal{N}}(\psi_* f) \tag{2.1}$$

where ψ_* is given by

$$(\psi_* f)(p) = \begin{cases} f(\psi^{-1}(p)) & \text{if } p \in \psi(\mathcal{M}) \\ 0 & \text{otherwise} \end{cases}.$$

The morphism $\mathscr{A}(\psi)$ is defined on the whole of $\mathscr{A}(\mathcal{M})$ by extending it to a homomorphism, which is possible since equation (2.1) is compatible with the relations. Two algebra elements $\boldsymbol{\Phi}(f)$ and $\boldsymbol{\Phi}(f')$ are equal if and only if $f - f' = P_{\mathcal{M}} g$ for some $g \in C_0^\infty(\mathcal{M})$, and $\psi_* f - \psi_* f' = P_{\mathcal{N}} g'$ for some $g' \in C_0^\infty(\mathcal{N})$ if and only if $f - f' = P_{\mathcal{M}} g$ for some $g \in C_0^\infty(\mathcal{M})$, hence $\mathscr{A}(\psi)$ is injective and therefore a valid morphism of Alg.

This theory satisfies the standard axioms of LCQFT; namely Einstein causality and the timeslice axiom. To see that Einstein causality is satisfied, we note that the advanced minus retarded Green function $E_{\mathcal{M}}$ has the property that $E_{\mathcal{M}}(f, g) = 0$ if the functions f and g have spacelike separated support. This is due to the fact that $E^\pm_{\mathcal{M}}$ are Green operators, so by definition $\mathrm{supp}(E^\pm_{\mathcal{M}} f) \subset J^\pm(\mathrm{supp} f)$. Therefore, the commutator for spacelike separated fields vanishes, hence Einstein causality is satisfied. To see that the timeslice axiom is satisfied we follow the proof outlined in [FV15, pp. 9-10]. Let $O(\Sigma)$ be any open causally convex neighbourhood of a Cauchy surface Σ, and let Σ^+ and Σ^- be Cauchy surfaces to the future/past of Σ that are contained in $O(\Sigma)$. Then let ρ be a function which vanishes to the future of Σ^+ and equals 1 to the past of Σ^-. For any $f \in C_0^\infty(\mathcal{M})$, we can construct the following function using ρ

$$\widetilde{f} = P_{\mathcal{M}}\; \rho\; E_{\mathcal{M}} f$$

which has compact support in $O(\Sigma)$ and can be shown to have the property that
$f - \widetilde{f} \in P_{\mathcal{M}} C_0^\infty(\mathcal{M})$. This implies that for the inclusion map $\psi : O(\Sigma) \to \mathcal{M}$ we have

$\mathscr{A}(\psi)\boldsymbol{\Phi}_{O(\Sigma)}(\psi^*\widetilde{f}) = \boldsymbol{\Phi}_{\mathcal{M}}(\widetilde{f}) = \boldsymbol{\Phi}_{\mathcal{M}}(f)$, hence $\mathscr{A}(\psi)$ is surjective onto generators and hence to the whole of $\mathscr{A}(\mathcal{M})$. Combining this with the fact that $\mathscr{A}(\psi)$ is injective (as noted in definition 2.3.2) implies $\mathscr{A}(\psi)$ is an isomorphism, therefore the timeslice axiom is satisfied.

We now switch our focus to an exposition of the Dirac field. In order to describe the quantum Dirac field, we must first describe classical spinor fields. For ease of exposition we will describe how they are defined on four dimensional globally hyperbolic spacetimes, however they can be defined on more general spacetimes. We use the following reference material [Wal84, §13] [San10] [Ish78] [Dim82] for the following exposition.

Spinor fields on Minkowski spacetime are vector valued fields i.e, sections of a vector bundle over Minkowski spacetime, which transform in the projective rep $(0, 1/2) \oplus (1/2, 0)$ of the restricted Lorentz group (denoted by $SO^+(1,3)$), this projective rep being a homomorphism $\widetilde{\rho} : SO^+(1,3) \to GL(4,\mathbb{C})/Z_2$. This rep can be viewed as an ordinary rep of the double cover of $SO^+(1,3)$ (denoted by $Spin(1,3)$) i.e, there exists a homomorphism $\rho : Spin(1,3) \to GL(4,\mathbb{C})$ such that $q \circ \rho = \widetilde{\rho} \circ \Lambda$ where $q : GL(4,\mathbb{C}) \to GL(4,\mathbb{C})/\mathbb{Z}_2$ is the quotient map and Λ is the double covering homomorphism from $Spin(1,3)$ to $SO^+(1,3)$. There is an isomorphism from $Spin(1,3)$ to $SL(2,\mathbb{C})$, so from now on we will simply refer to $Spin(1,3)$ as $SL(2,\mathbb{C})$. The rep ρ cannot be unitary since $SL(2,\mathbb{C})$ is non-compact, and therefore has no unitary finite dimensional reps. We can however define a sesquilinear form $\langle , \rangle$ on C^4, such that $\langle \rho(s)u_1, \rho(s)u_2\rangle = \langle u_1, u_2\rangle$ for all $s \in SL(2,\mathbb{C})$ and $u_1, u_2 \in C^4$, the cost being that this sesquilinear form is not positive definite. The sesquilinear form $\langle , \rangle$ can be used to define an antilinear isomorphism known as the Dirac adjoint from $\mathbb{C}^4$ to its dual space $\overline{\mathbb{C}^4}$.

In a curved Lorentzian spacetime $(\mathcal{M}, g)$, Lorentz symmetry becomes a local symmetry. Our theories should therefore be invariant under local Lorentz transformations, which consist of a set of Lorentz transformations at each tangent space $T_p\mathcal{M}$ such that the assignment of Lorentz transformations to each point forms a smooth function. In order to make this definition more precise, we can use the notion of fibre bundles introduced in definition 1.1.18. From this perspective, local Lorentz transformations are understood in terms of the frame bundle, which consists of the disjoint union of ordered orthonormal bases of the tangent space at each point in the manifold. The frame bundle can be identified with its associated principal bundle $\mathcal{FM} = \big(FM, \pi_{FM}, \mathcal{M}, SO^+(1,3)\big)$. A local Lorentz transformation then consists of a vertical bundle automorphism i.e, a diffeomorphism $\psi : FM \to FM$ such that for all $p \in FM$ and $g \in SO^+(1,3)$, $\psi(p \circ g) = \psi(p) \circ g$ and $\pi_{FM} = \pi_{FM} \circ \psi$.

To define spinor fields on curved spacetimes, we must adapt the transformation properties of spinors with respect to global Lorentz transformations outlined above, to transformation properties with respect to local Lorentz transformations. This requires us to find a way of lifting the local Lorentz transformations to local $SL(2,\mathbb{C})$ transformations. In order to do this, we need to lift a vertical bundle automorphism of $\mathcal{FM}$ to a vertical bundle automorphism of some principal bundle with structure group $SL(2,\mathbb{C})$. This requires a choice of spin structure which we will now define.

Definition 2.3.3. *A spin structure consists of a principal $SL(2,\mathbb{C})$ bundle $\mathcal{SM} = \big(SM, \pi_{SM}, \mathcal{M}, SL(2,\mathbb{C})\big)$ with right action $R : SM \times SL(2,\mathbb{C}) \to SM$ and a smooth mapping $f : SM \to FM$, such that $f(R(p,s)) = f(p) \circ \Lambda(s)\ \ \forall s \in SL(2,\mathbb{C}),\ p \in SM$. The* *map "preserves fibres" in the sense that the following diagram commutes*

$$\begin{tikzcd} SM \arrow[rr, "f"] \arrow[dr, "\pi_{SM}"'] & & FM \arrow[dl, "\pi_{FM}"] \\ & \mathcal{M} & \end{tikzcd} \quad .$$

Two spin structures $(\mathcal{SM}_1, f_1)$ and $(\mathcal{SM}_2, f_2)$ are considered equivalent if there exists a diffeomorphism $\mathcal{F} : SM_1 \to SM_2$ such that $\mathcal{F}(p \circ s) = \mathcal{F}(p) \circ s \ \forall s \in SL(2, \mathbb{C}),\ p \in SM$ and $f_1 = f_2 \circ \mathcal{F}$.

The existence of spin structures is not guaranteed however. For an orientable manifold to admit a spin structure, the second Stiefel-Whitney class of the tangent bundle $w_2(\mathcal{TM}) \in H^2(\mathcal{M}; \mathbb{Z}_2)$ is required to be trivial [BH59, pp. 350]. This condition is needed so that the transition functions of the frame bundle can be lifted to transition functions with values in $SL(2, \mathbb{C})$ in a way which preserves the cocycle condition $t_{ij} \circ t_{jk} = t_{ik}$, so that the lift gives a well defined principal bundle. This condition is always met in globally hyperbolic orientable four dimensional spacetimes [Ger70], so objects of Loc necessarily admit spin structures. We are now in a position to define the background category for the Dirac field.

Definition 2.3.4. *The category SpinLoc is the category whose objects are spin structures $(\mathcal{SM}, f)$ with the base space of $\mathcal{SM}$ being an object of Loc. A morphism between objects $(\mathcal{SM}_1, f_1)$ and $(\mathcal{SM}_2, f_2)$ in this category is given by a map $\chi : SM_1 \to SM_2$ that satisfies the following conditions:*

- *It covers a Loc morphism $\psi : \mathcal{M}_1 \to \mathcal{M}_2$ between the base spaces of $\mathcal{SM}_1$ and $\mathcal{SM}_2$ i.e, $\pi_2 \circ \chi = \psi \circ \pi_1$.*
- *It intertwines the right actions of $\mathcal{SM}_1$ and $\mathcal{SM}_2$ i.e, $R_2(\chi(p), s) = \chi(R_1(p, s))$.*
- *It satisfies $f_2 \circ \chi = \psi_* \circ f_1$ where ψ_* is the induced map on $\mathcal{FM}$ arising from the tangent map of ψ.*

The set of inequivalent spin structures are in one to one correspondence with elements of $H^1(\mathcal{M}; \mathbb{Z}_2)$ [Mil63b]. The group $H^1(\mathcal{M}; \mathbb{Z}_2)$ is in one to one correspondence with the set of homomorphisms from the fundamental group $\pi_1(\mathcal{M})$ to $\mathbb{Z}_2$, as is proven in corollary 1.1.15. A more detailed overview of these ideas can be found in [Ish78, San10].

Given a choice of spin structure, we can define the spinor and cospinor bundles whose sections define the Dirac field. The following construction is a special case of a more general construction known as the associated bundle construction.

Definition 2.3.5. *The spinor bundle is $\mathcal{DM} = \big(DM, \pi_{DM}, \mathcal{M}, \mathbb{C}^4, \rho(SL(2, \mathbb{C}))\big)$ where $\rho : Spin(1, 3) \to GL(4, \mathbb{C})$ is the rep $(0, 1/2) \oplus (1/2, 0)$ we discussed earlier, and DM consists of equivalence classes of pairs from SM and $\mathbb{C}^4$ satisfying the following equivalence relation*

$$(R(p, s), u) \sim (p, \rho(s)u) \quad .$$

Similarly the cospinor bundle is $\overline{\mathcal{DM}} = \left(\overline{DM}, \pi_{\overline{DM}}, \mathcal{M}, \overline{\mathbb{C}^4}, \rho(SL(2,\mathbb{C}))\right)$ *where* $\overline{DM}$ *consists of equivalence classes of pairs from* SM *and* $\overline{\mathbb{C}^4}$ *such that*

$$(R(p,s), v) \sim (p, \overline{\rho(s)}v)$$

where $\overline{\rho(s)}$ *is the Dirac adjoint of* $\rho(s)$*. The transition functions of these bundles are inherited from the spin bundle, which is specified by the choice of spin structure. Given a section* u *of* $\mathcal{DM}$ *and a section* v *of* $\overline{\mathcal{DM}}$*, there is a canonical pairing*

$$v[u] = \int_{\mathcal{M}} dvol_g(x)\ v(x)[u(x)] \quad .$$

We can define a covariant derivative on spinor and cospinor fields by pulling back the Levi-Civita connection on $\mathcal{FM}$ to $\mathcal{SM}$, which can then be used to define a connection on $\mathcal{DM}$ and $\overline{\mathcal{DM}}$ (a precise account can be found in [Dim82]). With all this in place, we can now define the quantum Dirac field.

Definition 2.3.6. *The theory of the Dirac field is given by the functor* $\mathscr{A} : \mathsf{SpinLoc} \to \mathsf{Alg}$ *which for a* $\mathsf{SpinLoc}$ *object* $\mathcal{SM}$ *(using the spin bundle to denote the object and suppressing the additional structure in the notation),* $\mathscr{A}(\mathcal{SM})$ *is the algebra with generators* $\overline{\Psi}_{\mathcal{SM}}(u)$ $\Psi_{\mathcal{SM}}(v)$ *indexed by smooth compactly supported test sections* $u \in C_0^\infty(\mathcal{M}, \mathcal{DM})$ *and* $v \in C_0^\infty(\mathcal{M}, \overline{\mathcal{DM}})$ *respectively, and an identity element* $\mathbb{1}$ *which satisfy the following relations:*

- *The mappings* $u \mapsto \overline{\Psi}_{\mathcal{SM}}(u)$ *and* $v \mapsto \Psi_{\mathcal{SM}}(v)$ *are linear.*
- $\overline{\Psi}_{\mathcal{SM}}(u)^* = \Psi_{\mathcal{SM}}(\overline{u})$.
- $\overline{\Psi}_{\mathcal{SM}}(P_{\mathcal{SM}}\, u) = 0$ *where* $P_{\mathcal{SM}} = -ie^\mu_a\gamma^a\nabla_\mu + m$ *where* e^μ_a *are frame components*[6]*,* γ^a *are a fixed set of gamma matrices satisfying the Clifford algebra relations, and* ∇_μ *is the covariant derivative on spinor fields discussed above.*
- $\left[\overline{\Psi}_{\mathcal{SM}}(u_1), \overline{\Psi}_{\mathcal{SM}}(u_2)\right]_+ := \overline{\Psi}_{\mathcal{SM}}(u_1)\overline{\Psi}_{\mathcal{SM}}(u_2) + \overline{\Psi}_{\mathcal{SM}}(u_2)\overline{\Psi}_{\mathcal{SM}}(u_1) = 0.$
- $\left[\overline{\Psi}_{\mathcal{SM}}(u), \Psi_{\mathcal{SM}}(v)\right] = iv[E_{\mathcal{SM}}u]\mathbb{1}$ *where* $E_{\mathcal{SM}}$ *is the advanced minus retarded Green function for the operator* $P_{\mathcal{SM}}$*.*

For a $\mathsf{SpinLoc}$ *morphism* $\chi : \mathcal{SM} \to \mathcal{SN}$*, the theory maps this morphism to* $\mathscr{A}(\chi)$ *where*

$$\mathscr{A}(\chi)\, \overline{\Psi}_{\mathcal{SM}}(u) = \overline{\Psi}_{\mathcal{SN}}(\chi_* u)$$
$$\mathscr{A}(\chi)\, \Psi_{\mathcal{SM}}(v) = \Psi_{\mathcal{SN}}(\chi_* v)$$

where the action of χ_* *on a spinor field* $u(p) = [S(p), U(p)]_{\mathcal{DM}}$ *(the equivalence class is defined in definition 2.3.5) is given by*

$$(\chi_* u)(p) = \begin{cases} [\,\chi^{-1}(S(p))\,,\,U(p)\,]_{\mathcal{DN}} & \text{if } S(p) \in \chi(\mathcal{SM}) \\ 0 & \text{otherwise} \end{cases}$$

[6] Frame components e^μ_a satisfy $g^{\mu\nu} = \sum_{a,b} \eta^{ab} e^\mu_a e^\nu_b$ where η^{ab} is the Minkowski metric.

and the action of χ_* *on a cospinor field* $v(p) = [\overline{S}(p), V(p)]_{\overline{\mathcal{DM}}}$ *is given by*

$$(\chi_* v)(p) = \begin{cases} [\, \chi^{-1}(\overline{S}(p))\,,\, V(p)\,]_{\overline{\mathcal{DN}}} & \text{if } \overline{S}(p) \in \chi(\mathcal{SM}) \\ 0 & \text{otherwise} \end{cases} .$$

Chapter 3

Einstein causality in universal algebras

In some situations a theory may only be defined on spacetimes with special properties, in which case it is natural to ask if one can extend the theory to more general spacetimes. For a given spacetime $\mathcal{M}$, this can be done by considering all the subregions of $\mathcal{M}$ on which the theory is defined. The theory assigns an algebra to each of these subregions and if one subregion embeds inside another, the theory gives us a map between the corresponding algebras. Using this set of algebras and the relations between them, which is called a *system of local algebras*, we can construct a *universal algebra* [Fre90] (which we define in definition 3.1.11) associated to $\mathcal{M}$. Given an inclusion morphism $\psi : \mathcal{M} \to \mathcal{N}$, there is a canonical way of assigning a morphism $\mathscr{U}(\psi)$ between the associated universal algebras (to be defined in the next section). The mapping of spacetimes to their corresponding universal algebras defines a functor, and thus defines an extension of the original theory that was used to construct the nets of local algebras [Lan12]. A precise account of this construction, which is known in category theory as a left Kan extension, will be given in the next section.

In this chapter we shall consider a general class of theories modelled on the theory of the free scalar/Dirac field. For a given theory $\mathscr{A}$ in this class, we define a restricted theory $\mathscr{A}'$ by restricting $\mathscr{A}$ to contractible spacetime regions. The motivation for considering contractible regions is that they comprise topologically simple regions of spacetime, which model "small" regions of spacetime that are not sensitive to the global topology. We then use the universal algebra techniques discussed above to get an extended theory $\mathscr{U}$ from $\mathscr{A}'$. The main result of the chapter is that the resulting class of universal theories satisfy Einstein causality, which is far from obvious given how $\mathscr{U}$ is defined. The motivation for proving this result relates to another method for defining extended theories. In [BSW17], the authors consider a framework for QFT which involves categories with an additional orthogonal structure defined by causal disjointness. Theories in this framework are defined as functors that preserve the orthogonal structure i.e, the induced commutator map for a pair of orthogonal morphisms is the zero map. The pair formed by a category and its orthogonal structure is called an orthogonal category, and to each orthogonal category one can associate a colored operad. Given an embedding of one orthogonal category into another, there exists an associated operad map that defines an adjunction between the corresponding categories of algebras. This is called the operadic left Kan extension in the literature. This adjunction can then be used to define extended theories in such a way that Einstein causality is necessarily preserved. It is shown [BSW17, Proposition 5.1], that if the left Kan extension satisfies Einstein causality, then it will be equivalent to the operad construction. This chapter therefore shows that the operad construction coincides with the left Kan extension, when extending the class of theories we consider from globally hyperbolic contractible spacetimes to globally hyperbolic connected spacetimes.

In order to prove Einstein causality for the class of extended theories we consider, we introduce some geometrical techniques which also have further applications to analysing universal algebras. We use these techniques to prove in addition to Einstein causality, the following result: Let $\mathscr{A}$ be a theory of unobservable fields (see definition 3.1.5), then the universal theory $\mathscr{U}$ built up from the restriction of $\mathscr{A}$ to contractible regions, is equivalent to $\mathscr{A}$. This is a useful result which implies that universal theories of this type will be non-trivial and satisfy Einstein causality. This result is also a significant generalisation of results obtained by Brunetti, Franceschini and Moretti [BFM09, Proposition B.0.8]. A similar result is also obtained by Lang for real p-form Klein-Gordon fields [Lan12, Proposition 4.5.6.], although his result is obtained by a longer and more abstract argument. The techniques used in the proof of the result are inspired by the techniques used by Lang and Dappiaggi, in particular [DL12, Proposition 3.1], to investigate the quantisation of electromagnetism in curved spacetime.

The geometrical techniques introduced here are a generalisation of the techniques introduced by Lang in his thesis [Lan12, Lemma 1.1.6]. Although his techniques are sufficient to prove our result for theories of unobservable fields, they are insufficient to prove Einstein causality for the universal algebras of "even theories" (see definition 3.1.6). Our techniques also have the advantage of being simpler and more widely applicable.

The results proven here also have applications to the next chapter, in which we investigate universal algebras formed from the even parts of algebras generated by Fermionic fields. In that chapter we construct central elements in the universal algebra, which then allow us to decompose our algebra into a product of subalgebras, one per element of $H^1(\mathcal{M}, \mathbb{Z}_2)$. This therefore links the universal algebra to the classification of spin structures on $\mathcal{M}$. In order to prove these elements are central however, we need to show that Einstein causality holds in the universal algebra, which we do in this chapter.

Our results apply to theories defined on globally hyperbolic spacetimes, and are heavily reliant on results in Lorentzian geometry which apply to this class of spacetimes. For technical reasons, outlined later and related to the geometric proofs, we restrict to theories on spacetimes whose Cauchy surfaces are of dimension three or higher. We discuss possible relaxations of these restrictions in the conclusions.

The outline of the remainder of this chapter is as follows: In Section 1 we outline the class of theories we are interested in and describe the universal algebra construction in more detail. In Section 2, we prove the geometrical results that will form the basis of the techniques used in the rest of the chapter. In section 3 we prove that universal theories of unobservable fields (see definition 3.1.5) are equivalent to the original theory they are built from. In the final section we prove Einstein causality for the universal theories of unobservable fields and "even theories" (see definition 3.1.6).

3.1 Preliminaries

In this section we will introduce the class of linear theories with hyperbolic equation of motion. We then take a general theory in this class and restrict it to contractible spacetimes, forgetting the global structure of the theory. Next we define and give a concrete characterisation of the universal algebra, and show how this can be used to extend the theory defined on contractible regions. We end the section by introducing what Einstein causality means in the context of the extended theories constructed in this section.

We consider a class of theories generalising the complex free scalar and Dirac fields defined on globally hyperbolic spacetimes. A concise description is provided which combines all such theories into a single theory on a category of bundles over globally hyperbolic spacetimes, equipped with Green hyperbolic operators. This strategy is adapted from [BG11]. Real hermitian free field theories may be described in a similar way but for simplicity of exposition we restrict to complex fields. We begin by introducing the following category of vector bundles.

Definition 3.1.1. *HVBundLoc is the category whose objects are given by: a Loc object $(\mathcal{M}, g)$ together with a hermitian vector bundle over $\mathcal{M}$, that is, a smooth vector bundle with finite-dimensional complex vector spaces as fibres, equipped with a smooth, nondegenerate, but possibly indefinite, hermitian form that is antlinear in its first argument. The fibre dimension is constant but arbitrary. Typically we denote an object of HVBundLoc by the pair $(\mathcal{M}, \mathcal{E})$ of its base and total space, leaving the projection $\pi : \mathcal{E} \to \mathcal{M}$, hermitian form $\langle , \rangle$ and other structures from the Loc object implicit.*

The morphisms of HVBundLoc are vector bundle morphisms that are fibrewise isometric isomorphisms that induce a Loc morphism on the base spaces. In particular, each $\Psi : (\mathcal{M}_1, \mathcal{E}_1) \to (\mathcal{M}_2, \mathcal{E}_2)$ induces a smooth map $\psi : \mathcal{M}_1 \to \mathcal{M}_2$, in the sense that $\pi_2 \circ \Psi = \psi \circ \pi_1$, and the map ψ is required to be a Loc morphism.

Part of the data required for the background structure of our theory, is a choice of linear differential operator acting on sections of a vector bundle. We now introduce some definitions which will be needed to describe the conditions we impose on these operators.

Definition 3.1.2. *A section $f : \mathcal{M} \to \mathcal{E}$ of a vector bundle $(\mathcal{M}, \mathcal{E}) \in$ HVBundLoc has spacelike compact support if there is some compact subset K of $\mathcal{M}$, such that* $\mathrm{supp}(f) \subset J(K)$. *We use $C_0^\infty(\mathcal{M}, \mathcal{E})$ and $C_{sc}^\infty(\mathcal{M}, \mathcal{E})$ to denote the set of sections with compact and spacelike compact support respectively.*

By integration using the volume form on $\mathcal{M}$, $C_0^\infty(\mathcal{M}, \mathcal{E})$ and $C_{sc}^\infty(\mathcal{M}, \mathcal{E})$ inherit a hermitian pairing from $\mathcal{E}$, denoted (f, g), defined when f and g have compactly intersecting supports.

It is worth highlighting that the pairings $\langle , \rangle$ and $(,)$ are between sections of the vector bundle, rather than being between a section of the vector bundle and a section of the dual vector bundle as in [BG11]. This is why our form is Hermitian rather than bilinear. We now generalise definition 2.3.1 to operators acting on sections of objects of HVBundLoc.

Definition 3.1.3. *Let P be a linear differential operator on sections of $(\mathcal{M}, \mathcal{E}) \in$ HVBundLoc. An advanced/retarded Green operator for P is a linear operator $E^{\mp} : C_0^\infty(\mathcal{M}, \mathcal{E}) \to C_{sc}^\infty(\mathcal{M}, \mathcal{E})$ which satisfies*

- $P \circ E^{\mp} = \mathcal{I}$, *where* $\mathcal{I}$ *is the inclusion map from* $C_0^\infty(\mathcal{M}, \mathcal{E})$ *to* $C_{sc}^\infty(\mathcal{M}, \mathcal{E})$.
- $E^{\mp} \circ P|_{C_0^\infty(\mathcal{M},\mathcal{E})} = \mathcal{I}$.
- $\mathrm{supp}(E^{\mp} f) \subset J^{\mp}(\mathrm{supp} f)$.

We use the notation $E := E^- - E^+$ *for the advanced minus retarded Green operator of* P *and use the notation* $E(f,g) := \int_{\mathcal{M}} \mathrm{dvol}_g \, \langle f, Eg \rangle$. *The operator* P *is said to be Green hyperbolic if it admits advanced and retarded Green operators (in which case the Green operators are unique). We say that* P *is formally self-adjoint if* $(f, Pg) = (Pf, g)$ *for all* $f, g \in C^\infty(\mathcal{M}, \mathcal{E})$ *with compactly intersecting supports.*

We can now define the background category which we will use to define the Bosonic and Fermionic linear theories.

Definition 3.1.4. *An object of* GlobHypGreen *consists of an object* $(\mathcal{M}, \mathcal{E}) \in$ HVBundLoc *and a formally self-adjoint Green hyperbolic operator* $P : C^\infty(\mathcal{M}, \mathcal{E}) \to C^\infty(\mathcal{M}, \mathcal{E})$. *A morphism in this category consists of a* HVBundLoc *morphism* $\Psi : (\mathcal{M}_1, \mathcal{E}_1, P_1) \to (\mathcal{M}_2, \mathcal{E}_2, P_2)$ *which satisfies*

$$P_1 \circ \Psi^* = \Psi^* \circ P_2 \quad \text{where} \quad \Psi^* f = \Psi^{-1} \circ f \circ \psi \quad .$$

The category GlobHypGreen differs from the category of the same name defined in [BG11], only in that the bundles in our category are complex vector bundles with hermitian form rather than real vector bundles with bilinear form. Now that we have our background category, we can define the theory (which defines a class of theories as will be explained below) which will be studied for the rest of the chapter.

Definition 3.1.5. *The linear Bosonic/Fermionic theory consists of a functor* $\mathscr{A}^{\mp}$: GlobHypGreen $\to$ Alg *which assigns to each object* $\mathbf{G} = (\mathcal{M}, \mathcal{E}, P)$ *of* GlobHypGreen*, the algebra generated by elements* $\mathcal{A}(f)_{\mathbf{G}}$ *with* $f \in C_0^\infty(\mathcal{M}, \mathcal{E})$, *and has the following relations imposed:*

- $\mathcal{A}(af + bg)_{\mathbf{G}} = a\mathcal{A}(f)_{\mathbf{G}} + b\mathcal{A}(g)_{\mathbf{G}}$ *for all* $a, b \in \mathbb{C}$.
- $\mathcal{A}(Pf)_{\mathbf{G}} = 0$.
- $[\mathcal{A}(f)_{\mathbf{G}}, \mathcal{A}(g)_{\mathbf{G}}]_{\mp} = 0$ *where* $[A, B]_{\mp} = AB \mp BA$.
- $[\mathcal{A}(g)^*_{\mathbf{G}}, \mathcal{A}(f)_{\mathbf{G}}]_{\mp} = iE(g, f) \, \mathbb{1}$.

The Alg *morphism* $\mathscr{A}^{\mp}(\Psi)$ *the theory associates to a* GlobHypGreen *morphism* $\Psi : \mathbf{G}_1 \to \mathbf{G}_2$, *is defined by its action on generators which is*

$$\mathscr{A}^{\mp}(\Psi)\mathcal{A}(f)_{\mathbf{G}_1} = \mathcal{A}(\Psi_* f)_{\mathbf{G}_2}$$

where Ψ_* *is the pushforward map which takes a section in* $C_0^\infty(\mathcal{M}_1, \mathcal{E}_1)$ *to a section in* $C_0^\infty(\mathcal{M}_2, \mathcal{E}_2)$ *and is defined as*

$$(\Psi_* f)(p) = \begin{cases} (\Psi \circ f \circ \psi^{-1})(p) & p \in \psi(\mathcal{M}_1) \\ 0 & \text{otherwise} \end{cases}$$

which is an extension by 0.

The theory $\mathscr{A}^{\mp}$ defined by definition 3.1.5 can be regarded as a class of theories, where each theory of the class is obtained by restricting $\mathscr{A}^{\mp}$ to a subcategory of $\mathsf{GlobHypGreen}$. For instance, if we want to model the theory of the free Dirac field, we can do so by defining a functor $\mathscr{D} : \mathsf{SpinLoc} \to \mathsf{GlobHypGreen}$ (where $\mathsf{SpinLoc}$ is defined in definition 2.3.4) and compose to get $\mathscr{A}^{+} \circ \mathscr{D}$ as our theory. This functor takes an object $(\mathcal{SM}, f) \in \mathsf{SpinLoc}$ to $(\mathcal{M}, \mathcal{DM}, P_{\mathcal{SM}}) \in \mathsf{GlobHypGreen}$ where $\mathcal{DM}$ is the spinor bundle defined in definition 2.3.5, and $P_{\mathcal{SM}}$ is the Dirac operator associated to $(\mathcal{SM}, f)$ defined in definition 2.3.6. The functor $\mathscr{D}$ takes a $\mathsf{SpinLoc}$ morphism χ to the bundle morphism χ_* defined in definition 2.3.6, which satisfies the conditions necessary to be regarded as a $\mathsf{GlobHypGreen}$ morphism.

We can see that $\mathscr{A}^{-}$ satisfies Einstein causality, given the fourth bullet point condition and the support properties of Ef given in definition 3.1.3, and similarly see that $\mathscr{A}^{+}$ satisfies graded Einstein causality, which means spacelike separated fields anti-commute. We also note that $\mathscr{A}^{+}$ assigns algebras to some objects of $\mathsf{GlobHypGreen}$ (in particular those objects which do not also belong to the category $\mathsf{GlobHypDef}$ defined in [BG11]) which do not admit states, meaning some of the theories obtained from $\mathscr{A}^{+}$ as outlined above describe ghost fields.

We will also be interested in "even theories", since these will be used in the next chapter to reconstruct spin structure information from the universal algebra. We therefore give the definition of the even subtheory of $\mathscr{A}^{\mp}$.

Definition 3.1.6. *The even linear Bosonic/Fermionic theory $\mathscr{A}^{\mp}_E$ assigns to each object $\mathbf{G}$ of $\mathsf{GlobHypGreen}$, the even subalgebra of $\mathscr{A}^{\mp}(\mathbf{G})$, which is the fixed point subalgebra with respect to the automorphism defined by $\mathcal{A}(f)_{\mathbf{G}} \mapsto -\mathcal{A}(f)_{\mathbf{G}}$. This subalgebra is generated by pairs of the form $\mathcal{A}(f)_{\mathbf{G}}\mathcal{A}(g)_{\mathbf{G}}$ or $\mathcal{A}(f)^*_{\mathbf{G}}\mathcal{A}(g)_{\mathbf{G}}$. The morphism $\mathscr{A}^{\mp}_E(\Psi)$ is defined by its action on generators*

$$\begin{aligned}\mathscr{A}^{\mp}_E(\Psi)\left(\mathcal{A}(f)_{\mathbf{G}}\mathcal{A}(g)_{\mathbf{G}}\right) &= \mathcal{A}(\Psi_* f)_{\mathbf{G}}\mathcal{A}(\Psi_* g)_{\mathbf{G}} \\ \mathscr{A}^{\mp}_E(\Psi)\left(\mathcal{A}(f)^*_{\mathbf{G}}\mathcal{A}(g)_{\mathbf{G}}\right) &= \mathcal{A}(\Psi_* f)^*_{\mathbf{G}}\mathcal{A}(\Psi_* g)_{\mathbf{G}} \quad .\end{aligned}$$

From now on to ease notation, we will drop the subscript on the algebra generators indicating the $\mathsf{GlobHypGreen}$ object the algebra corresponds to, opting to leave it implicit from context. For each of the theories in definitions 3.1.5 and 3.1.6, we would like to build up a *universal theory* associated to them. To do this we first need to introduce the following category.

Definition 3.1.7. *Given $\mathcal{M} \in \mathsf{Loc}$, we define the category $\mathsf{Loc}^{\mathcal{M}}$ whose objects are open and causally convex subsets of $\mathcal{M}$ equipped with the pulled back metric from $\mathcal{M}$. Morphisms in this category are the inclusion maps which embed regions of $\mathcal{M}$ into each other.*

We use $\mathsf{Loc}^{\mathcal{M}}_C$ to denote the full subcategory (see definition 1.4.3) of $\mathsf{Loc}^{\mathcal{M}}$ consisting of contractible spacetimes.

We now define what a net of local algebras is, which is needed to define *universal algebras.*

Definition 3.1.8. *A net of local algebras for a spacetime $\mathcal{M}$ is a functor $\mathscr{T}$ from a full subcategory of $\mathsf{Loc}^{\mathcal{M}}$ to $\mathsf{Alg}^{(h)}$.*

We use $\mathsf{Alg}^{(h)}$ in the definition of nets of local algebras in order to guarantee the existence of the associated *universal algebras* (defined later). We will be considering nets of local algebras from $\mathsf{Loc}_C^{\mathcal{M}}$, noting that bundles over contractible bases trivialise. The following family of functors will be useful for defining the nets of local algebras we will be analysing.

Definition 3.1.9. *Let $\mathbf{G} = (\mathcal{M}, \mathcal{E}, P)$ be any object of $\mathsf{GlobHypGreen}$, and let π be the bundle projection $\pi : \mathcal{E} \to \mathcal{M}$. The functor $I_{\mathbf{G}} : \mathsf{Loc}^{\mathcal{M}} \to \mathsf{GlobHypGreen}$ is defined to be the map which takes the object $\mathcal{O} \in \mathsf{Loc}^{\mathcal{M}}$ to the object $\left(\mathcal{O}, \pi^{-1}(\mathcal{O}), P|_{\mathcal{O}}\right) \in \mathsf{GlobHypGreen}$, where $\left(\mathcal{O}, \pi^{-1}(\mathcal{O})\right)$ is the bundle obtained by restricting $\mathcal{E}$ to $\mathcal{O}$. The functor $I_{\mathbf{G}}$ takes the $\mathsf{Loc}^{\mathcal{M}}$ morphism $\psi_{ij} : \mathcal{O}_i \to \mathcal{O}_j$ to the $\mathsf{GlobHypGreen}$ morphism $\Psi_{ij} : (\mathcal{O}_i, \pi^{-1}(\mathcal{O}_i), P|_{\mathcal{O}_i}) \to (\mathcal{O}_j, \pi^{-1}(\mathcal{O}_j), P|_{\mathcal{O}_j})$ which is the bundle inclusion map of $\pi^{-1}(\mathcal{O}_i)$ into $\pi^{-1}(\mathcal{O}_j)$.*

We now define the nets of local algebras analysed in this chapter.

Definition 3.1.10. *The net of local algebras $\mathscr{A}_{\mathbf{G}}$ is defined to be the composition of functors $F \circ \mathscr{A} \circ I_{\mathbf{G}} \circ \mathcal{I}$, where $\mathcal{I}$ is the inclusion functor $\mathcal{I} : \mathsf{Loc}_C^{\mathcal{M}} \to \mathsf{Loc}^{\mathcal{M}}$, $\mathscr{A}$ is a theory in the sense of definitions 3.1.5 and 3.1.6, and F is the forgetful functor from Alg to $\mathsf{Alg}^{(h)}$.*

The net of local algebras $\mathscr{A}_{\mathbf{G}}$ in definition 3.1.10 is the net of local algebras of open contractible causally convex subregions of the underlying spacetime of $\mathbf{G}$, with fields equations specified by the Green hyperbolic operator of $\mathbf{G}$. Given a net of local algebras, a global algebra can be obtained by the *universal algebra* construction of Fredenhagen [Fre90].

Definition 3.1.11. *Let C be a full subcategory of $\mathsf{Loc}^{\mathcal{M}}$. A cocone over a net of local algebras $\mathscr{T} : \mathsf{C} \to \mathsf{Alg}^{(h)}$ is a pair (A, h), where $A \in \mathsf{Alg}^{(h)}$ and h assigns to each $\mathcal{O} \in \mathsf{C}$ an $\mathsf{Alg}^{(h)}$ morphism $h_{\mathcal{O}} : \mathscr{T}(\mathcal{O}) \to A$, such that $h_{\mathcal{O}_i} = h_{\mathcal{O}_j} \circ \mathscr{T}(\psi_{ij})$ for every C morphism $\psi_{ij} : \mathcal{O}_i \to \mathcal{O}_j$.*

*The universal algebra for this net **(if it exists)**, is a cocone $(\mathscr{U}[\mathscr{T}], \phi)$ satisfying the following universal property: given any other cocone (A, h), there exists a unique $\mathsf{Alg}^{(h)}$ morphism $H : \mathscr{U}[\mathscr{T}] \to A$ such that the following diagram commutes for all $\psi_{ij} : \mathcal{O}_i \to \mathcal{O}_j$*

$$\begin{tikzcd}
\mathscr{T}(\mathcal{O}_i) \arrow[rr, "\mathscr{T}(\psi_{ij})"] \arrow[dr, "\phi_{\mathcal{O}_i}"] \arrow[ddr, "h_{\mathcal{O}_i}"'] & & \mathscr{T}(\mathcal{O}_j) \arrow[dl, "\phi_{\mathcal{O}_j}"'] \arrow[ddl, "h_{\mathcal{O}_j}"] \\
& \mathscr{U}[\mathscr{T}] \arrow[d, dashed, "H"'] & \\
& A &
\end{tikzcd}.$$

In categorical terms the universal algebra is the colimit of the net of local algebras. This colimit exists since the category $\mathsf{Alg}^{(h)}$ is cocomplete [Lan12, Theorem 2.2.10]. This is why we used the category $\mathsf{Alg}^{(h)}$ in definition 3.1.8, it guarantees the existence of the universal algebra $\mathscr{U}[\mathscr{T}]$ associated to $\mathscr{T}$. The universal property implies that the universal algebra is unique up to isomorphism. We now provide a concrete model for the universal algebra which will be used for calculations in subsequent sections.

Proposition 3.1.12. *For a complex vector space V, let $\mathcal{T}(V)$ denote the associative tensor algebra of V. The algebra $\mathcal{T}(V)$ has componentwise addition, componentwise multiplication with a scalar, componentwise antilinear involution $*$ and multiplication induced by the algebraic tensor product.*

A model for the universal algebra for the net of local algebras $\mathscr{A}_{\mathbf{G}}^{\mp}$ is given by the algebra

$$\mathcal{T}\left(\bigoplus_{\mathcal{O}\in\mathsf{Loc}_C^{\mathcal{M}}} U\left(\mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O})\right)\right)/\mathcal{I} \tag{3.1}$$

where U is the forgetful functor from $\mathsf{Alg}^{(h)}$ to the category of complex vector spaces, and $\mathcal{I}$ is the two-sided $*$-ideal generated by elements of the form

$$\left(0, -(AB)^{(\mathcal{O})}, A^{(\mathcal{O})}\otimes B^{(\mathcal{O})}, 0, \dots\right) \quad \forall A, B \in \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O}),$$
$$\left(1, -\mathbb{1}_{\mathcal{O}}^{(\mathcal{O})}, 0, \dots\right) \quad \textit{where } \mathbb{1}_{\mathcal{O}} \textit{ is the unit element of } \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O}),$$
$$\left(0, \left(\mathscr{A}_{\mathbf{G}}^{\mp}(\psi_{ij})(A)\right)^{(\mathcal{O}_j)} - A^{(\mathcal{O}_i)}, 0, \dots\right) \; \forall A \in \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O}_i) \textit{ and all } \mathsf{Loc}_C^{\mathcal{M}} \textit{ morphisms } \psi_{ij} : \mathcal{O}_i \to \mathcal{O}_j,$$

where for a local algebra element $A \in \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O})$, $A^{(\mathcal{O})}$ denotes the vector in $\bigoplus_{\mathcal{O}\in\mathsf{Loc}_C^{\mathcal{M}}} \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O})$ which has A in the entry indexed by $\mathcal{O}$ and zero in all other entries.

The morphisms $\phi_{\mathcal{O}}$ from the local algebras into the universal algebra that characterise the colimit are given by

$$\phi_{\mathcal{O}}(A) = \left(0, A^{(\mathcal{O})}, 0, \dots\right) \quad \forall A \in \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O}) \quad .$$

Proof. See [DL12, Theorem 3.1]. □

We see from proposition 3.1.12 that the universal algebra $\mathscr{U}^{\mp}(\mathbf{G}) := \mathscr{U}[\mathscr{A}_{\mathbf{G}}^{\mp}]$ is generated by the identity element together with elements $\mathcal{A}_{\mathcal{O}}(f) := \phi_{\mathcal{O}}(\mathcal{A}(f))$, with $\mathcal{O} \in \mathsf{Loc}_C^{\mathcal{M}}$ and $\mathcal{A}(f) \in \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O})$. We also see from the ideal $\mathcal{I}$ in equation (3.1) that the following facts hold: generators indexed by a fixed $\mathsf{Loc}_C^{\mathcal{M}}$ region $\mathcal{O}$ satisfy linearity and equation of motion relations, due to the first line of elements that generate $\mathcal{I}$ which impose local algebra relations. Two generators indexed by the same section $\mathcal{A}_{\mathcal{O}_1}(f)$ and $\mathcal{A}_{\mathcal{O}_2}(f)$ (with $\mathrm{supp} f \subset \mathcal{O}_1 \cap \mathcal{O}_2$) are equal if there is a $\mathsf{Loc}_C^{\mathcal{M}}$ morphism between the sub-indices i.e, $\mathcal{O}_1 \subset \mathcal{O}_2$ or vice-versa, due to the third line of elements that generate $\mathcal{I}$. Commutation/anti-commutation relations can be applied if the sub-indices of the generators agree or can be embedded in a common $\mathsf{Loc}_C^{\mathcal{M}}$ region. However, this is not a necessary condition for being able to apply commutation/anti-commutation relations; we can for example use linearity to split generators up into pieces with sections supported in smaller $\mathsf{Loc}_C^{\mathcal{M}}$ regions, resulting in a sum of commutators/anti-commutators which can be evaluated in the usual way. This is the main strategy that we adopt in this chapter to prove Einstein causality in universal algebras.

The universal algebra $\mathscr{U}_E^{\mp}(\mathbf{G}) := \mathscr{U}[\mathscr{A}_{E\,\mathbf{G}}^{\mp}]$ for even algebras is given by the same model as in proposition 3.1.12, except the net of local algebras $\mathscr{A}_{\mathbf{G}}^{\mp}$ is replaced with $\mathscr{A}_{E\,\mathbf{G}}^{\mp}$. Therefore, the elements $\mathcal{A}_{\mathcal{O}}^{\flat}(f,g) := \phi_{\mathcal{O}}(\mathcal{A}(f)\mathcal{A}(g))$ and $\mathcal{A}_{\mathcal{O}}^{\sharp}(f,g) := \phi_{\mathcal{O}}(\mathcal{A}(f)^*\mathcal{A}(g))$ generate $\mathscr{U}_E^{\mp}(\mathbf{G})$ (we omit the superscript $\sharp$ or $\flat$ for equations valid for both types of generator). Again, commutation/anti-commutation relations can be applied in this algebra if (but not only if) the sub-indices of the generators agree.

Given a $\mathsf{GlobHypGreen}$ morphism Ψ between $\mathbf{G}_1$ and $\mathbf{G}_2$, we get $\mathsf{Alg}^{(h)}$ morphisms $\mathscr{U}^{\mp}(\Psi) : \mathscr{U}^{\mp}(\mathbf{G}_1) \to \mathscr{U}^{\mp}(\mathbf{G}_2)$ and $\mathscr{U}^{\mp}_E(\Psi) : \mathscr{U}^{\mp}_E(\mathbf{G}_1) \to \mathscr{U}^{\mp}_E(\mathbf{G}_2)$ defined by their action on generators

$$\mathscr{U}^{\mp}(\Psi)\mathcal{A}_{\mathcal{O}}(f) = \mathcal{A}_{\psi(\mathcal{O})}(\Psi_* f) \tag{3.2}$$

$$\mathscr{U}^{\mp}_E(\Psi)\mathcal{A}_{\mathcal{O}}(f,g) = \mathcal{A}_{\psi(\mathcal{O})}(\Psi_* f, \Psi_* g) \quad . \tag{3.3}$$

This gives us functors $\mathscr{U}^{\mp}, \mathscr{U}^{\mp}_E : \mathsf{GlobHypGreen} \to \mathsf{Alg}^{(h)}$ which we regard as extensions of $\mathscr{A}^{\mp}|_{\mathsf{Loc}_C}$ and $\mathscr{A}^{\mp}_E|_{\mathsf{Loc}_C}$. This construction is a special case of a more general construction known in category theory as a left Kan extension; in this case of the functor $\mathscr{A} \circ I$ along the functor $I : \mathsf{GlobHypGreen}_C \to \mathsf{GlobHypGreen}$, where $\mathsf{GlobHypGreen}_C$ is the subcategory of $\mathsf{GlobHypGreen}$ consisting of objects with contractible base space $\mathcal{M}$, and I is the inclusion functor. For more details on the left Kan extension, see section 2.2 of [Lan12].

The main aim of this chapter is to prove that $\mathscr{U}^-$ and $\mathscr{U}^{\mp}_E$ satisfy Einstein causality, which means that for any $\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3 \in \mathsf{GlobHypGreen}$ with morphisms $\Psi_{13} : \mathbf{G}_1 \to \mathbf{G}_3$ and $\Psi_{23} : \mathbf{G}_2 \to \mathbf{G}_3$, if the base spaces of $\mathbf{G}_1$ and $\mathbf{G}_2$ are causally disjoint when embedded in the base space of $\mathbf{G}_3$, then the images of $\mathscr{U}(\Psi_{12})$ and $\mathscr{U}(\Psi_{23})$ in $\mathscr{U}(\mathbf{G}_3)$ commute.

3.2 Main geometric results

In the previous section we introduced the generators for the universal algebra and noted that commutation/anti-commutation relations could only be applied directly if the sub-indices of the generators agreed. Therefore the main impediment to proving Einstein causality in the universal algebra arises from spacelike separated $\mathsf{Loc}^{\mathcal{M}}_C$ regions that cannot be embedded in a common $\mathsf{Loc}^{\mathcal{M}}_C$ region. We therefore establish some results in Lorentzian geometry which will allow us to circumvent this problem. We begin with a definition.

Definition 3.2.1. *A Cauchy ball of a manifold Σ is an open subset $B \subset \Sigma$, such that there exists a chart (U, ψ) of Σ with $\overline{B} \subset U$ and $\psi(B)$ an open relatively compact ball centred at the origin. If Σ is an acausal spacelike hypersurface, then we use the term Cauchy diamond to denote the Cauchy development of a Cauchy ball.*

Cauchy balls of acausal spacelike hypersurfaces have the useful property that their corresponding Cauchy diamonds are objects of $\mathsf{Loc}^{\mathcal{M}}_C$, and using the results of [BS06], they can be extended to Cauchy surfaces of $\mathcal{M}$.

Our main strategy for proving Einstein causality will involve using partitions of unity to split up generators into parts which are each supported in Cauchy diamonds. We then need a way of embedding disjoint Cauchy diamonds into a larger object of $\mathsf{Loc}^{\mathcal{M}}_C$. Our method for doing so roughly corresponds to connecting regions by thin tubes in such a way that the combined region will remain contractible. The motivation for considering Cauchy balls comes from the fact that in general it may not be possible to connect two general disjoint contractible subsets of a manifold with a single contractible subset. We will however prove that disjoint Cauchy balls can be connected by a Cauchy ball in proposition 3.2.7. We make use of the fact that Cauchy balls come with coordinate systems, in order to control how a connecting tube enters each ball in such a way as to not spoil the contractibility of the connected region. We first introduce a couple of definitions.

Definition 3.2.2. *Let S be a submanifold of a Riemannian manifold X with metric g. A normal vector to S at p, is a tangent vector $n \in T_pX$ such that $g_p(n,v) = 0$ for all $v \in T_pS \subset T_pX$. The set of all normal vectors to S at p is denoted by N_pS. The normal bundle of S as a set is then defined to be*

$$NS = \bigsqcup_{p \in S} N_p$$

which is a subset of $T\mathcal{M}$ and can therefore be given the relative topology inherited from $T\mathcal{M}$. The atlas $\{(V_i, \phi_i)\}$ of S defines an atlas $\{(NV_i, \widetilde{\phi}_i)\}$ of NS where $NV_i = \bigsqcup_{p\in V_i} N_p$ and $\widetilde{\phi}_i : NV_i \to \mathbb{R}^n \times \mathbb{R}^n$ is given by

$$\widetilde{\phi}_i([p,i,v]) = (\phi_i(p), v) \quad .$$

Definition 3.2.3. *Let NS be the normal bundle of a submanifold S in a Riemannian manifold (X, g). The normal exponential map $\exp_\perp : NS \to X$ sends $(p,v) \in NS$ to $\gamma(1)$, where γ is the (X,g) geodesic such that $\gamma(0) = p$ and $\dot\gamma(0) = v$. A tubular neighbourhood of S is the diffeomorphic image under $\exp_\perp$ of a neighbourhood of the zero section in NS.*

We now establish a couple of lemmas which will then be used to prove a proposition which makes the notion of "connecting regions by thin tubes in such a way that the combined region will remain contractible" more precise and shows that it is always possible.

Lemma 3.2.4. *Let Σ be an n-manifold and $\gamma : (0,1) \to \Sigma$ be a smooth embedded curve, so that the image of γ is a contractible submanifold of Σ. Given tubular neighbourhoods Γ of γ and Γ' of γ restricted to a connected subset of $(0,1)$, such that $\overline{\Gamma'} \subset \Gamma$, there exists a region B such that $\Gamma' \subset B \subset \overline{B} \subset \Gamma$ and coordinates on Γ in which B is an open ball.*

Proof. We may choose normal coordinates $(x, x_\perp) \in \mathbb{R} \times \mathbb{R}^{n-1}$ in which Γ is given by $(-1,1) \times \mathrm{Ball}(1)$ where $\mathrm{Ball}(r)$ denotes a ball in $\mathbb{R}^{n-1}$ of radius r centred on the origin. By supposition $\overline{\Gamma'} \subset \Gamma$, hence we can find $a, r < 1$ such that $\Gamma' \subset (-a,a) \times \mathrm{Ball}(r)$. We choose $a_* \in (a,1)$ and $r_* \in (r,1)$ and set

$$b = \frac{r_* \tan(\pi a_*/2)}{\sqrt{r_*^2 - r^2}}$$

and then define B to be the set

$$B = \left\{ (x, x_\perp) \in \mathbb{R} \times \mathbb{R}^{n-1} \;\Big|\; \frac{\|x_\perp\|^2}{r_*^2} + \frac{\tan^2(\pi x/2)}{b^2} < 1 \right\}$$

which is an open ball of radius r_* in coordinates $(x', x_\perp)$ where $x' = \frac{r_*}{b}\tan(\pi x/2)$. We now consider a point $(x, x_\perp) \in \Gamma'$, which means $|x| < a < 1$ and $\|x_\perp\| < r$, and we get

$$\frac{\|x_\perp\|^2}{r_*^2} + \frac{\tan^2(\pi x/2)}{b^2} < \frac{r^2}{r_*^2} + \frac{\tan^2(\pi a/2)}{b^2} = 1 + \frac{r^2 - r_*^2}{r_*^2} + \frac{r_*^2 - r^2}{r_*^2}\frac{\tan^2(\pi a/2)}{\tan^2(\pi a_*/2)}$$
$$= 1 + \frac{r_*^2 - r^2}{r_*^2}\left(-1 + \frac{\tan^2(\pi a/2)}{\tan^2(\pi a_*/2)}\right) < 1$$

where the final inequality is due to the fact that $a < a_*$ and $r < r_*$, making the second term negative. This inequality defines points of B, hence $\Gamma' \subset B$. We can also see that any point in $\overline{B}$ must satisfy

$$\frac{\|x_\perp\|^2}{r_*^2} \leq 1 \quad \text{and} \quad \frac{\tan^2(\pi x/2)}{b^2} \leq 1$$

which implies $\|x_\perp\| \leq r_* < 1$ and $|x| < 1$ since tan diverges at $\pm\pi/2$. This implies $\overline{B} \subset \Gamma$, and therefore B satisfies the conditions of the lemma. □

Lemma 3.2.5. *Let Σ be a connected n dimensional manifold with $n > 1$. Let $\{B_n\}_{n=1}^N$ be a finite collection of Cauchy balls of Σ with disjoint closures, then $\Sigma \backslash \bigcup_{n \leq N} \overline{B_n}$ is a connected manifold.*

Proof. The region $\Sigma \backslash \bigcup_{n \leq N} \overline{B_n}$ is open since $\bigcup_{n \leq N} \overline{B_n}$ is closed, so it inherits a manifold structure from Σ. Since each B_n is a Cauchy ball, there exist Cauchy balls $\beta_n \supset \overline{B_n}$ because the closure of B_n is contained in the chart in which it is a ball. Moreover, we can chose the β_n to all be disjoint from each other, because Σ is a paracompact Hausdorff space, and therefore disjoint closed subsets of Σ have disjoint open neighbourhoods.

The region $\beta_1 \cap \Sigma \backslash \overline{B_1} = \beta_1 \backslash \overline{B_1}$ is homeomorphic to an n dimensional anulus, which is connected when $n > 1$. We can therefore use the last part of the reduced Mayer-Vietoris sequence

$$\widetilde{H}_0(\beta_1 \backslash \overline{B_1}) \to \widetilde{H}_0(\Sigma \backslash \overline{B_1}) \oplus \widetilde{H}_0(\beta_1) \to \widetilde{H}_0(\Sigma)$$

which is exact, to argue that $\Sigma \backslash \overline{B_1}$ is connected since $\widetilde{H}_0(\beta_1 \backslash \overline{B_1}) = \widetilde{H}_0(\beta_1) = \widetilde{H}_0(\Sigma) = \mathbb{1}$. Since β_2 is disjoint from $\overline{B_1}$, $\beta_2 \cap \Sigma \backslash (\overline{B_1} \cup \overline{B_2}) = \beta_2 \backslash \overline{B_2}$ is also homeomorphic to an anulus. We can therefore repeat the above Mayer-Vietoris sequence argument with the replacements $\beta_1 \mapsto \beta_2$, $\overline{B_1} \mapsto \overline{B_2}$ and $\Sigma \mapsto \Sigma \backslash \overline{B_1}$, to show that $\Sigma \backslash (\overline{B_1} \cup \overline{B_2})$ is connected. In this way we can inductively apply the Mayer-Vietoris sequence argument above to show that $\Sigma \backslash \bigcup_{n \leq N} \overline{B_n}$ is connected. □

Proposition 3.2.6. *Let Σ be a connected n dimensional manifold with $n \neq 2$. Let B_1 and B_2 be Cauchy balls of Σ, and let $d_1 : [0,1] \to \Sigma$ and $d_2 : [0,1] \to \Sigma$ be smooth embedded curves such that for $i \in \{1,2\}$, $\mathrm{imag}(d_i) \subset \overline{B_i}$ and d_i has end-points on the boundary of $\overline{B_i}$. For the $n = 1$ case, we additionally require that d_1 and d_2 have the same orientation. Then there exists a smooth curve $P : [0,1] \to \Sigma$ such that*

- *$P|_{(0,1)} \in \Sigma \backslash (\overline{B_1} \cup \overline{B_2})$.*
- *$P(0) = d_1(1)$, $P(1) = d_2(0)$, and $d_2 * (P * d_1)$ (see definition 1.1.9 for path multiplication) is a smooth embedded curve.*

Proof. If $n = 1$, then Σ must be $\mathbb{R}$ or S^1 since it is connected. It is then clear that such a P exists, since we have required the orientations of d_1 and d_2 to match in the $n = 1$ case, which ensures that we can choose P to have the same orientation as d_1 and d_2 so that $d_2 * P * d_1$ is a smooth embedded curve.

We now focus on the $n > 1$ cases. We see from lemma 3.2.5 that $\Sigma \backslash (\overline{B_1} \cup \overline{B_2})$ is a connected manifold, hence it is path connected, and thus there is a continuous curve

$P : [0,1] \to \Sigma$ such that: $P(0) = d_1(1)$, $P(1) = d_2(0)$ and $P|_{(0,1)} \in \Sigma \backslash (\overline{B_1} \cup \overline{B_2})$. The path $\gamma := d_2 * (P * d_1)$ applied to $[1/4, 1/2]$ yields the image of P. Let L denote $[0,1] \backslash (1/4, 1/2)$. Any map $\gamma' \in C^0([0,1], \Sigma)_{\gamma|_L}$ (see definition 1.2.13) can be expressed as $d_2 * (P' * d_1)$ for some $P' : [0,1] \to \Sigma$.

Since $\Sigma \backslash (\overline{B_1} \cup \overline{B_2})$ is open, there exists a neighbourhood $\mathcal{N} \subset C^0([0,1], \Sigma)_{\gamma|_L}$ of γ such that for all $\gamma' \in \mathcal{N}$, $\gamma'|_{(1/4,1/2)} \subset \Sigma \backslash (\overline{B_1} \cup \overline{B_2})$. If $\gamma' \in \mathcal{N}$ is a smooth embedding, then its associated P' (see previous paragraph) satisfies the the bullet point conditions of the proposition. The conditions of theorem 1.2.14 are met for the pair (γ, L), since $\gamma|_L$ consists of the segments given by d_1 and d_2, which are both embeddings and therefore self-transverse (see definition 1.2.11) immersions. If $n > 2$, $\dim(\Sigma) \geq 2\dim([0,1]) + 1$ hence we can use corollary 1.2.16 to find such a smooth embedding $\gamma' \in \mathcal{N}$.

□

We believe the result in proposition 3.2.6 also holds in the case that the manifold Σ is 2-dimensional, and we provide the following sketch proof.

Sketch Proof. Since $\dim(\Sigma) = 2$ by supposition, $\dim(\Sigma) \geq 2\dim([0,1])$ so we instead apply theorem 1.2.14 to find $\gamma' \in \mathcal{N}$ such that $\gamma'|_{(0,1)}$ is a smooth self-transverse immersion, and apply proposition 1.2.15 to conclude that γ' has finitely many double points. This implies that the map P' associated to γ' has finitely many double points. We can therefore obtain an injective map P'' which is a piecewise smooth immersion connecting d_1 and d_2, by removing the segments from P' that form closed loops at its double points. We then need to smooth out the kinks at the image of the former double points where the closed loops were removed. This can be done by interpolating between the two segments that meet at a kink using a bump function. Presumably, if the interpolations are done in sufficiently small neighbourhoods of each kink, the resulting P''' remains injective, although we have not been able to obtain a formal proof of this fact. Assuming that this can be done, $d_2 * (P''' * d_1)$ is an embedding, and therefore P''' satisfies the conditions of proposition 3.2.6. □

Proposition 3.2.7. *Let Σ be a connected manifold, then any finite union of Cauchy balls $\{B_n\}_{n=1}^N$ in Σ with disjoint closures is contained in a Cauchy ball of Σ.*

Proof. To prove the proposition we will show that B_1 and B_2 can be contained in a Cauchy ball of Σ whose closure is disjoint from $\bigcup_{2<n\leq N} \overline{B_n}$. It then follows by induction that all of the B_n can be contained in a common Cauchy ball of Σ. We begin by constructing a metric on Σ which "behaves well" near B_1 and B_2, so that the tubular neighbourhood of a segment connecting B_1 and B_2 constructed from this metric enters B_1 and B_2 in a controlled way.

The Cauchy balls B_1 and B_2 are balls in the charts (U_1, ψ_1) and (U_2, ψ_2) respectively. By supposition, the closures of the B_1 and B_2 are disjoint, so U_1 and U_2 can be chosen to be disjoint from each other since $\mathcal{M}$ is a paracompact Hausdorff space and therefore normal. Throughout this proof we will use the convention that a sub-index denoted by i belongs to the index set $\{1, 2\}$, and a sub-index denoted by n belongs to the index set $\{1, \ldots, N\}$. The definition of Cauchy ball requires $\overline{B_n} \subset U_n$, hence we can find slightly larger Cauchy balls B'_n and β_n such that $\overline{B_n} \subset B'_n$ and $\overline{B'_n} \subset \beta_n \subset U_n$.

We consider the open manifold $\widetilde{\Sigma} := \Sigma \backslash \left(\overline{\beta_1} \cup \overline{\beta_2}\right)$ and find an open cover $\mathcal{U}$ of $\widetilde{\Sigma}$ by charts $\{(V_m, \phi_m)\}$. Since $\widetilde{\Sigma}$ is an open subset of Σ, $\mathcal{U} \cup \{U_1, U_2\}$ is an open cover of Σ. We now use a partition of unity $\{\rho_\alpha\}$ subordinate to this cover to construct a Riemannian metric g for Σ

$$g = \sum_\alpha \rho_\alpha \; \delta_{ij} \, dx^i_\alpha \otimes dx^j_\alpha$$

where x_α are the coordinates in the chart with index α, which could be one of the (U_i, ψ_i) or (V_m, ϕ_m) charts. Paracompactness ensures that there are only finitely many nonzero terms in this sum for each point of Σ. Inside $\overline{\beta_1} \cup \overline{\beta_2}$, the sum over α for g collapses to a single term, since U_i is the only set in the cover that contains points of $\overline{\beta_i}$ due to the fact that $\overline{\beta_i} \cap V_m = \emptyset$ and $U_1 \cap U_2 = \emptyset$. This implies that ψ_i is an isometry between the metric spaces (β_i, D_g) and $(\psi_i(\beta_i), D_E)$, where D_g is the distance function associated to g and D_E is the usual Euclidean metric. We use R_i and r_i to denote the radii of $\psi_i(\beta_i)$ and $\psi_i(B'_i)$ in the D_E metric.

We now consider a pair of smooth paths $d_i : [0,1] \to \overline{\beta_i}$ which trace out diameters of the Cauchy balls so that $\psi_i \circ d_i$ is a diameter of $\psi_i(\overline{\beta_i})$. The manifold $\Sigma' := \Sigma \backslash \bigcup_{2<n\leq N} \overline{B_n}$ is connected by lemma 3.2.5, hence we can apply proposition 3.2.6 to this manifold to show that there exists a smooth embedded curve P in Σ' that connects d_1 and d_2 i.e,

$$P(0) = d_1(1), \quad P(1) = d_2(0), \quad \text{and} \quad P|_{(0,1)} \subset \Sigma'$$

such that $\gamma := d_2 * (P * d_1)$ which connects the diameters of the Cauchy balls is a smooth embedded curve.

Since $\text{imag}(\gamma) \cap \overline{\Sigma'}$ and $\overline{B'_1} \cup \overline{B'_2}$ are compact, D_g applied to $(\text{imag}(\gamma) \cap \overline{\Sigma'}) \times (\overline{B'_1} \cup \overline{B'_2})$ has a minimum. Since $\text{imag}(\gamma) \cap \overline{\Sigma'}$ and $\overline{B'_1} \cup \overline{B'_2}$ are disjoint, the distance between them, denoted by δ_1, must be strictly greater than zero, hence the minimum distance between them must also be greater than zero. Similarly, the minimum distance between $\text{imag}(\gamma)$ and $\bigcup_{2<n\leq N} \overline{\beta_n}$, denoted by δ_2, is greater than zero since they are both compact and disjoint.

Since γ is an embedding, its image is a submanifold [Hir76, Theorem 1.3.1], so we can use the normal exponential map $\exp_\perp$ (see definition 3.2.3) defined by the metric g to find a tubular neighbourhood of $\gamma|_{(0,1)}$ [Hir76, Theorem 4.5.2]. By applying $\exp_\perp$ to a sufficiently small neighbourhood of the zero section of the normal bundle of γ, we get a tubular neighbourhood Γ for $\gamma|_{(0,1)}$ with fibres of length less than δ_1 and δ_2. Having fibres of length less than δ_1 ensures that fibres of points of γ outside β_i cannot enter B'_i, and having fibres of length less than δ_2 ensures none of the fibres of Γ enter $\bigcup_{2<n\leq N} \overline{\beta_n}$.

We now show that the points of γ whose fibres intersect B'_i are precisely those in $\gamma \cap B'_i$. Working in the chart (U_i, ψ_i), the fibres of points of γ in β_i have an initial segment which is a straight line perpendicular to diameter of $\psi_i(\beta_i)$ traced out by $\psi_i \circ d_i$, since ψ_i is an isometry between (β_i, D_g) and $(\psi_i(\beta_i), D_E)$. Therefore, in order for a fibre of $\beta_i \backslash B'_i$ to intersect B'_i, it cannot be entirely contained in β_i, it must have a segment which extends from the boundary of β_i to the boundary of B'_i, and this segment will have a length of at least $R_i - r_i$. The situation is shown in figure 3.1. The tubular neighbourhood $\widetilde{\Gamma} \subset \Gamma$ with fibres of length less that $\epsilon < \min\{\delta_1, \delta_2, R_1 - r_1, R_2 - r_2\}$, therefore has the property that

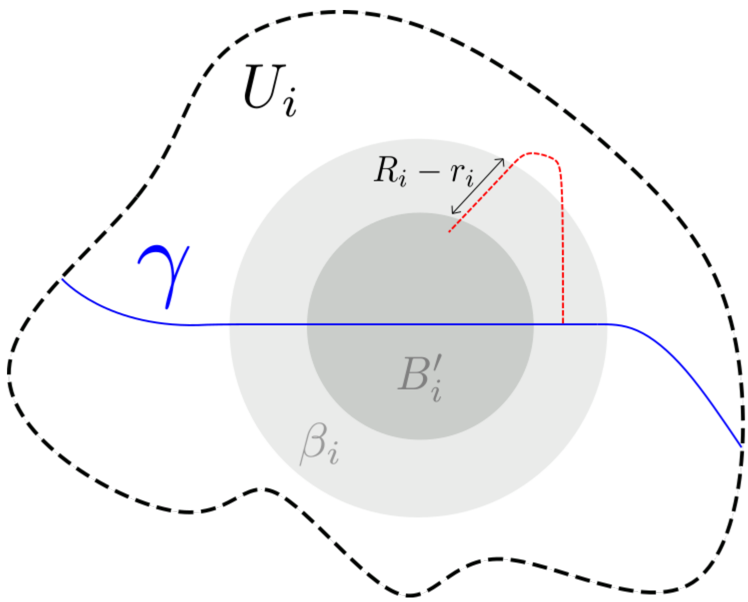

Figure 3.1: Fibre of a point $p \in \beta_i \backslash B_i'$ entering B_i'.

only points of γ inside each B_i' can have fibres which intersect B_i'. Having fibres of length less than δ_2 ensures that the closure of $\widetilde{\Gamma}$ is disjoint from $\bigcup_{2<n\leq N} \overline{\beta_n}$.

The region $\hat{\Gamma} := \widetilde{\Gamma} \cup B_1' \cup B_2'$ is the image of $\exp_\perp$ applied to $\mathcal{N} := \exp_\perp^{-1}(\widetilde{\Gamma} \cup B_1' \cup B_2')$. In the (U_i, ψ_i) coordinates, the geodesics between points in $\overline{\beta_i}$ are just straight lines, and γ traverses the diameter of $\overline{\beta_i}$. Therefore, the restriction of $\exp_\perp$ to $\widetilde{\mathcal{N}} := \exp_\perp^{-1}(\overline{B_1'} \cup \overline{B_2'})$ is a smooth immersion which is also injective, since every point in a sphere is uniquely specified by a position on its diameter and a vector orthogonal to the diameter. Moreover, $\widetilde{\mathcal{N}}$ is compact since $\overline{B_i'}$ is compact and Σ is Hausdorff, which implies $\exp_\perp|_{\widetilde{\mathcal{N}}}$ has smooth inverse since it is a smooth immersion on a compact domain. We also see that $\exp_\perp$ is injective on the whole of $\mathcal{N}$, since we have arranged $\widetilde{\Gamma}$ so that only points in $\gamma \cap B_i'$ can have fibres which can intersect B_i'. We therefore see that $\exp_\perp|_{\mathcal{N}}$ is a smooth injective map with smooth inverse, and is therefore a diffeomorphism, hence $\hat{\Gamma}$ is a tubular neighbourhood of $\gamma|_{(0,1)}$. This tubular neighbourhood still retains the property that its closure is disjoint from $\bigcup_{2<n\leq N} \overline{\beta_n}$, since $\overline{B_1'}$ and $\overline{B_2'}$ are disjoint from $\bigcup_{2<n\leq N} \overline{\beta_n}$.

We will use γ' to denote the segment of γ that has the initial segment connecting B_1 to β_1, and the final segment connecting B_2 to β_2 removed. This ensures the closure of $\text{imag}(\gamma')$ is contained in $\hat{\Gamma}$. We obtain a tubular neighbourhood Γ' of $\gamma'|_{(0,1)}$, by applying $\exp_\perp$ to the subset $\mathcal{N}' \subset \mathcal{N}$ that only contains fibres of γ', and these fibres differ from those in $\mathcal{N}$ by a constant scale factor λ, where $\sup_{i\in\{1,2\}}(\widetilde{r}_i/r_i) < \lambda < 1$ and $\widetilde{r}_i$ is the radius of $\psi_i(B_i)$ in the D_E metric. The lower bound on λ implies $B_1 \cup B_2 \subset \Gamma'$ and the upper bound implies $\overline{\Gamma'} \subset \hat{\Gamma}$.

We can then use lemma 3.2.4 to find a region $\mathcal{B}$ such that $\Gamma' \subset \mathcal{B} \subset \overline{\mathcal{B}} \subset \hat{\Gamma}$ and a chart $(\hat{\Gamma}, \hat{\psi})$ such that $\mathcal{B}$ is a ball in this chart. We therefore have a Cauchy ball $\mathcal{B}$ of Σ which contains B_1 and B_2, and has closure which is disjoint from $\bigcup_{2<n\leq N} \overline{\beta_n} \supset \bigcup_{2<n\leq N} \overline{B_n}$. We can therefore continue by induction to find a Cauchy ball of Σ that contains all of the B_n.

□

Using proposition 3.2.7, we can establish the following useful result.

Lemma 3.2.8. *Given a connected manifold Σ and a compact subset $\Sigma_K \subset \Sigma$, there exists a finite cover of Σ_K by Cauchy balls $\{B_i\}$ of Σ such that any two of the Cauchy balls in the cover can be contained in a larger Cauchy ball.*

Proof. The manifold Σ can be equipped with a Riemannian metric to give it the structure of a Riemannian manifold. Every point in a Riemannian manifold has a geodesically convex neighbourhood with respect to this metric [Mil63a, lemma 10.3], so we can cover Σ_K by a set of geodesically convex Cauchy balls $\{B_{r_i}(\sigma_i)\}$ of radius r_i centred at σ_i. As Σ_K is compact, there exists a Lebesgue number $\delta > 0$ for this cover, which means for all $\sigma \in \Sigma_K$ there exists some j such that $B_\delta(\sigma) \subset B_{r_j}(\sigma_j)$. Since $B_{r_j}(\sigma_j)$ is geodesically convex, any two points in it are connected by a unique geodesic, this implies the exponential map $\exp_\sigma : B_\delta(0) \subset T_\sigma\Sigma \to B_\delta(\sigma)$ is injective for all $\sigma \in \Sigma_K$. Therefore, every point of Σ_K has an injectivity radius greater than $\delta > 0$, which in turn implies $B_\epsilon(\sigma)$ is a Cauchy ball of Σ for all $\sigma \in \Sigma_K$ and $\epsilon \in (0, \delta)$.

We now consider the cover $\{B_{\delta/2}(\widetilde{\sigma}_i)\}$ of Σ_K by Cauchy balls of radius $\delta/2$. If any two members of this cover intersect, we can pick some $\widetilde{\sigma}_{ij} \subset B_{\delta/2}(\widetilde{\sigma}_i) \cap B_{\delta/2}(\widetilde{\sigma}_j)$ and see that $B_\delta(\widetilde{\sigma}_{ij})$ (which as argued in the previous paragraph is a Cauchy ball) contains $B_{\delta/2}(\widetilde{\sigma}_i) \cup B_{\delta/2}(\widetilde{\sigma}_j)$, since the geodesic distance between a point of $B_{\delta/2}(\widetilde{\sigma}_i)$ and $\widetilde{\sigma}_{ij}$ is less than δ. If they do not intersect, we use proposition 3.2.7 to find a Cauchy ball containing them, so in either case there is a larger Cauchy ball that contains any two members of the cover. □

We finish this section with a couple of results in Lorentzian geometry which will be used in the subsequent sections.

Lemma 3.2.9. *Given a Cauchy surface Σ and a cover of an open subset $\Sigma' \subseteq \Sigma$ by Cauchy balls $\{B_n\}$, the union $\mathcal{O}_{\Sigma'} := \bigcup_n D(B_n)$ is a relatively time compact (see definition 1.3.5) $\mathsf{Loc}^{\mathcal{M}}$ region.*

Proof. For any $p \in \mathcal{M}$, $J^\pm(p)$ will only intersect finitely many of the $\{B_n\}$ since the intersection of $J^\pm(p)$ with Σ is compact. Therefore $J^\pm(p)$ will only intersect finitely many of the Cauchy diamonds $D(B_n)$, each of which are relatively compact, and hence the intersection of $\mathcal{O}_{\Sigma'}$ with $J^\pm(p)$ is relatively compact. This proves that $\mathcal{O}_{\Sigma'}$ is relatively time compact.

Consider a causal curve γ with end-points in $\mathcal{O}_{\Sigma'}$, these past and future end-points must be in Cauchy diamonds $D(B_i)$ and $D(B_j)$ respectively, for some i and j in the index set for the cover. Any inextendible extension of this causal curve must intersect B_i and B_j (definition of Cauchy development). Let us assume that these intersections occur at different points. Since B_i and B_j are subregions of Σ, this would mean the causal curve intersects Σ twice, but this contradicts the fact that Σ is a spacelike Cauchy surface. This therefore implies that any inextendible extension of γ intersects B_i and B_j at the same point in Σ. The causal curve γ is therefore the union of the segments which lie to the future/past of this intersection point. The past segment of γ has end-points in $D(B_i)$ which is causally convex, hence the past segment is entirely contained in $D(B_i)$, and similarly the future segment is entirely contained in $D(B_j)$. The curve as a whole is therefore contained in $D(B_i) \cup D(B_j)$. This establishes the causal convexity of $\mathcal{O}_{\Sigma'}$, which is also open since each of the Cauchy diamonds $D(B_n)$ is open, hence $\mathcal{O}_{\Sigma'} \in \mathsf{Loc}^{\mathcal{M}}$. □

Proposition 3.2.10. *Given any $\mathcal{O}_1, \mathcal{O}_2 \in \mathsf{Loc}_C^{\mathcal{M}}$ and causally disjoint compact sets $K_1 \subset \mathcal{O}_1$ and $K_2 \subset \mathcal{O}_2$, there exists a Cauchy surface Σ such that $J(K_1) \cap \Sigma$ and $J(K_2) \cap \Sigma$ are disjoint and contained in $\mathcal{O}_1$ and $\mathcal{O}_2$ respectively.*

Proof. We begin by finding open neighbourhoods $U_1 \supset K_1$ and $U_2 \supset K_2$ such that $U_1 \subset \mathcal{O}_1$ and $U_2 \subset \mathcal{O}_2$ are also causally disjoint. To do this we equip $\mathcal{M}$ with a Riemannian metric g_R with corresponding distance function $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$, which by [O'N83, Proposition 5.18] will generate the topology of $\mathcal{M}$. Each point $p \in K_i$ is a finite distance $\delta_p > 0$ (as measured by d) from $J(K_2)$, since K_1 and $J(K_2)$ are disjoint (by supposition K_1 and K_2 are causally disjoint) and closed ($J(K_2)$ is closed due to theorem 1.3.12). The open ball $B_{\epsilon_p}(p)$ of radius $\epsilon_p < \delta_p$ centred at p will therefore be disjoint from $J(K_2)$, and the set $\{B_{\epsilon_p}(p) \cap \mathcal{O}_1\}$ over all $p \in K_1$ defines an open cover K_1 contained in $\mathcal{O}_1$. Since K_1 is compact, this cover can be refined to a finite open subcover. Let U_1 be the union of all the balls in the finite subcover of K_1, then $K_1 \supset U_1 \subset \mathcal{O}_1$, and U_1 has compact closure which is causally disjoint from K_2, since the closure of a finite number of relatively compact balls is compact, and $\epsilon_p < \delta_p$ ensures disjointness from $J(K_2)$. We then repeat this procedure to get U_2 which is causally disjoint from $\overline{U}_1$ (which is possible since $\overline{U}_1$ is compact) and therefore U_1 and U_2 are causally disjoint.

The regions $\widetilde{U}_1 := J^+(U_1) \cap J^-(U_1)$ and $\widetilde{U}_2 := J^+(U_2) \cap J^-(U_2)$ are open (since the causal past/future of an open set in a globally hyperbolic spacetime is open, see for example [FV12, Lemma A.8]) causally convex subsets of $\mathcal{M}$, and due to theorem 1.3.15 they are globally hyperbolic and therefore admit Cauchy surfaces Σ_1 and Σ_2 respectively. We also have $\widetilde{U}_1 \subset \mathcal{O}_1$ and $\widetilde{U}_2 \subset \mathcal{O}_2$ since the causal convex hull of a subset of a causally convex region will remain a subset. Since $J(K_1) \cap \Sigma_1 \subset \widetilde{U}_1$, this implies $J(K_1) \cap \Sigma_1 \subset \mathcal{O}_1$ (and similarly $J(K_2) \cap \Sigma_2 \subset \mathcal{O}_2$). The regions $\widetilde{U}_1$ and $\widetilde{U}_2$ are causally disjoint due to the fact that any future-directed causal curve from $\widetilde{U}_1$ to $\widetilde{U}_2$ has past endpoint in $J^+(U_1)$ and future endpoint in $J^-(U_2)$, and therefore extends to a causal curve from U_1 to U_2. This therefore implies that $J(K_1) \cap \Sigma_1$ and $J(K_2) \cap \Sigma_2$ are disjoint.

We now use the smooth Urysohn's lemma [AMR88, proposition 5.5.8] to construct a smooth function $\chi_1 : \Sigma_1 \to \mathbb{R}$ which satisfies $\chi_1 = 1$ on $J(K_1) \cap \Sigma_1$ and $\chi_1 = 0$ outside of an open neighbourhood of $J(K_1) \cap \Sigma_1$. By Sard's theorem, the set of regular values of χ_1 is dense in $\mathbb{R}$, hence we can find $y \in (0, 1)$ which is a regular value of χ_1. This implies $\widetilde{\Sigma}_1 := \chi_1^{-1}([y, \infty)) \supset J(K_1) \cap \Sigma_1$ is closed, since the preimage of a continuous function on a closed set is closed. The closure of $\widetilde{U}_1$ is compact (see [HM19, Proposition 2.3]), and a closed subset of a compact set in a Hausdorff space is also compact, which implies $\widetilde{\Sigma}_1$ is compact. The region $\widetilde{\Sigma}_1$ is also a submanifold with boundary, since the boundary of $[y, \infty)$ is a regular value, and the preimage of a smooth function on a regular value is a submanifold [Hir76, Theorem 1.3.2]. We then repeat this procedure on $\widetilde{U}_2$ to get a compact submanifold with boundary $\widetilde{\Sigma}_2 \subset \Sigma_2$, which contains $J(K_2) \cap \Sigma_2$. We can now use [BS06, Theorem 1.1] to find a Cauchy surface Σ of $\mathcal{M}$ which contains $\widetilde{\Sigma}_1$ and $\widetilde{\Sigma}_2$, which implies $J(K_1) \cap \Sigma = J(K_1) \cap \Sigma_1$ and $J(K_2) \cap \Sigma = J(K_2) \cap \Sigma_2$, and therefore Σ satisfies the requirements of the proposition. □

3.3 Universal field algebras

Before proving the main result of the chapter, we will show another application of these geometrical techniques. The aim of this section is to analyse the universal theory $\mathscr{U}^{\mp}$ (which as mentioned in section 3.1 can be viewed as a class of theories) and show that it is equivalent to the original theory $\mathscr{A}^{\mp}$. This result establishes the non-triviality of the universal algebra constructed from $\mathscr{A}^{\mp}$, a result which is far from obvious and for general theories is not guaranteed. We begin by establishing the following result, which will be used in this section and the next section.

Lemma 3.3.1. *For any $\mathbf{G} = (\mathcal{M}, \mathcal{E}, P) \in$ GlobHypGreen, relatively time compact $\mathsf{Loc}^{\mathcal{M}}$ region $\mathcal{O}$, and $f \in C_0^\infty(\mathcal{M}, \mathcal{E})$ such that $J(\mathrm{supp} f) \subset J(\mathcal{O})$, there exists $\widetilde{f}$ with support in $\mathcal{O}$ such that $\widetilde{f} = f + Pg$ with $g \in C_0^\infty(\mathcal{M}, \mathcal{E})$.*

Proof. The collection $\{J^+(\mathcal{O})\,,\, J^-(\mathcal{O})\,,\, \mathcal{M}\backslash J(\mathrm{supp} f)\}$ forms an open cover of $\mathcal{M}$, since by supposition $J(\mathrm{supp} f) \subset J(\mathcal{O})$. We use the partition of unity $\{\chi^+, \chi^-, \chi^\perp\}$ subordinate to this cover to construct the following section

$$\widetilde{f} = f - P\chi^+ E^- f - P\chi^- E^+ f \quad . \tag{3.4}$$

When restricting to $J(\mathrm{supp} f)\backslash J^-(\mathcal{O})$ we get $\chi^- = \chi^\perp = 0$ and therefore $\widetilde{f} = f - PE^- f = 0$ when restricted to $J(\mathrm{supp} f)\backslash J^-(\mathcal{O})$, similarly we get $\widetilde{f} = 0$ when restricted to $J(\mathrm{supp} f)\backslash J^+(\mathcal{O})$. Since $\mathrm{supp}(E^\pm f) \subset J^\pm(\mathrm{supp} f)$, we also see that $\widetilde{f} = 0$ outside of $J(\mathrm{supp} f)$. It follows that $\mathrm{supp}(\widetilde{f}) \subset J^+(\mathcal{O}) \cap J^-(\mathcal{O})$ and therefore $\mathrm{supp}(\widetilde{f}) \subset \mathcal{O}$ since $\mathcal{O}$ is a causally convex subregion of $\mathcal{M}$. Since $\mathcal{O}$ is relatively time compact, $J^\pm(\mathcal{O}) \cap J^\mp(\mathrm{supp} f)$ is relatively compact by lemma 1.3.20. Since $\mathrm{supp}(P\chi^\pm E^\mp f) \subset J^\pm(\mathcal{O}) \cap J^\mp(\mathrm{supp} f)$ which is relatively compact, the latter two terms in equation (3.4) sum to a single term of the form Pg with $g \in C_0^\infty(\mathcal{M}, \mathcal{E})$. □

To prove the equivalence between $\mathscr{A}^{\mp}$ and $\mathscr{U}^{\mp}$, we will establish a natural transformation from $\mathscr{A}^{\mp}$ to $\mathscr{U}^{\mp}$. To do this we define a collection of homomorphisms in the following proposition and show that they satisfy a certain property. This property will then be used to show these homomorphisms are isomorphisms. We then also show that the homomorphisms form the components of a natural transformation.

Proposition 3.3.2. *For $\mathbf{G} = (\mathcal{M}, \mathcal{E}, P) \in$ GlobHypGreen, there is an $\mathsf{Alg}^{(h)}$ morphism $H_{\mathbf{G}} : \mathscr{A}^{\mp}(\mathbf{G}) \to \mathscr{U}^{\mp}(\mathbf{G})$ such that for any $\mathcal{O} \in \mathsf{Loc}_C^{\mathcal{M}}$, we have $H_{\mathbf{G}} \circ [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}}) = \phi_{\mathcal{O}}$ where $i_{\mathcal{O}}$ is the inclusion map $i_{\mathcal{O}} : O \to \mathcal{M}$ and $\phi_{\mathcal{O}} : \mathscr{A}^{\mp}_{\mathbf{G}}(\mathcal{O}) \to \mathscr{U}^{\mp}(\mathbf{G})$ are the morphisms from definition 3.1.11.*

Proof. We consider the following elements of $\mathscr{U}^{\mp}(\mathbf{G})$

$$\mathcal{A}_U(f) := \sum_i \mathcal{A}_{\mathcal{O}_i}(i^*_{\mathcal{O}_i}[\chi_i f]) \tag{3.5}$$

where $f \in C_0^\infty(\mathcal{M}, \mathcal{E})$, and $\{\chi_i\}$ is a partition of unity subordinate to a cover $\{\mathcal{O}_i\}$ of $\mathcal{M}$, with $\mathcal{O}_i \in \mathsf{Loc}_C^{\mathcal{M}}$. This definition is independent of which partition of unity is chosen. To see this we consider another partition of unity $\{\rho_j\}$ subordinate to $\{\mathcal{O}'_j\}$, and let $I_i^{(ij)}$ and

$I_j^{(ij)}$ be the inclusion maps from $\mathcal{O}_i \cap \mathcal{O}'_j$ into $\mathcal{O}_i$ and $\mathcal{O}'_j$ respectively. We then perform the following calculation

$$\begin{aligned}\mathcal{A}_U(f) &:= \sum_i \mathcal{A}_{\mathcal{O}_i}(i^*_{\mathcal{O}_i}[\chi_i f]) = \sum_{i,j} \mathcal{A}_{\mathcal{O}_i}(i^*_{\mathcal{O}_i}[\rho_j \chi_i f]) = \sum_{i,j} \mathcal{A}_{\mathcal{O}_i \cap \mathcal{O}'_j}(I_i^{(ij)\,*} i^*_{\mathcal{O}_i}[\rho_j \chi_i f]) \\ &= \sum_{i,j} \mathcal{A}_{\mathcal{O}'_j}(I^{(ij)}_{j\ *} I_i^{(ij)\,*} i^*_{\mathcal{O}_i}[\rho_j \chi_i f]) = \sum_{i,j} \mathcal{A}_{\mathcal{O}'_j}(i^*_{\mathcal{O}'_j}[\rho_j \chi_i f]) = \sum_j \mathcal{A}_{\mathcal{O}'_j}(i^*_{\mathcal{O}'_j}[\rho_j f])) \quad .\end{aligned}$$

The universal algebra $\mathscr{U}^{\mp}(\mathbf{G})$ is generated by elements of the form $\mathcal{A}_{\mathcal{O}}(i^*_{\mathcal{O}} f))$ with $f \in C_0^\infty(\mathcal{O}, \pi^{-1}(\mathcal{O}))$ (and the identity), and we can see that by picking a trivial partition of unity we get $\mathcal{A}_U(f) = \mathcal{A}_{\mathcal{O}}(i^*_{\mathcal{O}} f)$. We therefore see that $\mathscr{U}^{\mp}(\mathbf{G})$ is generated by the $\mathcal{A}_U(f)$ elements (and the identity), and therefore investigate their properties. It is easy to see from equation (3.5) that the generators are linear i.e,

$$\mathcal{A}_U(af + bg) = a\mathcal{A}_U(f) + b\mathcal{A}_U(g) \tag{3.6}$$

and consequently the field equations are satisfied

$$\begin{aligned}\mathcal{A}_U(Pf) &= \mathcal{A}_U(P\big[\sum_i \chi_i\big] f) = \sum_i \mathcal{A}_U(P\chi_i f) = \sum_i \mathcal{A}_{\mathcal{O}_i}(i^*_{\mathcal{O}_i}[P\chi_i f]) \\ &= \sum_i \mathcal{A}_{\mathcal{O}_i}(P_{\mathcal{O}_i} i^*_{\mathcal{O}_i}[\chi_i f]) = 0\end{aligned} \tag{3.7}$$

where we have used the fact that P is a linear operator that does not increase support. We can also prove that these generators satisfy the same commutation relations as the generators of the original theory. To see this we consider generators $\mathcal{A}_U(f)$ and $\mathcal{A}_U(g)$. We then consider a Cauchy surface Σ of $\mathcal{M}$ and note that $\Sigma \cap J(\operatorname{supp} f \cup \operatorname{supp} g)$ is compact by corollary 1.3.17, so we can use lemma 3.2.8 to get a finite cover $\{B_i\}$ of $\Sigma \cap J(\operatorname{supp} f \cup \operatorname{supp} g)$ by Cauchy balls such that any two members of the cover can be contained in a larger Cauchy ball. We then consider the region $\mathcal{O}_\Sigma := \bigcup_i D(B_i)$, which by lemma 3.2.9 is a relatively time compact $\mathsf{Loc}^{\mathcal{M}}$ region. Combining this with the fact that the generators $\mathcal{A}_U(f)$ satisfy linearity and equation of motion relations, we can use lemma 3.3.1 to find sections $\widetilde{f}$ and $\widetilde{g}$ supported in $\mathcal{O}_\Sigma$ such that $\mathcal{A}_U(f) = \mathcal{A}_U(\widetilde{f})$, $\mathcal{A}_U(g) = \mathcal{A}_U(\widetilde{g})$ and $\langle \widetilde{g}, E\widetilde{f}\rangle = \langle g, Ef\rangle$. We can then use a partition of unity $\{\rho_i\}$ subordinate to $\{D(B_i)\}$ to get

$$\begin{aligned}[\mathcal{A}_U(g)^*, \mathcal{A}_U(f)]_{\mp} &= [\mathcal{A}_U(\widetilde{g})^*, \mathcal{A}_U(\widetilde{f})]_{\mp} = \sum_{i,j} [\mathcal{A}_U(\rho_j \widetilde{g})^*, \mathcal{A}_U(\rho_i \widetilde{f})]_{\mp} \\ &= \sum_{i,j} [\mathcal{A}_{D(B_{ij})}(i^*_{D(B_{ij})}[\rho_j \widetilde{g}])^*, \mathcal{A}_{D(B_{ij})}(i^*_{D(B_{ij})}[\rho_i \widetilde{f}])]_{\mp} \\ &= \sum_{i,j} iE(\rho_j \widetilde{g}, \rho_i \widetilde{f})\mathbb{1} = iE(\widetilde{g}, \widetilde{f})\,\mathbb{1} = iE(g, f)\,\mathbb{1}\end{aligned} \tag{3.8}$$

where B_{ij} is a Cauchy ball which contains B_i and B_j (which exists due to our choice of cover). These generators can also be seen to satisfy

$$[\mathcal{A}_U(f), \mathcal{A}_U(g)]_{\mp} = 0 \tag{3.9}$$

by using the same techniques as in the previous case. We define a map $H_{\mathbf{G}} : \mathscr{A}^{\mp}(\mathbf{G}) \to \mathscr{U}^{\mp}(\mathbf{G})$ by its action on generators: $H_{\mathbf{G}}\big(\mathcal{A}(f)\big) = \mathcal{A}_U(f)$ and $H_{\mathbf{G}}(\mathbb{1}) = \mathbb{1}$. From equations (3.6)-(3.9) it follows that $H_{\mathcal{O}}$ is a well-defined $\mathsf{Alg}^{(h)}$ morphism. We can also see that for all $\mathcal{O} \in \mathsf{Loc}_C^{\mathcal{M}}$ and $\mathcal{A}(f) \in \mathscr{A}_{\mathbf{G}}^{\mp}(\mathcal{O})$

$$\big(H_{\mathbf{G}} \circ [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}})\big)\big(\mathcal{A}(f)\big) = H_{\mathbf{G}}\big(\mathcal{A}(i_{\mathcal{O}*}f)\big) = \mathcal{A}_U(i_{\mathcal{O}*}f) = \mathcal{A}_{\mathcal{O}}(f) := \phi_{\mathcal{O}}\big(\mathcal{A}(f)\big)$$

□

We also establish the following useful property of the morphisms $\mathscr{A}^{\mp}(i_{\mathcal{O}})$.

Lemma 3.3.3. *The morphisms $\mathscr{A}^{\mp}(i_{\mathcal{O}})$ are jointly-epic i.e, $\alpha \circ \mathscr{A}^{\mp}(i_{\mathcal{O}}) = \beta \circ \mathscr{A}^{\mp}(i_{\mathcal{O}})$ for all $\mathcal{O} \in \mathsf{Loc}_C^{\mathcal{M}}$ implies $\alpha = \beta$.*

Proof. Let $D : \mathsf{HVBundLoc} \to \mathsf{Vec}$ be the functor which takes an object of $\mathsf{HVBundLoc}$ to its vector space of test sections, and let $\mathcal{A}$ be the generating field of $\mathscr{A}^{\mp}$ which is a natural transformation $D \to F \circ \mathscr{A}^{\mp}$, where $\mathcal{F} : \mathsf{Alg} \to \mathsf{Vec}$ is the forgetful functor. Suppose $\alpha, \beta : \mathscr{A}^{\mp}(\mathcal{M}) \to A$ are Alg morphisms such that $\alpha \circ \mathscr{A}^{\mp}(i_{\mathcal{O}}) = \beta \circ \mathscr{A}^{\mp}(i_{\mathcal{O}})$. Then

$$\mathcal{F}(\alpha) \circ \mathcal{A}_M \circ D(i_{\mathcal{O}}) = \mathcal{F}(\alpha \circ \mathscr{A}^{\mp}(i_{\mathcal{O}})) \circ \mathcal{A}_{\mathcal{O}} = \mathcal{F}(\beta \circ \mathscr{A}^{\mp}(i_{\mathcal{O}})) \circ \mathcal{A}_{\mathcal{O}} = \mathcal{F}(\beta) \circ \mathcal{A}_M \circ D(i_{\mathcal{O}})$$

by naturality and the supposition. We then use the fact that the morphisms $D(i_{\mathcal{O}})$ are jointly epic (as can be seen by splitting f into parts supported in each $\mathcal{O}_i$ with a partition of unity) to get $\mathcal{F}(\alpha) \circ \mathcal{A}_M = \mathcal{F}(\beta) \circ \mathcal{A}_M$.

Now let $T(\mathcal{M})$ be the free unital *-algebra over $D(\mathcal{M})$, with canonical map $\Psi : D(\mathcal{M}) \to \mathcal{F}(T(\mathcal{M}))$ [in Vec], and $q : T(\mathcal{M}) \to \mathscr{A}^{\mp}(\mathcal{M})$ [in Alg] the quotient, so that $\mathcal{A}_{\mathcal{M}} = \mathcal{F}(q) \circ \Psi_{\mathcal{M}}$. Then $\mathcal{F}(\alpha) \circ \mathcal{A}_M = \mathcal{F}(\beta) \circ \mathcal{A}_M$ implies $\mathcal{F}(\alpha \circ q) \circ \Psi_{\mathcal{M}} = \mathcal{F}(\beta \circ q) \circ \Psi_{\mathcal{M}}$ and the universal property of the free unital *-algebra construction entails that $\alpha \circ q = \beta \circ q$. Since q is an epimorphism, we get $\alpha = \beta$. □

The algebra $\mathscr{A}^{\mp}(\mathbf{G})$ together with the inclusion morphisms from its local algebras form a cocone over the net of local algebras. The universal property of $\mathscr{U}^{\mp}(\mathbf{G})$ implies there must exist a unique $\mathsf{Alg}^{(h)}$ morphism $\widetilde{H}_{\mathbf{G}} : \mathscr{U}^{\mp}(\mathbf{G}) \to \mathscr{A}^{\mp}(\mathbf{G})$ such that $\widetilde{H}_{\mathbf{G}} \circ \phi_{\mathcal{O}} = [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}})$. The collection of morphisms $\{\widetilde{H}_{\mathbf{G}}\}$ defines a canonical natural transformation $\widetilde{H} : \mathscr{A}^{\mp} \to \mathscr{U}^{\mp}$. We are now in a position to prove the main result of this section.

Theorem 3.3.4. *The canonical natural transformation $\widetilde{H} : \mathscr{A}^{\mp} \to \mathscr{U}^{\mp}$ is a natural isomorphism.*

Proof. We use the map $H_{\mathbf{G}}$ from proposition 3.3.2 to get the following pair of commutative triangles

$$
\begin{tikzcd}
\mathscr{A}^{\mp}_{\mathbf{G}}(\mathcal{O}) \arrow[rr, "\phi_{\mathcal{O}}"] \arrow[drr, "H_{\mathbf{G}} \circ \widetilde{H}_{\mathbf{G}} \circ \phi_{\mathcal{O}}"'] & & \mathscr{U}^{\mp}(\mathbf{G}) \arrow[d, shift right, "H_{\mathbf{G}} \circ \widetilde{H}_{\mathbf{G}}"'] \arrow[d, shift left, "\mathrm{Id}_{\mathscr{U}^{\mp}(\mathbf{G})}"] \\
& & \mathscr{U}^{\mp}(\mathbf{G})
\end{tikzcd}
.
$$

The inner triangle commutes by definition, the outer triangle commutes due to the fact that $H_{\mathbf{G}} \circ [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}}) = \phi_{\mathcal{O}}$ (by proposition 3.3.2) combined with $\widetilde{H}_{\mathbf{G}} \circ \phi_{\mathcal{O}} = [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}})$. The universal algebra together with the morphisms $H_{\mathbf{G}} \circ \widetilde{H}_{\mathbf{G}} \circ \phi_{\mathcal{O}}$ also form a cocone over the net of local algebras, therefore the universal property of $\mathscr{U}^{\mp}(\mathbf{G})$ implies $H_{\mathbf{G}} \circ \widetilde{H}_{\mathbf{G}} = \mathrm{Id}_{\mathscr{U}^{\mp}(\mathbf{G})}$. We also get $\widetilde{H}_{\mathbf{G}} \circ H_{\mathbf{G}} \circ [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}}) = [\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}})$, which by the joint-epic property of the $[\mathscr{A}^{\mp} \circ I_{\mathbf{G}}](i_{\mathcal{O}})$ morphisms (see lemma 3.3.3) implies $\widetilde{H}_{\mathbf{G}} \circ H_{\mathbf{G}} = \mathrm{Id}_{\mathscr{A}^{\mp}(\mathbf{G})}$. Therefore $H_{\mathbf{G}}$ is an isomorphism.

Given a $\mathsf{GlobHypGreen}$ morphism $\Psi : \mathbf{G_1} \to \mathbf{G}_2$, we perform the following calculations (leaving the algebra that each generator belongs to implicit)

$$
\begin{aligned}
&\big(\mathscr{U}^{\mp}(\Psi) \circ H_{\mathbf{G}_1}\big)\mathcal{A}(f) = \mathscr{U}^{\mp}(\Psi)\mathcal{A}_U(f) = \mathscr{U}^{\mp}(\Psi) \sum_i \mathcal{A}_{\mathcal{O}_i}(i^*_{\mathcal{O}_i}[\chi_i f]) \\
&= \sum_i \mathcal{A}_{\mathcal{O}_i}(\Psi_* i^*_{\mathcal{O}_i}[\chi_i f]) = \mathcal{A}_U(\Psi_* f)
\end{aligned}
$$

and

$$
\big(H_{\mathbf{G}_2} \circ \mathscr{A}^{\mp}(\Psi)\big)\mathcal{A}(f) = H_{\mathbf{G}_2}\mathcal{A}(\Psi_* f) = \mathcal{A}_U(\Psi_* f)
$$

from which it follows that

$$
\big(\mathscr{U}^{\mp}(\Psi) \circ H_{\mathbf{G}_1}\big)\mathcal{A}(f) = \big(H_{\mathbf{G}_2} \circ \mathscr{A}^{\mp}(\Psi)\big)\mathcal{A}(f) \ ,
$$

and therefore we see that the collection $H_{\mathbf{G}}$ over all $\mathbf{G} \in \mathsf{GlobHypGreen}$ defines a natural isomorphism between $\mathscr{A}^{\mp}$ and $\mathscr{U}^{\mp}$. □

3.4 Einstein causality for universal even algebras

We shall first prove the following result for generators of $\mathscr{U}^{\mp}_E(\mathbf{G})$.

Proposition 3.4.1. *Let $\mathbf{G} \in$ GlobHypGreen and $\mathcal{O}_1$ and $\mathcal{O}_2$ be objects of $\mathsf{Loc}^{\mathcal{M}}_C$ and let $\mathcal{A}_{\mathcal{O}_1}(u,v)$ and $\mathcal{A}_{\mathcal{O}_2}(f,g)$ be algebra elements of $\mathscr{U}^{\mp}_E(\mathbf{G})$ of the form specified below definition 3.1.11. If* $\mathrm{supp}(u) \cup \mathrm{supp}(v)$ *and* $\mathrm{supp}(f) \cup \mathrm{supp}(g)$ *are causally disjoint, the following relation holds in $\mathscr{U}^{\mp}_E(\mathbf{G})$*

$$
[\mathcal{A}_{\mathcal{O}_1}(u,v), \mathcal{A}_{\mathcal{O}_2}(f,g)]_- = 0 \quad .
$$

Proof. The idea of the proof here is similar to the computation of the commutation/anticommuatation relations in proposition 3.3.2, where we used lemma 3.3.1 to get generators with sections supported in a convenient neighbourhood. This allowed us to split the algebra generators as sums over generators labelled by smaller $\mathsf{Loc}_C^{\mathcal{M}}$ regions that could be embedded, pairwise, in common $\mathsf{Loc}_C^{\mathcal{M}}$ regions. We could then evaluate each commutator separately using the relations from the local algebras and sum the results. Summing the results was possible because each local commutator resulted in something proportional to the identity, so there were no issues with combining the terms. In the even algebra case, the local commutator

$$[\mathcal{A}_{\mathcal{O}'}(u',v'),\mathcal{A}_{\mathcal{O}'}(f',g')]_-$$

can result in something that has field content, which would make summing the end results an issue, because the linearity relation for fields can only be applied if they can be embedded in a common $\mathsf{Loc}_C^{\mathcal{M}}$ region. The local commutator will only result in something proportional to the identity if $\mathrm{supp}(u')\cup\mathrm{supp}(v')$ and $\mathrm{supp}(f')\cup\mathrm{supp}(g')$ are causally disjoint. We therefore need to ensure that when we use lemma 3.3.1, the resulting sections remain causally disjoint from each other.

Let $K_1=\mathrm{supp}(u)\cup\mathrm{supp}(v)$ and $K_2=\mathrm{supp}(f)\cup\mathrm{supp}(g)$. Since K_1 and K_2 are causally disjoint we can use proposition 3.2.10 to find a Cauchy surface Σ of $\mathcal{M}$ such that $J(K_1)\cap\Sigma$ and $J(K_2)\cap\Sigma$ are disjoint and contained in $\mathcal{O}_1$ and $\mathcal{O}_2$ respectively. Since $J(K_1)\cap\Sigma$ and $J(K_2)\cap\Sigma$ are compact by corollary 1.3.17, we can use lemma 3.2.8 to construct open covers $\{B_n^{(1)}\}$ and $\{B_n^{(2)}\}$ of $J(K_1)\cap\Sigma$ and $J(K_2)\cap\Sigma$ respectively (superscripts are used to denote which cover a given Cauchy ball belongs to), with the property that any two members of a given cover can be contained in a larger Cauchy ball. The covers $\{B_n^{(1)}\}$ and $\{B_n^{(2)}\}$ can be made disjoint from each other and contained in $\mathcal{O}_1$ and $\mathcal{O}_2$ respectively, by sufficiently shrinking the Cauchy balls, and this does not change the properties of the covers.

Let Σ_1 and Σ_2 be Cauchy surfaces of $\mathcal{O}_1$ and $\mathcal{O}_2$ respectively. Every inextendible causal curve that intersects $B_n^{(1)}$ must intersect Σ_1 since $B_n^{(1)}\subset\mathcal{O}_1$, and therefore $D(B_n^{(1)})\subset D(\Sigma_1)$. This implies that $\widetilde{\mathcal{O}}_1:=\bigcup_n D(B_n^{(1)})$ and $\widetilde{\mathcal{O}}_2:=\bigcup_n D(B_n^{(2)})$ are contained in the $\mathsf{Loc}^{\mathcal{M}}$ regions $D(\Sigma_1)$ and $D(\Sigma_2)$ respectively. The regions $\widetilde{\mathcal{O}}_1$ and $\widetilde{\mathcal{O}}_2$ are also causally disjoint because the hypersurfaces $\bigcup_n B_n^{(1)}$ and $\bigcup_n B_n^{(2)}$ are disjoint regions of a single Cauchy surface Σ. By construction $J(K_1)\subset J(\widetilde{\mathcal{O}}_1)$, and by lemma 3.2.9 $\widetilde{\mathcal{O}}_1$ is a relatively time compact $\mathsf{Loc}^{\mathcal{M}}$ region, so we can use lemma 3.3.1 to find sections $\widetilde{u}$ and $\widetilde{v}$ supported in $\widetilde{\mathcal{O}}_1$ such that $\mathcal{A}_{D(\Sigma_1)}(u,v)=\mathcal{A}_{D(\Sigma_1)}(\widetilde{u},\widetilde{v})$. Similarly we can find sections $\widetilde{f}$ and $\widetilde{g}$ supported in $\widetilde{\mathcal{O}}_2$ such that $\mathcal{A}_{D(\Sigma_2)}(f,g)=\mathcal{A}_{D(\Sigma_2)}(\widetilde{f},\widetilde{g})$.

We then construct partitions of unity $\{\rho_n\}$ and $\{\phi_n\}$ subordinate to the covers $\{D(B_n^{(1)})\}$ and $\{D(B_n^{(2)})\}$ of $\widetilde{\mathcal{O}}_1$ and $\widetilde{\mathcal{O}}_2$. Given Cauchy balls $B_i^{(1)}$, $B_j^{(1)}$, $B_k^{(2)}$ and $B_l^{(2)}$ (superscript denoting the cover which the Cauchy ball comes from) from these covers, we would like to find a single Cauchy ball that contains all of them. There is a larger Cauchy ball $B_{ij}\subset\bigcup_n B_n^{(1)}$ that contains $B_i^{(1)}$ and $B_j^{(1)}$ (this follows from the properties of the cover that we've used), and we similarly find $B_{kl}\subset\bigcup_n B_n^{(2)}$ that contains $B_k^{(2)}$ and $B_l^{(2)}$. The Cauchy balls B_{ij} and B_{kl} are disjoint, since $\bigcup_n B_n^{(1)}$ and $\bigcup_n B_n^{(2)}$ are disjoint, so we can therefore use proposition 3.2.7 to find a Cauchy ball B_{ijkl} that contains B_{ij} and B_{kl}. We can now prove the proposition by using our partitions of unity as follows

$$\begin{aligned}
[\mathcal{A}_{\mathcal{O}_1}(u,v),\mathcal{A}_{\mathcal{O}_2}(f,g)]_- &= [\mathcal{A}_{D(\Sigma_1)}(u,v),\mathcal{A}_{D(\Sigma_2)}(f,g)]_- = [\mathcal{A}_{D(\Sigma_1)}(\widetilde{u},\widetilde{v}),\mathcal{A}_{D(\Sigma_2)}(\widetilde{f},\widetilde{g})]_- \\
&= \sum_{i,j,k,l}[\mathcal{A}_{D(\Sigma_1)}(\rho_i\widetilde{u},\rho_j\widetilde{v}),\mathcal{A}_{D(\Sigma_2)}(\phi_k\widetilde{f},\phi_l\widetilde{g})]_- \\
&= \sum_{i,j,k,l}[\mathcal{A}_{D(B_{ijkl})}(\rho_i\widetilde{u},\rho_j\widetilde{v}),\mathcal{A}_{D(B_{ijkl})}(\phi_k\widetilde{f},\phi_l\widetilde{g})]_- = 0
\end{aligned}$$

where the second equality follows from the fact that $\mathcal{O}_1 \subset D(\Sigma_1)$ and $\mathcal{O}_2 \subset D(\Sigma_2)$. The final equality is a consequence of using the commutation relations from the local algebras (yielding 0 because $\widetilde{\mathcal{O}}_1$ and $\widetilde{\mathcal{O}}_2$ are causally disjoint), which can be applied in this context because the sub-indices of the generators agree for each term in the sum. □

We are now in a position to prove the main result of the chapter.

Theorem 3.4.2. *Einstein causality holds in the theories $\mathscr{U}^-$ and $\mathscr{U}_E^{\mp}$, and graded Einstein causality holds in $\mathscr{U}^+$.*

Proof. Since $\mathscr{A}^-$ and $\mathscr{U}^-$ are isomorphic by theorem 3.3.4, and Einstein causality holds in $\mathscr{A}^-$, it automatically follows that Einstein causality also holds in $\mathscr{U}^-$. Similarly $\mathscr{U}^+$ satisfies graded Einstein causality since $\mathscr{A}^+$ does (see comments below definition 3.1.5). The result for $\mathscr{U}_E^{\mp}$ follows from the fact that the result holds for generators, as proven in proposition 3.4.1. □

Chapter 4

Spin structures and Fermionic quantisation

In this chapter we shall consider the class of linear Fermionic QFTs defined in the previous chapter. The aim of this chapter is to show how topological information, in particular the number of inequivalent spin structures a given spacetime admits, can be extracted from a given linear Fermionic theory.

We obtain spin information from the theory $\mathscr{A}$, by considering extensions of subtheories of $\mathscr{A}$. From $\mathscr{A}$ we define an even subtheory $\mathscr{A}_E$, which associates to $\mathcal{M}$ the subalgebra of $\mathscr{A}(\mathcal{M})$ generated by pairs of quantum fields (see definition 3.1.6). Starting from $\mathscr{A}_E$, we then define a restricted theory $\mathscr{A}'_E$, by restricting the subtheory $\mathscr{A}_E$ to contractible causally convex spacetime regions. We can then obtain an extended theory from $\mathscr{A}'_E$, which we will call the universal even theory associated to $\mathscr{A}$. This extension technique relies on the *universal algebra* construction [Fre90] (see definition 3.1.11), which is known in category theory as a left Kan extension (see section 1 of chapter 3 for details).

For each linear Fermionic theory in the class of theories defined in definition 3.1.5, there is a corresponding universal theory. The main result of this chapter is that each of the resulting universal even theories maps a spacetime $\mathcal{M}$, to an algebra that decomposes into a product of subalgebras. Moreover, these subalgebras are indexed by the set $H^1(\mathcal{M}, \mathbb{Z}_2)$. This cohomology set counts the number of distinct spin structures the spacetime manifold $\mathcal{M}$ permits, assuming its second Stiefel-Whitney class is trivial. A spin structure (see definition 2.3.3) encodes information about how the spin bundle covers the frame bundle, and is required to define fields that transform in half-integer spin representations of the Lorentz group.

The idea of using the universal algebra to analyse topological aspects of quantum field theories has also been explored by Lang in his PhD thesis [Lan12]. In this thesis Lang uses categorical techniques to prove the existence of *twisted* variants of a quantum field theory, and to classify them. The analysis is restricted to analysing twisted variants of theories defined on locally constant bundles, which have constant transition functions. We do not impose such a constraint in this analysis. Another analysis using the universal algebra construction has been conducted for the source-free Maxwell field by Dappiaggi and Lang [DL12]. Their analysis has a similar flavour; they find that the centre of their universal algebra corresponds to topological invariants of the theory. Specifically, they find that each element of the second De Rham cohomology group has a corresponding central element in the universal algebra.

The chapter is organised as follows: In section 1 we introduce a class of elements in the universal algebra, denoted by $\mathcal{Y}$, which will eventually be shown to be central elements and will be used to define projection operators to decompose the universal algebra. We show that each element of $\mathcal{Y}$ can be defined solely in terms of four $\mathsf{Loc}_C^{\mathcal{M}}$ regions (J^+, J^-, I_1, I_2) satisfying $I_1 \cup I_2 \subset J^+ \cap J^-$, which we call a quadruple (see definition 4.1.2). In section 2 we then show how we can relate the elements of $\mathcal{Y}$ to the topology of the spacetime the universal algebra was built from. This is done by defining an equivalence relation on the set of quadruples, which ensures that equivalent quadruples correspond to equal elements of $\mathcal{Y}$, and then defining a map $\mathcal{L}$ from equivalence classes of quadruples to the set $\widetilde{\pi}_1(\mathcal{M})$ which is closely related to the fundamental group of the background spacetime (see definition 4.2.3). In section 3 we develop some geometrical techniques that are then used in section 4 to prove that the map $\mathcal{L}$ is a bijection. In section 5 we define a semi-group structure on $\mathcal{Y}$, and use the results from previous section to show that the resulting semi-group is isomorphic to $\pi_1(\mathcal{M}, p)/\pi_1(\mathcal{M}, p)^2$ (and is therefore a group). In section 6 we use the results from the preceding section to prove additional properties of the elements of $\mathcal{Y}$, in particular that they are central. From these properties, it follows that the elements of $\mathcal{Y}$ can be used to define projections that are mutually orthogonal, which then allows us to decompose the universal algebra into a product of subalgebras. The homomorphism established in the previous section then shows that these subalgebras are in one to one correspondence with the elements of $H^1(\mathcal{M}, \mathbb{Z}_2)$. We then finish with some concluding remarks, outlining further work that could build on these results.

4.1 Distinguished elements of the universal algebra

Throughout this chapter we will be analysing the theory $\mathscr{A}_E^+ : \mathsf{GlobHypGreen} \to \mathsf{Alg}$ defined in definition 3.1.6. As was noted in section 3.1, this single theory encompasses a whole class of theories, and in particular it contains the even subtheory of the Dirac field, which is obtained by restricting $\mathscr{A}_E^+$ to the image of $\mathsf{SpinLoc}$ in $\mathsf{GlobHypGreen}$ (see comments below definition 3.1.5). We will be focusing on $\mathscr{U}_E^+(\mathbf{G})$ for a general object $\mathbf{G} = (\mathcal{M}, \mathcal{E}, P)$ in $\mathsf{GlobHypGreen}$. In this section we will define certain elements of $\mathscr{U}_E^+(\mathbf{G})$ which will help us find topological invariants of $\mathcal{M}$. To do this, we first consider algebra elements of the local algebra $\mathscr{A}_{E\,\mathbf{G}}^+(\mathcal{O})$ of the form

$$b(f, g) = \big(\mathcal{A}(f) + \mathcal{A}(f)^*\big)\big(\mathcal{A}(g) + \mathcal{A}(g)^*\big) \tag{4.1}$$

where f and g are compactly supported sections of $\mathcal{E}$, and are chosen to be normalised as $E(f, f) = E(g, g) = -i$ (see definition 3.1.3 for $E(\cdot, \cdot)$). This is mostly done for convenience, but also to exclude sections that cannot be normalised which would make the above product collapse to zero. It can be seen from the anti-commutation relations in definition 3.1.5 that

$$b(f, f) = \big(\mathcal{A}(f) + \mathcal{A}(f)^*\big)^2 = \mathcal{A}(f)^2 + \mathcal{A}(f)\mathcal{A}(f)^* + \mathcal{A}(f)\mathcal{A}(f)^* + \mathcal{A}(f)^{*\,2} = \mathbb{1} \quad .$$

Using this result we see these elements also satisfy another useful relation (h also denotes a compactly supported section of $\mathcal{E}$)

$$\begin{aligned} b(f,h)b(h,g) &= \big(\mathcal{A}(f)+\mathcal{A}(f)^*\big)\big(\mathcal{A}(h)+\mathcal{A}(h)^*\big)\big(\mathcal{A}(h)+\mathcal{A}(h)^*\big)\big(\mathcal{A}(g)+\mathcal{A}(g)^*\big) \\ &= \big(\mathcal{A}(f)+\mathcal{A}(f)^*\big)b(h,h)\big(\mathcal{A}(g)+\mathcal{A}(g)^*\big) = b(f,g) \quad . \end{aligned}$$

If $E(f,g)=0$, we can also use the anti-commutation relations to get the following relation

$$\begin{aligned} b(f,g) &= \big(\mathcal{A}(f)+\mathcal{A}(f)^*\big)\big(\mathcal{A}(g)+\mathcal{A}(g)^*\big) \\ &= -\big(\mathcal{A}(g)+\mathcal{A}(g)^*\big)\big(\mathcal{A}(f)+\mathcal{A}(f)^*\big) = -b(g,f) \quad . \end{aligned}$$

It is also easy to see that $b(f,g)^* = b(g,f)$, for convenience we will summarise these relations here:

i) $b(f,f) = \mathbb{1}$.

ii) $b(f,h)b(h,g) = b(f,g)$.

iii) $b(f,g)^* = b(g,f)$.

iv) $b(f,g) = -b(g,f)$ if $E(f,g)=0$.

These local algebra elements also satisfy another useful property.

Lemma 4.1.1. *The elements $b(f,g)$ generate $\mathscr{A}^+_{E\,\mathbf{G}}(\mathcal{O})$.*

Proof. We first prove that normalised generators $\mathcal{A}(f)\mathcal{A}(g)$ and $\mathcal{A}(f)\mathcal{A}(g)^*$ (meaning $E(f,f) = E(g,g) = -i$) of $\mathscr{A}^+_{E\,\mathbf{G}}(\mathcal{O})$ can be obtained from the $b(f,g)$ elements. To do this we expand out the following elements

$$\begin{aligned} b(f,g) &= \mathcal{A}(f)\mathcal{A}(g)+\mathcal{A}(f)\mathcal{A}(g)^*+\mathcal{A}(f)^*\mathcal{A}(g)+\mathcal{A}(f)^*\mathcal{A}(g)^* \\ ib(if,g) &= -\mathcal{A}(f)\mathcal{A}(g)-\mathcal{A}(f)\mathcal{A}(g)^*+\mathcal{A}(f)^*\mathcal{A}(g)+\mathcal{A}(f)^*\mathcal{A}(g)^* \\ ib(f,ig) &= -\mathcal{A}(f)\mathcal{A}(g)+\mathcal{A}(f)\mathcal{A}(g)^*-\mathcal{A}(f)^*\mathcal{A}(g)+\mathcal{A}(f)^*\mathcal{A}(g)^* \\ b(if,ig) &= -\mathcal{A}(f)\mathcal{A}(g)+\mathcal{A}(f)\mathcal{A}(g)^*+\mathcal{A}(f)^*\mathcal{A}(g)-\mathcal{A}(f)^*\mathcal{A}(g)^* \end{aligned}$$

and then take linear combinations to get

$$\begin{aligned} \mathcal{A}(f)\mathcal{A}(g)^* &= \frac{1}{4}\left[b(f,g)-ib(if,g)+ib(f,ig)+b(if,ig)\right] \\ \mathcal{A}(f)\mathcal{A}(g) &= \frac{1}{4}\left[b(f,g)-ib(if,g)-ib(f,ig)-b(if,ig)\right] \quad . \end{aligned}$$

From these normalised generators we can generate $\mathscr{A}^+_{E\,\mathbf{G}}(\mathcal{O})$, hence the elements $b(f,g)$ generate $\mathscr{A}^+_{E\,\mathbf{G}}(\mathcal{O})$. □

This implies that the elements $b_{\mathcal{O}}(f,g) := \phi_{\mathcal{O}}(b(f,g))$ with f and g normalised generate $\mathscr{U}^+_E(\mathbf{G})$. With these elements and their relations now established, we introduce a definition which will be used to define the elements of the universal algebra we want to study.

Definition 4.1.2. *A quadruple consists of an ordered set of four $\mathsf{Loc}_C^{\mathcal{M}}$ regions (J^+, J^-, I_1, I_2) satisfying the condition (which is depicted in figure 4.1)*

$$I_1 \cup I_2 \subset J^+ \cap J^- \tag{4.2}$$

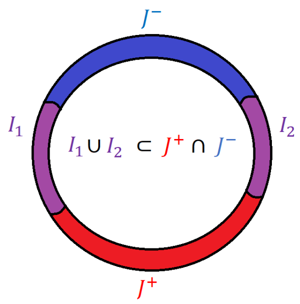

Figure 4.1: Constraint on the regions used to define a quadruple (note that I_1 and I_2 need not be disjoint and can belong to the same path-component of $J^+ \cap J^-$).

We are now in a position to introduce the algebra elements that we will be analysing for the rest of the chapter.

Definition 4.1.3. *Given a quadruple (J^+, J^-, I_1, I_2) and a pair of compactly supported normalised sections f and g of $\mathcal{E}$ with supports in I_1 and I_2 respectively, we can construct the following algebra elements of $\mathscr{U}_E^+(\mathbf{G})$*

$$Y_{(J^+,J^-,I_1,I_2)}(f,g) = b_{J^+}(f,g)b_{J^-}(g,f) \quad . \tag{4.3}$$

We will use $\mathcal{Y} \subset \mathscr{U}_E^+(\mathbf{G})$ to denote the set of all elements of the form (4.3).

We will now explore some of the properties of the elements of $\mathcal{Y}$.

Lemma 4.1.4. *$Y_{(J^+,J^-,I_1,I_2)}(f,g)^* = Y_{(J^-,J^+,I_1,I_2)}(f,g)$ for any normalised sections f and g compactly supported in I_1 and I_2 respectively.*

Proof. Using relation iii) we get

$$\begin{aligned} Y_{(J^+,J^-,I_1,I_2)}(f,g)^* &= \big(b_{J^+}(f,g)b_{J^-}(g,f)\big)^* = b_{J^-}(g,f)^* b_{J^+}(f,g)^* = b_{J^-}(f,g)b_{J^+}(g,f) \\ &= Y_{(J^-,J^+,I_1,I_2)}(f,g) \quad . \end{aligned}$$

□

Lemma 4.1.5. *$Y_{(J^+,J^-,I_1,I_2)}(f,g) = Y_{(J^+,J^-,I_1,I_2)}(f,g')$ for any normalised sections g, g' compactly supported in I_2 and any normalised section f compactly supported in I_1.*

Proof. Using the relations i), ii) and the cocone property of the universal algebra we get

$$\begin{aligned} &Y_{(J^+,J^-,I_1,I_2)}(f,g) = b_{J^+}(f,g)b_{J^-}(g,f) \overset{i)+ii)}{=} b_{J^+}(f,g)b_{I_2}(g,g')b_{I_2}(g',g)b_{J^-}(g,f) \\ &= b_{J^+}(f,g)b_{J^+}(g,g')b_{J^-}(g',g)b_{J^-}(g,f) \overset{i)}{=} b_{J^+}(f,g')b_{J^-}(g',f) = Y_{(J^+,J^-,I_1,I_2)}(f,g') \quad . \end{aligned}$$

Note that to get from the first line to the second line we have used the fact that I_2 embeds into J^+ and into J^- (this is part of the definition of a quadruple). $\square$

Proposition 4.1.6. *$Y_{(J^+,J^-,I_1,I_2)}(f,g) = Y_{(J^+,J^-,I_1,I_2)}(f',g')$ for any normalised sections f, f' compactly supported in I_1 and g, g' compactly supported in I_2.*

Proof. Let g_1, g_2, g_3 be linearly independent normalised sections with compact support in I_2, using these we construct the following sections

$$g_{12} = E(f,g_1)g_2 - E(f,g_2)g_1 \quad \text{and} \quad g_{23} = E(f,g_2)g_3 - E(f,g_3)g_2$$

which we then use to construct the following section

$$g'' = E(f',g_{12})g_{23} - E(f',g_{23})g_{12}$$

which has the property that $E(f,g'') = 0 = E(f',g'')$ (which still holds after normalising g''). Relation iv) implies $Y_{(J^+,J^-,I_1,I_2)}(f,g) = Y_{(J^+,J^-,I_2,I_1)}(g,f)$ if $E(f,g) = 0$, so we can use this and the fact that $E(f,g'') = 0 = E(f',g'')$ in the following calculation

$$\begin{aligned} &Y_{(J^+,J^-,I_1,I_2)}(f,g) = Y_{(J^+,J^-,I_1,I_2)}(f,g'') = Y_{(J^+,J^-,I_2,I_1)}(g'',f) \\ &= Y_{(J^+,J^-,I_2,I_1)}(g'',f') = Y_{(J^+,J^-,I_1,I_2)}(f',g'') = Y_{(J^+,J^-,I_1,I_2)}(f',g') \end{aligned}$$

where we have used lemma 4.1.5 to change g to g'' in the first equality, to change f to f' in the first equality on the second line, and to change g'' to g' in the final equality. $\square$

Corollary 4.1.7. *$Y_{(J^+,J^-,I_1,I_2)}(f,g) = Y_{(J^+,J^-,I_2,I_1)}(g',f')$ for any normalised sections f, f' and g, g' compactly supported in I_1 and I_2 respectively.*

Proof. As was shown in the proof of proposition 4.1.6, there exists a section g'' such that

$$Y_{(J^+,J^-,I_1,I_2)}(f,g) = Y_{(J^+,J^-,I_1,I_2)}(f,g'') = Y_{(J^+,J^-,I_2,I_1)}(g'',f) \quad .$$

We can then apply proposition 4.1.6 to the RHS of this equation to change g'' to g' and f to f', thus proving the result. $\square$

Lemma 4.1.8. *For $\mathsf{Loc}_C^{\mathcal{M}}$ regions J, J^+, J^-, I_1 and I_2 such that $I_1 \cup I_2 \subset J \cap J^+ \cap J^-$, the following equality holds*

$$Y_{(J^+,J,I_1,I_2)}(f,g)\, Y_{(J,J^-,I_1,I_2)}(f',g') = Y_{(J^+,J^-,I_1,I_2)}(f'',g'') \quad ,$$

for all normalised sections f, f', f'' and g, g', g'' compactly supported in I_1 and I_2 respectively.

Proof. Using proposition 4.1.6 we get

$$\begin{aligned} &Y_{(J^+,J,I_1,I_2)}(f,g)\, Y_{(J,J^-,I_1,I_2)}(f',g') = Y_{(J^+,J,I_1,I_2)}(f'',g'')\, Y_{(J,J^-,I_1,I_2)}(f'',g'') \\ &= b_{J^+}(f'',g'')b_J(g'',f'')b_J(f'',g'')b_{J^-}(g'',f'') = b_{J^+}(f'',g'')b_{J^-}(g'',f'') \\ &= Y_{(J^+,J^-,I_1,I_2)}(f'',g'') \quad . \end{aligned}$$

$\square$

We have now established the properties we will need of the elements of $\mathcal{Y}$ to allow us to decompose $\mathscr{U}_E^+(\mathbf{G})$. It is worth noting at this point that the decomposition will be trivial if the elements of $\mathcal{Y}$ only include the identity. This is because, as we will show later, the elements of $\mathcal{Y}$ define projection operators onto subalgebras. The reason that we are trying to extend $\mathscr{A}_E^+$ rather than $\mathscr{A}^+$, is to avoid the elements of $\mathcal{Y}$ being trivial. To see why this is, we introduce the map $F(f) = \mathcal{A}(f) + \mathcal{A}(f)^*$ that takes sections of $\pi^{-1}(\mathcal{O})$ (where π is the bundle projection $\pi : \mathcal{E} \to \mathcal{M}$) to elements of $\mathscr{A}_{\mathbf{G}}^+(\mathcal{O})$. The map F has the property that $F(f)$ admits an inverse $F(f)^{-1}$. We then note that the elements $b(f,g)$ take the form

$$b(f,g) = F(f)F(g)^{-1} \quad .$$

If we were to analyse the elements of $\mathcal{Y}$ in $\mathscr{U}^+(\mathbf{G})$ instead of $\mathscr{U}_E^+(\mathbf{G})$, we would see that they would all collapse to the identity, since we could use the cocone property of the universal algebra as follows

$$\begin{aligned}
&Y_{(J^+,J^-,I_1,I_2)}(f,g) = \phi_{J^+}(b(f,g))\phi_{J^-}(b(g,f)) = \phi_{J^+}(F(f)F(g)^{-1})\phi_{J^-}(F(g)F(f)^{-1}) \\
&= \phi_{J^+}(F(f))\phi_{J^+}(F(g))^{-1}\phi_{J^-}(F(g))\phi_{J^-}(F(f))^{-1} \\
&= \phi_{I_1}(F(f))\phi_{I_2}(F(g))^{-1}\phi_{I_2}(F(g))\phi_{I_1}(F(f))^{-1} = \mathbb{1} \quad .
\end{aligned}$$

Such a manipulation cannot be done in $\mathscr{U}_E^+(\mathbf{G})$ since the equality going from the first line to the second would not be possible: the element $F(f)$ does not belong to the local even algebra $\mathscr{A}_{E\,\mathbf{G}}^+(J^+)$. With this clarified, we will now begin to establish a relation between the elements of $\mathcal{Y}$ and the spacetime topology.

4.2 Relating the universal algebra to the spacetime topology

By virtue of proposition 4.1.6, the elements of the set $\mathcal{Y}$ from definition 4.1.3 only depend on the quadruple used to define them. We can therefore define a map from quadruples to elements of $\mathcal{Y}$. The aim of this section is to relate the universal algebra to the topology of $\mathcal{M}$ by studying the quadruples associated to elements of $\mathcal{Y}$. We begin by defining a notion of equivalence on the set of quadruples.

Definition 4.2.1. *We define an equivalence relation on the set of quadruples, generated by the following relations*

i) $j^+ \subset J^+$, $j^- \subset J^-$, $i_1 \subset I_1$, $i_2 \subset I_2 \;\Rightarrow\; (j^+, j^-, i_1, i_2) \sim (J^+, J^-, I_1, I_2)$.

ii) $(J^+, J^-, I_1, I_2) \sim (J^+, J^-, I_2, I_1)$.

Thus, two quadruples are considered equivalent if they can be linked by a sequence of quadruples related by one of the above conditions. We use $\mathcal{Q}$ to denote the set of all equivalence classes of quadruples.

We now define a map from equivalence classes of quadruples into the subset $\mathcal{Y}$ of the universal algebra $\mathscr{U}_E^+(\mathbf{G})$.

Definition 4.2.2. *We define the map $\mathfrak{q} : \mathcal{Q} \to \mathcal{Y}$ as*

$$\mathfrak{q}([Q]) = Y_{(J^+,J^-,I_1,I_2)}(f,g)$$

where $(J^+, J^-, I_1, I_2) \in [Q]$, and f and g are normalised sections compactly supported in I_1 and I_2 respectively.

This map is well defined because as mentioned above, elements of $\mathcal{Y}$ only depend on the quadruple used to define them, and also the notion of equivalence in definition 4.2.1 has the property that equivalent quadruples are mapped by $\mathfrak{q}$ to the same element of $\mathcal{Y}$. To see this we note that quadruples related by i) give the same algebra element, due to the cocone property of the universal algebra, and quadruples related by ii) give the same algebra element due to corollary 4.1.7. The map $\mathfrak{q} : \mathcal{Q} \to \mathcal{Y}$ is also surjective, by virtue of definition 4.1.3.

We now switch our attention from $\mathcal{Y}$ to $\mathcal{Q}$, and will establish a relation between $\mathcal{Q}$ and the topology of $\mathcal{M}$. In order to do this we first note that each quadruple resembles a loop; there are two segments represented by J^+ and J^-, and they are stitched together at I_1 and I_2 (see figure 4.1). Before we can make this association of loops to quadruples precise, we need to make a choice of equivalence classes of loops that we will associate to each quadruple. These equivalence classes of loops should be chosen so that quadruples in the same equivalence class in $\mathcal{Q}$, map to the same equivalence class of loops. For this purpose we introduce the following set.

Definition 4.2.3. *Let $\widetilde{\pi}_1(\mathcal{M})$ denote the set of equivalence classes of continuous functions $\gamma : S^1 \to \mathcal{M}$ where equivalence is given by a combination of free homotopy equivalence (need not preserve base-point) and identifying loops with opposite orientations.*

Our choice of notation for $\widetilde{\pi}_1(\mathcal{M})$ stems from the fact that it is closely related to the fundamental group $\pi_1(\mathcal{M})$, although we have not placed a group structure on $\widetilde{\pi}_1(\mathcal{M})$ yet. As we shall see, two quadruples related by ii) in definition 4.2.1 will have the same loop associated to them but with opposite orientations. We therefore identify loops with opposite orientations so that equivalent quadruples are associated to equivalent loops (see lemma 4.2.7). We now define a set of curves associated to a given quadruple, and then prove the set is non-empty and defines a unique homotopy class.

Definition 4.2.4. *A (J^+, J^-, I_1, I_2)-curve is a continuous map $\gamma : S^1 \to \mathcal{M}$ such that γ can be decomposed as*

$$\begin{aligned}&\gamma = p_5 * p_4 * p_3 * p_2 * p_1 \quad \text{with} \\ &\text{imag}(p_1) \subset I_1 \quad \text{imag}(p_2) \subset J^+ \quad \text{imag}(p_3) \subset I_2 \quad \text{imag}(p_4) \subset J^- \quad \text{imag}(p_5) \subset I_1\end{aligned}$$

and we refer to the above decomposition as a (J^+, J^-, I_1, I_2)-decomposition of γ.

Proposition 4.2.5. *For any quadruple (J^+, J^-, I_1, I_2), there exists a curve γ which is a (J^+, J^-, I_1, I_2)-curve, and any two (J^+, J^-, I_1, I_2)-curves are homotopic.*

Proof. The existence of (J^+, J^-, I_1, I_2)-curves follows from the fact that each quadruple region is path-connected, which combined with the fact that $I_1 \cup I_2 \subset J^+ \cap J^-$ ensures that any point in I_1 or I_2 can be connected by a continuous path to any point in J^+ or J^-.

Now consider two (J^+, J^-, I_1, I_2) curves f and g with the following (J^+, J^-, I_1, I_2)-decompositions:

$$\begin{aligned} f &= p_5 * p_4 * p_3 * p_2 * p_1 \\ g &= q_5 * q_4 * q_3 * q_2 * q_1 \quad . \end{aligned}$$

Let δ_1 be a path from $p_1(0)$ to $q_1(0)$ with image in I_1, δ_2 be a path from $p_3(0)$ to $q_3(0)$ with image in I_2 and δ_3 be a path from $p_5(0)$ to $q_5(0)$ with image in I_1. These δ paths exist because each region of the quadruple is path-connected. As multiplication of paths is associative up to homotopy, we may suppress the notation concerning the grouping of multiplication. First we consider the following path

$$\overline{\delta}_1 * \overline{q}_1 * \overline{q}_2 * \delta_2 * p_2 * p_1$$

where the bar notation refers to a path being reversed i.e, $\overline{p}(t) = p(1-t)$. The above path is a closed loop contained in $I_1 \cup J^+ = J^+$. Since J^+ is contractible, this loop must be contractible to a point. This implies

$$p_2 * p_1 \sim_{\mathrm{rel}\{0,1\}} \overline{\delta}_2 * q_2 * q_1 * \delta_1$$

where $\sim_{\mathrm{rel}\{0,1\}}$ means there is an end point fixing homotopy between the two paths. We can then repeat this argument to find

$$\begin{aligned} p_4 * p_3 &\sim_{\mathrm{rel}\{0,1\}} \overline{\delta}_3 * q_4 * q_3 * \delta_2 \\ p_5 &\sim_{\mathrm{rel}\{0,1\}} \overline{\delta}_1 * q_5 * \delta_3 \end{aligned}$$

since these paths are confined to the contractible sets J^- and I_1 respectively. We can now use the fact that given paths a_1, a_2, b_1 and b_2 such that $a_1 \sim_{\mathrm{rel}\{0,1\}} b_1$ and $a_2 \sim_{\mathrm{rel}\{0,1\}} b_2$ this implies $a_1 * a_2 \sim_{\mathrm{rel}\{0,1\}} b_1 * b_2$. Using this we find

$$\begin{aligned} f &\sim p_5 * (p_4 * p_3) * (p_2 * p_1) \sim (\overline{\delta}_1 * q_5 * \delta_3) * (\overline{\delta}_3 * q_4 * q_3 * \delta_2) * (\overline{\delta}_2 * q_2 * q_1 * \delta_1) \\ &\sim \delta_1 * g * \overline{\delta}_1 \sim g \end{aligned}$$

where the final homotopy is a non base-point preserving homotopy. We have therefore shown that any two (J^+, J^-, I_1, I_2) curves are homotopic. □

We now define a map from $\mathcal{Q}$ to $\widetilde{\pi}_1(\mathcal{M})$ which we prove is well-defined in the subsequent lemma.

Definition 4.2.6. *The map $\mathcal{L} : \mathcal{Q} \to \widetilde{\pi}_1(\mathcal{M})$ maps each equivalence class $[Q] \in \mathcal{Q}$ to the equivalence class $[g] \in \widetilde{\pi}_1(\mathcal{M})$ where g is a Q-curve.*

Lemma 4.2.7. *The map $\mathcal{L} : \mathcal{Q} \to \widetilde{\pi}_1(\mathcal{M})$ is well defined.*

Proof. Let q and Q be quadruples related by i) in definition 4.2.1, so that the regions of q are contained in the regions of Q. Then clearly any q-curve will also be a Q-curve, and since any two Q-curves are homotopic by proposition 4.2.5, this implies q-curves and Q-curves belong to the same equivalence class in $\widetilde{\pi}_1(\mathcal{M})$.

Let Q and $\overline{Q}$ be quadruples related by ii) in definition 4.2.1, and let g be a Q-curve and f be a $\overline{Q}$ curve. This implies $\overline{f}$ will also be a Q-curve, which by proposition 4.2.5 implies $\overline{f}$ is homotopic to g, and therefore f and g belong to the same equivalence class in $\widetilde{\pi}_1(\mathcal{M})$. $\square$

We now have the following pair of maps

$$\mathcal{Y} \xleftarrow{\mathfrak{q}} \mathcal{Q} \xrightarrow{\mathcal{L}} \widetilde{\pi}_1(\mathcal{M}) \quad .$$

Our strategy for relating the spacetime topology to $\mathscr{U}_E^+(\mathbf{G})$ will be to show that $\mathcal{L}$ is a bijection, and use its inverse to obtain a map from $\widetilde{\pi}_1(\mathcal{M})$ to $\mathcal{Y}$. In order to do this we will need to introduce some geometrical techniques.

4.3 Lorentzian geometry techniques for relating quadruples

The geometrical techniques introduced in this section will be used to prove injectivity of the map $\mathcal{L}$ in the next section. To prove injectivity of $\mathcal{L}$, we need to prove $\mathcal{L}([Q]) = \mathcal{L}([q]) \Rightarrow [Q] = [q]$. This is done by finding a sequence of equivalent quadruples linking $Q \in [Q]$ and $q \in [q]$. Since we are free to pick any representatives of $[Q]$ and $[q]$ for our map $\mathcal{L}$, we should choose convenient ones. In this section we show that every equivalence class $[Q]$ has a representative whose regions are Cauchy developments of subregions of an arbitrary Cauchy surface. This allows us to compare $[Q]$ and $[q]$ using such representatives based on a common Cauchy surface, thereby ensuring each region of the intermediate representatives is causally convex (a condition necessary for them to be $\mathsf{Loc}_C^{\mathcal{M}}$ regions).

One of the main techniques we will use in this section is to expand Cauchy balls (see definition 3.2.1) so that their Cauchy development (see definition 1.3.7) contains the original Cauchy ball translated in time, which will be used to transport quadruples towards a Cauchy surface. We first define how we will expand Cauchy balls and then prove a result about this expansion.

Definition 4.3.1. *Given a Cauchy ball B of Σ, we define the open ball* $\mathrm{Ball}(B, \delta)$ *containing B to be the set $\{\sigma \in \Sigma \mid \inf_{\widetilde{\sigma} \in B} d_k(\sigma, \widetilde{\sigma}) < \delta\}$, where d_k is the instantaneous optical metric on Σ (see definition 1.3.11).*

Lemma 4.3.2. *Given a Cauchy ball B of Σ, there exists $\delta > 0$ such that* $\mathrm{Ball}(B, \delta)$ *is contained in a Cauchy ball B' of Σ.*

Proof. Let (U, ψ) be the chart which makes B a Euclidean ball so that $\psi(B) = \{x \in \mathbb{R}^n \ : \ d_E(0, x) < r\}$ where 0 is the origin of the ball and d_E is the Euclidean metric. Since the closure of B is contained in U, there exists some $\epsilon > 0$ such that $B' := \psi^{-1}(\{x \in \mathbb{R}^n \ : \ d_E(0, x) < r + \epsilon\}) \subset U$. The topology of Σ is equivalent to the topology induced by the instantaneous optical metric d_k [Lee03, Theorem 13.29], hence ψ can also be regarded as a continuous function between the metric spaces $(\overline{B'}, d_k)$ and $\big(\psi(\overline{B'}), d_E\big)$. Since $\overline{B'}$ is compact (U is relatively compact) we can use the Heine-Cantor theorem[1] [Rud86, Theorem 4.19] to further assert that ψ, as a mapping between metric spaces is uniformly continuous. Therefore given ϵ as above we can find $\delta > 0$ such that

[1]A continuous function between two metric spaces with compact domain space is uniformly continuous.

the open δ-ball (in the d_k metric) about any point $x \in \overline{B}$ gets mapped inside the ϵ-ball (in the d_E metric) about $\psi(x)$. This implies $\mathrm{Ball}(B, \delta) \subset B'$ which proves the lemma. $\square$

Quadruples can be related by a chain of embeddings, which is defined as follows.

Definition 4.3.3. *A chain of embeddings is a sequence of $\mathsf{Loc}_C^{\mathcal{M}}$ regions such that for each pair of regions next to each other in the sequence, one is a subset of the other.*

We can now establish a lemma which we subsequently use to prove a result which will allow us to find chains of embeddings relating regions that differ in their time coordinate.

Lemma 4.3.4. *Let $\widetilde{\Sigma}$ be a relatively compact subset of a Cauchy surface Σ. Then there exists $\delta > 0$ such that $\mathrm{Ball}(\widetilde{\Sigma}, \delta)$ is relatively compact.*

Proof. The radius of the largest ball B about the origin in $T_p\Sigma$ such that the exponential map restricted to B is a diffeomorphism is called the injectivity radius at p. The minimum of the injectivity radii over all points in $\widetilde{\Sigma}$, which we denote by 2δ, is positive since $\widetilde{\Sigma}$ is relatively compact [Kli11, Proposition 2.1.10]. By compactness of $\overline{\widetilde{\Sigma}}$, we can cover $\widetilde{\Sigma}$ by a finite set $\{B(p_n, \delta) | n \leq N\}$ where $p_n \in \widetilde{\Sigma}$ and $B(p_n, \delta)$ is the image of the exponential map applied to a ball in $T_{p_n}\widetilde{\Sigma}$ centered on the origin of radius δ. We then see that $\{B(p_n, 2\delta) | n \leq N\}$ forms a finite cover of $\mathrm{Ball}(\widetilde{\Sigma}, \delta)$ by relatively compact sets, since any point in $\mathrm{Ball}(\widetilde{\Sigma}, \delta)$ is within a distance of 2δ from at least one of the points of $\{p_n | n \leq N\}$. A union of finitely many relatively compact sets is compact, hence $\mathrm{Ball}(\widetilde{\Sigma}, \delta)$ is relatively compact.

$\square$

Lemma 4.3.5. *Let $\mathcal{M}$ be a globally hyperbolic spacetime in standard form with manifold $\mathbb{R} \times \Sigma$, with relatively compact $\widetilde{\Sigma} \subset \Sigma$. Then for any compact subset $S \subset \mathbb{R}$ and $\delta > 0$ such that $\mathrm{Ball}(\widetilde{\Sigma}, \delta)$ is compact (such a δ exists by lemma 4.3.4), there exists $\epsilon > 0$ such that*

$$t \in S,\, t' \in \mathbb{R} \ \text{ and } \ |t - t'| \leq \epsilon \quad \Rightarrow \quad \{t\} \times \widetilde{\Sigma} \subset D\big(\{t'\} \times \mathrm{Ball}(\widetilde{\Sigma}, \delta)\big) \quad . \tag{4.4}$$

Proof. We will use a slight generalisation of the proof of [Few15, Lemma 2.5], which proves the above result in the special case where S consists of a single point, and requires the additional constraint on δ that $\mathrm{Ball}(\widetilde{\Sigma}, \delta)$ is relatively compact and has non-empty exterior. The assumption of a non-empty exterior is needed for a subsequent result in [Few15, Lemma 2.5], but is unnecessary for our purposes.

Since we are concerned with Cauchy developments which depend only on the causal structure of $(\mathcal{M}, g)$, it will be convenient to replace g with the conformally related metric $\widetilde{g}$ in definition 1.3.11 which is of the form

$$\widetilde{g} = d\mathcal{T} \otimes d\mathcal{T} - k_\Sigma(\mathcal{T}) \quad . \tag{4.5}$$

This means that $\widetilde{g}$ restricted to each Cauchy surface $\mathcal{T}^{-1}(\tau)$ yields a family of metrics $k_\tau = -i_\tau^* \, g|_{\mathcal{T}^{-1}(\tau)}$ on Σ, where i_τ is the inclusion map $i_\tau : \Sigma \to \mathbb{R} \times \Sigma$ such that $i(\sigma) = (\tau, \sigma)$ for $\sigma \in \Sigma$.

Let $\hat{T}_{(\delta)}\widetilde{\Sigma}$ denote the subset of the tangent bundle $T\Sigma$ consisting of unit vectors (in the k_0 metric) in the tangent spaces of points in the closure of $\mathrm{Ball}(\widetilde{\Sigma},\delta)$. By supposition, δ has been chosen so that $\mathrm{Ball}(\widetilde{\Sigma},\delta)$ is relatively compact, hence $\hat{T}_{(\delta)}\widetilde{\Sigma}$ is compact. Using the fact that a continuous function defined on a compact domain has image contained in a compact set, which is therefore bounded, there exists a constant $C>1$ such that

$$\frac{k_t(v,v)}{k_\tau(v,v)} = \frac{k_t(\hat{v},\hat{v})}{k_\tau(\hat{v},\hat{v})} \leq C \quad \forall\,(t,\tau,\hat{v}) \in S \times S_\epsilon \times \hat{T}_{(\delta)}\widetilde{\Sigma} \tag{4.6}$$

where $\hat{v} = v/\sqrt{k_0(v,v)}$ and S_ϵ consists of all points with a distance $\leq \epsilon$ from S. To show that $\{t\}\times\widetilde{\Sigma} \subset D\big(\{t'\}\times\mathrm{Ball}(\widetilde{\Sigma},\delta)\big)$, we must show that any inextendible causal curve that intersects $\{t\}\times\widetilde{\Sigma}$ necessarily intersects $\{t'\}\times\mathrm{Ball}(\widetilde{\Sigma},\delta)$. We therefore consider a generic causal curve γ which we parameterise, in the coordinates used to express the metric in equation (4.5), as $\gamma(\tau) = (\tau,\sigma(\tau)) \in \mathbb{R}\times\Sigma$ where $\sigma(\tau)$ is a path traced out in Σ. In these coordinates, the causality of γ with respect to the metric $\widetilde{g}$ (which ensures causality with respect to g) implies

$$k_\tau(\dot{\sigma}(\tau),\dot{\sigma}(\tau)) \leq 1 \quad .$$

We further require $\sigma(t) \in \widetilde{\Sigma}$ so that γ intersects $\{t\}\times\widetilde{\Sigma}$. Without loss of generality we assume $t'>t$, and let $\widetilde{t}$ be the first point after t such that σ leaves $\mathrm{Ball}(\widetilde{\Sigma},\delta)$, which implies $\sigma(\widetilde{t})$ is in the closure of $\mathrm{Ball}(\widetilde{\Sigma},\delta)$. We can then apply the bound obtained in equation (4.6) to get

$$d_{k_t}(\sigma(t),\sigma(\widetilde{t}\,)) \leq \int_t^{\widetilde{t}} d\tau\ \sqrt{k_t(\dot{\sigma}(\tau),\dot{\sigma}(\tau))} \leq \int_t^{\widetilde{t}} d\tau\ \sqrt{C}\sqrt{k_\tau(\dot{\sigma}(\tau),\dot{\sigma}(\tau))} \leq \sqrt{C}|t-\widetilde{t}|$$

where the first inequality becomes an equality if σ is a geodesic, and the third inequality is obtained from the causality of γ.

We now choose ϵ such that $\epsilon < \delta/\sqrt{C}$. If $\widetilde{t} \leq t'$, this implies $d_{k_t}(\sigma(t),\sigma(\widetilde{t}\,)) \leq \sqrt{C}\epsilon < \delta$, hence $\sigma(\widetilde{t}) \in \mathrm{Ball}(\widetilde{\Sigma},\delta)$ which is a contradiction, therefore $\widetilde{t} > t'$. This means that $\sigma(t') \in \mathrm{Ball}(\widetilde{\Sigma},\delta)$, and therefore $\epsilon < \delta/\sqrt{C}$ satisfies the conditions of the lemma. $\square$

We want to find a chain of embeddings through $\mathsf{Loc}_C^{\mathcal{M}}$ regions relating a causally convex neighbourhood of a subset of one Cauchy surface to a causally convex neighbourhood of a subset of another Cauchy surface. We therefore need to find a Cauchy temporal function (see definition 1.3.10) which has both Cauchy surfaces as level sets, since then the region just needs to be translated along the time coordinate given by the temporal function. We therefore prove the following lemma.

Lemma 4.3.6. *Given two Cauchy surfaces Σ and $\widetilde{\Sigma} \subset I^-(\Sigma)$ with disjoint neighbourhoods, there exists a Cauchy temporal function $\mathcal{T}$ such that $\mathcal{T}^{-1}(0) = \widetilde{\Sigma}$ and $\mathcal{T}^{-1}(1) = \Sigma$.*

This result is a modification of the result [BS06, Theorem 1.2], which states that for any given Cauchy surface, one can find a Cauchy temporal function such that the given Cauchy surface appears as a level set of the function. In [BW15], the authors state (on page 21) that "A minor modification of the proof also shows that one can prescribe two disjoint Cauchy hypersurfaces as level sets", referring to the proof of [BS06, Theorem 1.2].

Rather than just stating the above lemma is true, we will give an outline of a proof using the tools developed in [BS06].

Sketch Proof. The regions $I^+(\Sigma), I^-(\Sigma), I^+(\widetilde{\Sigma})$ and $I^-(\widetilde{\Sigma})$ are all globally hyperbolic, and hence have onto Cauchy temporal functions (meaning the range of each function is the whole real line) $T^+_\Sigma, T^-_\Sigma, T^+_{\widetilde{\Sigma}}$ and $T^-_{\widetilde{\Sigma}}$ respectively. With these we can construct the following time function

$$\mathcal{T}(p) = \begin{cases} 1 + \exp\big(T^+_\Sigma(p)\big) & p \in I^+(\Sigma) \\ 1 & p \in \Sigma \\ \Big(1 + \exp\Big(-T^-_\Sigma(p) - T^+_{\widetilde{\Sigma}}(p)\Big)\Big)^{-1} & p \in I^-(\Sigma) \cap I^+(\widetilde{\Sigma}) \\ 0 & p \in \widetilde{\Sigma} \\ -\exp\big(-T^+_\Sigma(p)\big) & p \in I^-(\widetilde{\Sigma}) \end{cases}$$

which is inspired by the choice of time function in [BS06, Proposition 5.17], and can be seen to be continuous by using the methods in [BS06, Proposition 5.17]. The above time function is almost a Cauchy temporal function, except for the fact that it might not be smooth at Σ and $\widetilde{\Sigma}$. This can be dealt with by smoothing out $\mathcal{T}$ in disjoint neighbourhoods of Σ and $\widetilde{\Sigma}$ (see proof of [BS06, Theorem 5.15]) and then further modifying it within disjoint neighbourhoods of Σ and $\widetilde{\Sigma}$ to a function $\widetilde{\mathcal{T}}$ that is a Cauchy temporal function (see proof of [BS06, Proposition 6.20]), all whilst maintaining the property that $\widetilde{\mathcal{T}}^{-1}(0) = \widetilde{\Sigma}$ and $\widetilde{\mathcal{T}}^{-1}(1) = \Sigma$. The function $\widetilde{\mathcal{T}}$ therefore satisfies the requirements of the lemma. □

We say that a quadruple is localised about a Cauchy surface Σ, if each of the regions of the quadruple are Cauchy developments of Cauchy balls of Σ. We now have the tools in place to prove one of the main results of this section, which can be used to relate a quadruple localised about one Cauchy surface to a quadruple localised about any other Cauchy surface.

Proposition 4.3.7. *Consider a $\mathsf{Loc}^{\mathcal{M}}_C$ region $\mathcal{O}$ with a Cauchy surface $\Sigma_{\mathcal{O}}$ that is a Cauchy ball of some acausal spacelike hypersurface Σ. Given any Cauchy surface $\widetilde{\Sigma}$, there exists a chain of embeddings of $\mathsf{Loc}^{\mathcal{M}}_C$ regions between $\mathcal{O}$ and the Cauchy development of a Cauchy ball of $\widetilde{\Sigma}$.*

Proof. First we assume that Σ and $\widetilde{\Sigma}$ are disjoint, and without loss of generality further assume that $\Sigma_{\mathcal{O}} \subset I^-(\widetilde{\Sigma})$. The closure of $\Sigma_{\mathcal{O}}$ is a compact acausal spacelike hypersurface with boundary, since $\Sigma_{\mathcal{O}}$ is a Cauchy ball of Σ. This means we can use [BS06, Theorem 1.1] to find a Cauchy surface Σ' which contains $\Sigma_{\mathcal{O}}$. We then use lemma 4.3.6 to find a Cauchy temporal function $\mathcal{T}$ such that $\mathcal{T}^{-1}(0) = \Sigma'$ and $\mathcal{T}^{-1}(1) = \widetilde{\Sigma}$. We will use $\mathcal{T}$-foliation coordinates for our spacetime so that we can regard it as $\mathbb{R} \times \Sigma$.

We then use lemma 4.3.2 to find some $\delta > 0$ such that $\mathrm{Ball}(\Sigma_{\mathcal{O}}, \delta)$ is contained in a Cauchy ball B of Σ'. We define $\mathcal{O}_t := D\big(\{t\} \times \Sigma_{\mathcal{O}}\big)$ and $\widetilde{\mathcal{O}}_t := D\big(\{t\} \times B\big)$. Given δ, we use lemma 4.3.5 find an $\epsilon > 0$ such that $\mathcal{O}_t \subset \widetilde{\mathcal{O}}_{t+\epsilon}$ for all $t \in [0,1]$, and choose an integer $N > 1/\epsilon$. This gives us the following chain of embeddings of $\mathsf{Loc}^{\mathcal{M}}_C$ regions

$$\mathcal{O} = \mathcal{O}_0 \subset \widetilde{\mathcal{O}}_{1/N} \supset \mathcal{O}_{1/N} \subset \ldots \supset \mathcal{O}_{(N-1)/N} \subset \widetilde{\mathcal{O}}_1 \supset \mathcal{O}_1$$

with $\mathcal{O}_1 = D\big(\{1\} \times \Sigma_{\mathcal{O}}\big) \subset \mathcal{T}^{-1}(1) = \widetilde{\Sigma}$. Moreover, the Cauchy surface $\{1\} \times \Sigma_{\mathcal{O}}$ of $\mathcal{O}_1$ is a Cauchy ball, since it is the time translation of a Cauchy ball and time translation defines a diffeomorphism between Cauchy surfaces. We have therefore proved the lemma in the case that Σ and $\widetilde{\Sigma}$ are disjoint.

If Σ and $\widetilde{\Sigma}$ aren't disjoint, we can use an argument adapted from [BW15, Corollary 18]. We consider $I^+(\Sigma') \cap I^+(\widetilde{\Sigma})$, which is globally hyperbolic and therefore has a Cauchy surface $\widetilde{\Sigma}'$. Since $\widetilde{\Sigma}'$ is disjoint from Σ' and $\widetilde{\Sigma}$, we can repeat the above steps to get a chain of embeddings from $\mathcal{O}$ to $D(B')$ with $B' \subset \widetilde{\Sigma}'$ and another chain of embeddings from $D(B')$ to $D(B'')$ with $B'' \subset \widetilde{\Sigma}$. Combining these two chains gives a single chain of embeddings and completes the proof of the proposition. □

We note that by choosing N large enough in the above proof, we can apply the same strategy to construct chains of embeddings simultaneously for finitely many Cauchy balls. This therefore gives us the means of relating a quadruple localised about one Cauchy surface to a quadruple localised about any other Cauchy surface. We therefore need to prove that for every quadruple Q, there exists a Cauchy surface Σ such that Q can be related to a quadruple which is localised about Σ. We now prove this for a subset of quadruples, and will later prove that all quadruples can be related to a quadruple in this subset.

Proposition 4.3.8. *Consider a quadruple (J^+, J^-, I_1, I_2) where J^+ and J^- have Cauchy surfaces Σ^+ and Σ^- which are Cauchy balls of Σ_0 and Σ_1 respectively, $I_1 = D(\Sigma_{I_1})$ and $I_2 = D(\Sigma_{I_2})$ with $\Sigma_{I_1} \cup \Sigma_{I_2} \subset \Sigma^+ \cap \Sigma^-$. This quadruple is in the same equivalence class as $(\widetilde{J}^+, J^-, I_1, I_2)$ where $\widetilde{J}^+$ and J^- are Cauchy developments of Cauchy balls of a common Cauchy surface.*

Proof. The structure of the proof is as follows: we will define a function that deforms subsets of Σ_0 to Σ_1, but leaves points in $\Sigma^+ \cap \Sigma^-$ invariant. This function will then be used to gradually lift Σ^+ to a Cauchy ball of Σ_1, which is done in sufficiently small steps so that the Cauchy developments of the intermediate Cauchy balls form a chain of embeddings. Since the function leaves $\Sigma^+ \cap \Sigma^-$ invariant, the intersection of J^- with each of the intermediate $\mathsf{Loc}_C^{\mathcal{M}}$ regions representing the shift of J^+ will contain $I_1 \cup I_2$. We therefore get a chain of equivalent quadruples relating our original quadruple to one whose regions are all Cauchy developments of Cauchy balls of Σ_1.

We begin by using [BS06, Theorem 1.2] to find Cauchy temporal functions $\mathcal{T}_0$ and $\mathcal{T}_1$ such that $\mathcal{T}_0^{-1}(0) = \Sigma_0$ and $\mathcal{T}_1^{-1}(0) = \Sigma_1$. These Cauchy temporal functions have future-directed timelike tangents, and are therefore strictly increasing along future-directed timelike curves. We then consider the convex combination

$$\mathcal{T}_\lambda = \lambda \mathcal{T}_1 + (1 - \lambda)\mathcal{T}_0$$

with $\lambda \in [0, 1]$. We define $\Sigma_\lambda = \mathcal{T}_\lambda^{-1}(0)$ and note that the Cauchy surfaces defined by $\lambda = 0$ and 1 are equivalent to the original Σ_0 and Σ_1 we started with. We also note that $\lambda \mapsto \Sigma_\lambda$ is not intended to be a foliation, but rather gives a set of Cauchy surfaces that interpolate between Σ_0 and Σ_1. Each Σ_λ is a smooth embedded submanifold since level sets of smooth submersions are smooth embedded submanifolds [Hir76, Theorem 1.3.2].

We also see that every inextendible timelike curve intersects Σ_λ exactly once. This is because any inextendible timelike curve γ crosses both Σ_0 and Σ_1 exactly once. This means that sufficiently to the past along γ, $\mathcal{T}_0$ and $\mathcal{T}_1$ are both negative while sufficiently to the future, they are both positive. The same is therefore true of $\mathcal{T}_\lambda$, and by continuity this implies that the timelike curve crosses Σ_λ, and it also crosses exactly once because a convex combination of monotonic functions is monotonic. We therefore see that Σ_λ is a smooth Cauchy surface of $\mathcal{M}$.

From now on we will use $\mathcal{T}_0$-foliation coordinates for $\mathcal{M}$ so that we can regard $\mathcal{M}$ as $\mathbb{R} \times \Sigma_0$, and we replace the metric g on $\mathcal{M}$ with the conformally related metric $\widetilde{g}$ given by replacing $\mathcal{T}$ with $\mathcal{T}_0$ in equation (4.5). The metric $\widetilde{g}$ induces a family of instantaneous optical metrics k_τ on Σ_0 (as described below equation (4.5)). We also define the map $\pi_{\Sigma_0} : \mathcal{M} \to \Sigma_0$ as the projection of $\mathcal{M} = \mathbb{R} \times \Sigma_0$ on to the second component.

We now define a function $\alpha : [0,1] \times \Sigma_0 \to \mathbb{R} \times \Sigma_0$ such that $\alpha(\lambda, \cdot)$ drags points in Σ_0 to Σ_λ, and will be used to deform the Cauchy ball Σ^+ of Σ_0 to a Cauchy ball of Σ_1. This is done by solving the following implicit equation for $t : \mathbb{R} \times \Sigma_0 \to \mathbb{R}$

$$\mathcal{F}\big(\lambda, \sigma, t(\lambda,\sigma)\big) := \mathcal{T}_\lambda\big(t(\lambda,\sigma), \sigma\big) = 0 \tag{4.7}$$

and then defining α in terms of the function t as follows

$$\alpha(\lambda,\sigma) = (t(\lambda,\sigma), \sigma) \quad , \tag{4.8}$$

noting that $\alpha(\lambda,\sigma) \in \Sigma_\lambda$ by virtue of equation (4.7). For any $(\lambda_*, \sigma_*) \in [0,1] \times \Sigma_0$, we can use the implicit function theorem to find a unique smooth solution $t(\lambda,\sigma)$ for equation (4.7) in a neighbourhood of (λ_*, σ_*), since $\partial\mathcal{F}/\partial t$ is non-zero everywhere ($\mathcal{T}_\lambda$ is a temporal function). Any two solutions must agree on the overlap of their domains, otherwise there would be two points in Σ_λ which only differ in the time coordinate, and would therefore be connected by a timelike curve, violating the achronality of Σ_λ. Since the unique smooth local solutions to equation (4.7) must agree on their overlap, we get a unique smooth global solution defined on $[0,1] \times \Sigma_0$.

We use the smooth global solution for t to define $\alpha : \Sigma_0 \to \mathbb{R} \times \Sigma_0$ using equation (4.8), which for fixed λ is a smooth map between Σ_0 and Σ_λ that is an injective immersion because $\pi_{\Sigma_0} \circ \alpha(\lambda, \cdot)$ is the identity map on Σ_0. We see that α is constant with respect to λ on $[0,1] \times \pi_{\Sigma_0}(\Sigma^+ \cap \Sigma^-)$, because $\Sigma^+ \cap \Sigma^- \subset \Sigma_0 \cap \Sigma_1$ and $\mathcal{T}_0$ and $\mathcal{T}_1$ both vanish on $\Sigma_0 \cap \Sigma_1$, meaning the implicit equation for t in equation (4.7) is satisfied by $t(\lambda,\sigma) = 0$ for all $(\lambda,\sigma) \in [0,1] \times \pi_{\Sigma_0}(\Sigma^+ \cap \Sigma^-)$. This means as we deform Σ_0 into Σ_1 by varying λ, the points in $\Sigma_{I_1} \cup \Sigma_{I_2} \subset \Sigma^+ \cap \Sigma^-$ are left invariant.

We now use lemma 4.3.2 to find a Cauchy ball $\widetilde{\Sigma}^+ \supset \mathrm{Ball}(\Sigma^+, \delta)$ of Σ_0 for some $\delta > 0$. We define $\Sigma^+_\lambda = \alpha(\lambda, \pi_{\Sigma_0}(\Sigma^+))$ and $\widetilde{\Sigma}^+_\lambda = \alpha(\lambda, \pi_{\Sigma_0}(\widetilde{\Sigma}^+))$ which are subsets of Σ_λ, and note that $\Sigma^+ = \Sigma^+_0$. Since $\alpha(\lambda, \cdot)$ is an injective immersion, it is a diffeomorphism onto its image and therefore preserves the property of being a Cauchy ball, hence Σ^+_λ and $\widetilde{\Sigma}^+_\lambda$ are Cauchy balls for all $\lambda \in [0,1]$.

We now want to relate $J^+ = D(\Sigma^+_0)$ to $D(\Sigma^+_1)$ by a chain of embeddings of $\mathsf{Loc}^{\mathcal{M}}_C$ regions by working in steps, finding sufficiently small $\Delta\lambda$ such that $D(\Sigma^+_\lambda) \subset D(\widetilde{\Sigma}^+_{\lambda+\Delta\lambda})$ for all $\lambda \in [0, 1-\Delta\lambda]$. We do this by finding bounds that quantify how much Σ^+_λ can expand as we vary λ. For this purpose we consider the region where the dragging of points from $\widetilde{\Sigma}^+$ to Σ_1 occurs, which is given by

$$\widetilde{\mathcal{M}} := \bigcup_{\lambda\in[0,1]} \widetilde{\Sigma}^+_\lambda \subset \left(J^+\left(\overline{\widetilde{\Sigma}^+}\right) \cap J^-(\Sigma_1)\right) \cup \left(J^-\left(\overline{\widetilde{\Sigma}^+}\right) \cap J^+(\Sigma_1)\right) \quad .$$

The set on the right-hand side is compact by virtue of corollary 1.3.19, hence $\widetilde{\mathcal{M}}$ is relatively compact. This implies that the projection of $\widetilde{\mathcal{M}}$ to Σ_0 given by $\Sigma_0' := \pi_{\Sigma_0}(\widetilde{\mathcal{M}})$ is also relatively compact. Given a point $\sigma \in \Sigma_0$ and tangent vector $v \in T_\sigma\Sigma_0$, the pushforward of v to $T_{\alpha(\lambda,\sigma)}\widetilde{\mathcal{M}}$ by $\alpha(\lambda,\cdot)$ is given by

$$v_\lambda := d\alpha(\lambda,\cdot)v = (\nabla_v t)(\lambda,\sigma)\frac{\partial}{\partial \mathcal{T}_0} + v$$

where $\nabla_v t$ is the derivative of the scalar function t along v. The first term arises due to the fact that the $\mathcal{T}_0$ component of $\alpha(\lambda,\sigma)$ is given by $t(\lambda,\sigma)$. We see that v_λ must be spacelike, since it is a tangent vector to the Cauchy surface Σ_λ. This gives us the following inequality

$$\begin{aligned} \widetilde{g}(v_\lambda, v_\lambda) = (\nabla_v t)(\lambda,\sigma)^2 - \ k_{t(\lambda,\sigma)}(v,v) < 0 \quad \Rightarrow \\ \sqrt{k_{t(\lambda,\sigma)}(v,v)} > |(\nabla_v t)(\lambda,\sigma)| \end{aligned} \tag{4.9}$$

where the inequality is strict, since otherwise v_λ would be null. Using the fact that a continuous function defined on a relatively compact domain has image contained in a compact set, which is therefore bounded, there exists a constant $C_1 > 1$ such that

$$\frac{|(\nabla_v t)(\lambda,\sigma)|}{\sqrt{k_{t(\lambda,\sigma)}(v,v)}} = \frac{|(\nabla_{\hat{v}} t)(\lambda,\sigma)|}{\sqrt{k_{t(\lambda,\sigma)}(\hat{v},\hat{v})}} \leq C_1 \quad \forall\ (\lambda,\sigma) \in [0,1]\times\Sigma_0', \quad v \in T_\sigma\Sigma_0' \tag{4.10}$$

where $\hat{v} = v/\sqrt{k_0(v,v)}$. Equation (4.9) implies that C_1 may be chosen to be < 1. Similarly there exists a constant C_2 such that

$$\frac{\sqrt{k_\tau(v,v)}}{\sqrt{k_{\tau'}(v,v)}} \leq C_2 \quad \forall\ (\tau,\tau') \in \Big[\inf_{p\in\widetilde{\mathcal{M}}} \mathcal{T}_0(p), \sup_{q\in\widetilde{\mathcal{M}}} \mathcal{T}_0(q)\Big], \quad v \in T_\sigma\Sigma_0' \quad . \tag{4.11}$$

To show that $D(\Sigma^+_\lambda) \subset D(\widetilde{\Sigma}^+_{\lambda+\Delta\lambda})$ for a given $\Delta\lambda$, we must show that any inextendible causal curve from Σ^+_λ intersects $\widetilde{\Sigma}^+_{\lambda+\Delta\lambda}$. We fix $\Delta\lambda > 0$ and let γ be some arbitrary inextendible causal curve that intersects Σ^+_λ. We can parametrise γ within $\widetilde{\mathcal{M}}$ so that $\gamma(\mu) = \alpha(\mu,\sigma(\mu))$ for $\mu \in [\lambda,\lambda+\Delta\lambda]$, where $\sigma(\mu)$ is a curve in Σ_0. The fact that γ intersects Σ^+_λ implies $\sigma(\lambda) \in \pi_{\Sigma_0}(\Sigma^+)$. Proving γ intersects $\widetilde{\Sigma}^+_{\lambda+\Delta\lambda}$ is equivalent to proving $\sigma(\lambda+\Delta\lambda) \in \pi_{\Sigma_0}(\widetilde{\Sigma}^+)$. The causality of γ implies

$$\sqrt{k_{t(\mu,\sigma(\mu))}\big(\dot{\sigma}(\mu),\dot{\sigma}(\mu)\big)} \leq d\mathcal{T}_0(\dot{\gamma}(\mu)) = \frac{\partial t}{\partial\lambda}(\mu,\sigma(\mu)) + (\nabla_{\dot{\sigma}(\mu)}t)(\mu,\sigma(\mu))$$

and this can be combined with the bound we have in equation (4.10) to get

$$\begin{aligned}\sqrt{k_{t(\mu,\sigma(\mu))}\big(\dot\sigma(\mu),\dot\sigma(\mu)\big)} &\le \frac{\partial t}{\partial\lambda}(\mu,\sigma(\mu)) + C_1\sqrt{k_{t(\mu,\sigma(\mu))}\big(\dot\sigma(\mu),\dot\sigma(\mu)\big)} \quad \Rightarrow \\ \sqrt{k_{t(\mu,\sigma(\mu))}\big(\dot\sigma(\mu),\dot\sigma(\mu)\big)} &\le \frac{1}{1-C_1}\frac{\partial t}{\partial\lambda}(\mu,\sigma(\mu)) \le \frac{C_3}{1-C_1}\end{aligned}$$

where C_3 is the supremum of $\partial t/\partial\lambda$ over $[0,1]\times\Sigma_0'$, which exists since Σ_0' is relatively compact. In order to prove $\sigma(\lambda+\Delta\lambda)\in\pi_{\Sigma_0}(\widetilde\Sigma^+)$ (which proves $D(\Sigma^+_\lambda)\subset D(\widetilde\Sigma^+_{\lambda+\Delta\lambda})$ as mentioned above), it is sufficient to show that the distance defined by the instantaneous optical metric k_0 on Σ_0 between $\sigma(\lambda+\Delta\lambda)$ and $\sigma(\lambda)$ is less than δ, since $\widetilde\Sigma^+\supset\mathrm{Ball}(\Sigma^+,\delta)$. This can be done as follows

$$\begin{aligned}&d_{k_0}\big(\sigma(\lambda),\sigma(\lambda+\Delta\lambda)\big)\le\int_\lambda^{\lambda+\Delta\lambda}d\mu\,\sqrt{k_0\big(\dot\sigma(\mu),\dot\sigma(\mu)\big)}\\ &\le\Delta\lambda\sup_{\mu\in[\lambda,\lambda+\Delta\lambda]}\sqrt{k_0\big(\dot\sigma(\mu),\dot\sigma(\mu)\big)}\le\Delta\lambda\;C_2\sup_{\mu\in[\lambda,\lambda+\Delta\lambda]}\sqrt{k_{t(\mu,\sigma(\mu))}\big(\dot\sigma(\mu),\dot\sigma(\mu)\big)}\le\Delta\lambda\frac{C_2C_3}{1-C_1}\end{aligned}$$

and we therefore find that

$$\Delta\lambda<\frac{1-C_1}{C_2C_3}\delta\quad\Rightarrow\quad D(\Sigma^+_\lambda)\subset D(\widetilde\Sigma^+_{\lambda+\Delta\lambda})\quad\forall\lambda\in[0,1-\Delta\lambda]$$

and since $C_1<1$, we find a positive integer N large enough so that $N>\frac{C_2C_3}{1-C_1}/\delta$. With this choice of N we get $D(\Sigma^+_{n/N})\subset D(\widetilde\Sigma^+_{(n+1)/N})$. As mentioned earlier in the proof, $\alpha(\lambda,\cdot)$ leaves $\Sigma_{I_1}\cup\Sigma_{I_2}\subset\Sigma^+\cap\Sigma^-$ invariant, so $\Sigma_{I_1}\cup\Sigma_{I_2}\subset\Sigma^+_\lambda\cap\Sigma^-$ for all $\lambda\in[0,1]$. This means that

$$I_1\cup I_2=D(\Sigma_{I_1}\cup\Sigma_{I_2})\subset D(\Sigma^+_\lambda\cap\Sigma^-)=D(\Sigma^+_\lambda)\cap J^-\subset D(\widetilde\Sigma^+_\lambda)\cap J^-$$

for all $\lambda\in[0,1]$. This implies $(D(\Sigma^+_\lambda),J^-,I_1,I_2)$ and $(D(\widetilde\Sigma^+_\lambda),J^-,I_1,I_2)$ are valid quadruples. We therefore get the following chain of equivalences

$$\begin{aligned}&(J^+,J^-,I_1,I_2)\sim(D(\Sigma^+_0),J^-,I_1,I_2)\sim(D(\widetilde\Sigma^+_{1/N}),J^-,I_1,I_2)\\ &\sim(D(\Sigma^+_{1/N}),J^-,I_1,I_2)\sim\ldots\sim(D(\Sigma^+_{(N-1)/N}),J^-,I_1,I_2)\sim(D(\widetilde\Sigma^+_1),J^-,I_1,I_2)\end{aligned}$$

and this final quadruple has regions which are all Cauchy developments of subregions of Σ_1. We can therefore see that $\widetilde J^+=D(\widetilde\Sigma^+_1)$ satisfies the conditions of the proposition. □

We can now prove the main result of this section, which will be used for the proof of injectivity of $\mathcal{L}$ in the next section.

Theorem 4.3.9. *Given any Cauchy surface Σ of $\mathcal{M}$, and any equivalence class of quadruples $[Q]$, there exists a representative $(\widetilde J^+,\widetilde J^-,\widetilde I_1,\widetilde I_2)\in[Q]$ such that each of its $\mathsf{Loc}_C^{\mathcal{M}}$ regions are Cauchy developments of subregions of Σ.*

Proof. We start with some arbitrary representative $(J^+,J^-,I_1,I_2)\in[Q]$ and construct a chain of equivalences to the desired $(\widetilde J^+,\widetilde J^-,\widetilde I_1,\widetilde I_2)$ in four main steps as follows (with proofs below):

1) $(J^+, J^-, I_1, I_2) \sim (J^+, J^-, I_1', I_2')$ where I_1' and I_2' are Cauchy developments of disjoint Cauchy balls of a common Cauchy surface Σ^{+-} of $J^+ \cap J^-$.

2) $(J^+, J^-, I_1', I_2') \sim (J'^+, J'^-, I_1', I_2')$ where J'^+ and J'^- are Cauchy developments of Cauchy balls of Cauchy surfaces Σ_0 and Σ_1 of $\mathcal{M}$ respectively.

3) $(J'^+, J'^-, I_1', I_2') \sim (J''^+, J'^-, I_1', I_2')$ where J''^+ and J'^- are Cauchy developments of Cauchy balls of Σ_1.

4) $(J''^+, J'^-, I_1', I_2') \sim (\widetilde{J}^+, \widetilde{J}^-, \widetilde{I}_1, \widetilde{I}_2)$ with the regions of the latter quadruple all being Cauchy developments of subregions of the Cauchy surface Σ.

Proof for 1) The intersection of J^+ and J^- is a causally convex subset of a globally hyperbolic spacetime, hence $J^+ \cap J^-$ is globally hyperbolic. This means there is a Cauchy surface Σ^{+-} of $J^+ \cap J^-$. Let Σ_{I_1} be a Cauchy surface of I_1, then $(J^+, J^-, I_1, I_2) \sim (J^+, J^-, D(B_1), I_2)$ where B_1 is a Cauchy ball of Σ_{I_1}. We can then use proposition 4.3.7 to get a chain of embeddings of $\mathsf{Loc}_C^{\mathcal{M}}$ regions within $J^+ \cap J^-$, and hence a chain of equivalent quadruples which relates $(J^+, J^-, D(B_1), I_2)$ to (J^+, J^-, I_1', I_2) where $I_1' = D(B_1')$ and B_1' is a Cauchy ball of Σ^{+-}. We then repeat this procedure for I_2 to get $I_2' = D(B_2')$ where B_2' is a Cauchy ball of Σ^{+-}. If B_1' and B_2' are not disjoint, we shrink them about distinct points until they are disjoint.

Proof for 2) Since B_1' and B_2' are disjoint Cauchy balls of a common acausal spacelike hypersurface, we can use [BS06, Theorem 1.1] to find Cauchy surfaces Σ_{J+} and Σ_{J-} of J^+ and J^- respectively that contain B_1' and B_2'. We can use proposition 3.2.7 which we established in the previous chapter using the techniques of differential topology, to find Cauchy balls Σ^+ and Σ^- of Σ_{J+} and Σ_{J-} respectively, each containing $B_1' \cup B_2'$. We can further shrink B_1' and B_2' if necessary to ensure $B_1' \cup B_2' \subset \Sigma^+ \cap \Sigma^-$. We then get $J'^+ = D(\Sigma^+)$ and $J'^- = D(\Sigma^-)$, and we can use [BS06, Theorem 1.1] to find Cauchy surfaces Σ_0 and Σ_1 of $\mathcal{M}$ containing Σ^+ and Σ^- respectively.

Proof for 3) See proposition 4.3.8.

Proof for 4) We can use proposition 4.3.7 to find a chain of embeddings which links each $\mathsf{Loc}_C^{\mathcal{M}}$ region in $(J''^+, J'^-, I_1', I_2')$ to a $\mathsf{Loc}_C^{\mathcal{M}}$ region which is a Cauchy development of a subregion of Σ. Since these Cauchy surfaces of these regions all belong to Σ_1, these chains can be constructed in tandem, as mention below proposition 4.3.7, so that at each step the regions from the four chains give a well defined quadruple. □

4.4 Proving bijectivity of the map $\mathcal{L} : \mathcal{Q} \to \widetilde{\pi}_1(\mathcal{M})$

The difficult part of proving bijectivity of $\mathcal{L}$ is proving injectivity, since surjectivity just requires us to prove that we can construct a quadruple associated to any homotopy class of loops. The latter can be done by proving each homotopy class has an element which is an embedding (see definition 1.2.7) with image in a Cauchy surface Σ, because given such a curve we can construct tubular neighbourhoods in Σ and take Cauchy developments of these tubes to get $\mathsf{Loc}_C^{\mathcal{M}}$ regions that form a quadruple (a detailed account of this construction is provided by lemma 4.4.2 below).

To prove injectivity of $\mathcal{L}$, we must be able to relate any two quadruples q and Q such that each q-curve is homotopic to each Q-curve (see definition 4.2.4). We can use theorem

4.3.9 from the last section to translate q and Q to a common Cauchy surface Σ, so that the q-curve and Q-curve both have image in Σ and the homotopy between them is also confined to Σ. We would then like to apply the construction sketched in the previous paragraph to the intermediate loops in Σ defined by the homotopy to get intermediate quadruples relating q and Q, thereby proving injectivity. The issue is that the above construction seems to require the loops to be embedded, since this is required for the construction of tubular neighbourhoods, and in general it is not possible to arrange the homotopy so that each of its intermediate loops are embedded. This issue can be circumvented by noting that the construction does not actually require the entire loop to be embedded, it only requires the individual segments defining the $\mathsf{Loc}_C^{\mathcal{M}}$ components of the quadruple to be embedded. The intermediate loops can therefore have self-intersections, so long as they occur between segments defining different components of the quadruple. We will therefore use results from differential topology to prove that in dimension ≥ 3, any homotopy can be modified such that each of its intermediate loops has finitely many self-intersections. This will ensure that we can always shift the segments of the intermediate loops about, so that the self-intersections occur between segments defining different components of its associated quadruple.

We begin by defining a parameterisation of S^1 and a notion of quadruples on S^1, and then prove a lemma that defines a construction of quadruples associated to embedded loops in Cauchy surfaces.

Definition 4.4.1. *We model S^1 as $\mathbb{R}/\mathbb{Z}$ with quotient map $\mathrm{mod}_1 : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$, which takes an element of $\mathbb{R}$ to its equivalence class modulo 1. Contractible open subsets of S^1 are of the form $\mathrm{mod}_1\left((c - d/2, c + d/2)\right)$ with $0 < d \leq 1$ and without loss of generality $0 < c \leq 1$, and we refer to c as the centre of the subset and d as the width of the subset. Any contractible open subset of S^1 is uniquely specified by its centre and width, so we use $\lfloor c, d \rceil$ to denote the open subset of S^1 with centre c and width d.*

A quadruple of S^1 consist of four open contractible regions (j^+, j^-, i_1, i_2) of S^1, such that $i_1 \cup i_2 \subset j^+ \cap j^-$. We define an equivalence relation on quadruples of S^1 exactly as we did for quadruples of $\mathsf{Loc}_C^{\mathcal{M}}$ regions in definition 4.2.1. We will refer to

$$Q^C_{S^1} := \left(\ \lfloor 1/4, 3/4 \rceil, \lfloor 3/4, 3/4 \rceil, \lfloor 0, 1/4 \rceil, \lfloor 1/2, 1/4 \rceil\ \right) \tag{4.12}$$

as the canonical quadruple of S^1.

Lemma 4.4.2. *Let γ be an embedding of S^1 into a Cauchy surface Σ of $\mathcal{M}$, and Q_{S^1} be a quadruple of S^1. There exists $[Q] \in \mathcal{Q}$ and a subset $S_{\mathcal{M}}(\gamma, Q_{S^1}) \subset [Q]$ such that for each $q \in S_{\mathcal{M}}(\gamma, Q_{S^1})$ the components of q are Cauchy developments of subsets of Σ, γ is a q-curve (see definition 4.2.4) and the components of Q_{S^1} are equal to the pre-image under γ of the corresponding components of q.*

Proof. The image of an embedding is a submanifold, so we can use [Hir76, Theorem 4.5.2] to guarantee the existence of tubular neighbourhoods (see definition 3.2.3) of γ in Σ. Using γ and Q_{S^1}, we construct a set $S_\Sigma(\gamma, Q_{S^1})$ whose objects each consist of: tubular neighbourhoods in Σ of γ restricted to each component of $Q_{S^1} = (j^+, j^-, i_1, i_2)$, with the requirement that the tubular neighbourhoods of $\gamma|_{i_1}$ and $\gamma|_{i_2}$ are both contained in the intersection of the tubular neighbourhoods of $\gamma|_{j^+}$ and $\gamma|_{j^-}$. These tubular neighbourhoods are contractible, since a bundle over a contractible base is just a product of the base and fibre spaces, which in this case are both contractible so the bundle as a whole is contractible.

We can then form another set $S_{\mathcal{M}}(\gamma, Q_{S^1})$ by taking each object of $S_\Sigma(\gamma, Q_{S^1})$ and taking Cauchy developments of its tubular neighbourhoods, which define components of a quadruple q (since Cauchy developments preserve contractibility and are causally convex) which we take to be the corresponding object of $S_{\mathcal{M}}(\gamma, Q_{S^1})$. Each quadruple $q \in S_{\mathcal{M}}(\gamma, Q_{S^1})$ has the property that γ is a q-curve and the components of Q_{S^1} are equal to the pre-image under γ of the corresponding components of q.

For any $q_1, q_2 \in S_{\mathcal{M}}(\gamma, Q_{S^1})$ there exists $q_3 \in S_{\mathcal{M}}(\gamma, Q_{S^1})$ such that $q_3 \subset q_1$ and $q_3 \subset q_2$ (meaning each component of the first quadruple is a subset of the corresponding component of the second quadruple). This is because we can always arrange the tubular neighbourhoods used to construct the components of q_3 to have sufficiently small radii to be contained in those of q_1 and q_2. Therefore all quadruples in $S_{\mathcal{M}}(\gamma, Q_{S^1})$ belong to the same equivalence class in $\mathcal{Q}$. $\square$

Definition 4.4.3. *Let γ be an embedding of S^1 into a Cauchy surface Σ of $\mathcal{M}$, and Q_{S^1} be a quadruple of S^1. Let $S_{\mathcal{M}}(\gamma, Q_{S^1})$ denote the set of quadruples of $\mathcal{M}$ constructed from γ and Q_{S^1} in lemma 4.4.2.*

With these preliminaries, we can now establish surjectivity of the map $\mathcal{L}$.

Proposition 4.4.4. *The map $\mathcal{L} : \mathcal{Q} \to \widetilde{\pi}_1(\mathcal{M})$ is surjective if* $\dim(\mathcal{M}) \geq 4$.

Proof. Our spacetime is globally hyperbolic so $\mathcal{M} \cong \mathbb{R} \times \Sigma$ which implies that for any $[\gamma] \in \widetilde{\pi}_1(\mathcal{M}) \cong \widetilde{\pi}_1(\Sigma)$, there exists a representative $[\gamma] \ni \widehat{\gamma} : S^1 \to \Sigma_0 := \{0\} \times \Sigma$. We can find a smooth embedding $\widetilde{\gamma}$ contained in an arbitrarily small neighbourhood of $\widehat{\gamma}$ (in the sense of the topology defined in definition 1.2.8), since $C^\infty(S^1, \Sigma_0)$ is open and dense in $C^0(S^1, \Sigma_0)$ [Hir76, Theorem 2.2.6] and $\dim(\Sigma_0) \geq 3 = 2\dim(S^1) + 1$ by supposition so $\mathrm{Emb}^\infty(S^1, \Sigma_0)$ is dense in $C^\infty(S^1, \Sigma_0)$ [Hir76, Theorem 2.2.13].

To construct a homotopy between $\widetilde{\gamma}$ and $\widehat{\gamma}$, we embed Σ_0 as a submanifold of $\mathbb{R}^N$ where $N = 2\dim(\Sigma_0) + 1$ using the Whitney embedding theorem [Hir76, Theorem 2.2.14], and use [Hir76, Theorem 4.5.2] to guarantee the existence of a tubular neighbourhood T of Σ_0 in $\mathbb{R}^N$. We can arrange for $\widehat{\gamma}$ and $\widetilde{\gamma}$ to be as close to each other as we like since $\mathrm{Emb}^\infty(S^1, \Sigma_0)$ is dense in $C^0(S^1, \Sigma_0)$, so the homotopy $H : S^1 \times [0, 1] \to \mathbb{R}^N$ given by

$$H(s, t) = t\widehat{\gamma}(s) + (1 - t)\widetilde{\gamma}(s)$$

has image contained in the tubular neighbourhood T. We get a homotopy between $\widehat{\gamma}$ and $\widetilde{\gamma}$ by composing H with the retraction of T onto Σ_0. We then pick an element in $Q \in S_{\mathcal{M}}(\widetilde{\gamma}, Q^C_{S^1})$ and see that $\mathcal{L}([Q]) = [\gamma]$ since $\widetilde{\gamma} \in [\gamma]$ is a Q-curve. $\square$

We now prove that for any pair of equivalent quadruples of S^1, the construction in definition 4.4.3 yields two sets of quadruples that belong to the same equivalence class in $\mathcal{Q}$, and then classify the equivalence classes of quadruples of S^1. These results will be used in the part of the strategy outlined at the start of the section where we shift segments about to avoid self-intersections.

Lemma 4.4.5. *If q_{S^1} is equivalent to Q_{S^1}, then $S_{\mathcal{M}}(\gamma, q_{S^1})$ and $S_{\mathcal{M}}(\gamma, Q_{S^1})$ are both subsets of the same equivalence class in $\mathcal{Q}$.*

Proof. If $q_{S^1} \subset Q_{S^1}$, then the tubular neighbourhoods used to construct a quadruple of $S_{\mathcal{M}}(\gamma, q_{S^1})$ can be arranged to be contained in those of a quadruple of $S_{\mathcal{M}}(\gamma, Q_{S^1})$, hence the quadruples are equivalent meaning all quadruples in both sets must be equivalent. Similarly if q_{S^1} and Q_{S^1} are related by swapping their last two component regions, then every element of $S_{\mathcal{M}}(\gamma, q_{S^1})$ has an equivalent element in $S_{\mathcal{M}}(\gamma, Q_{S^1})$ obtained by swapping the last two component regions. □

Lemma 4.4.6. *All quadruples of S^1 are either equivalent to the trivial quadruple (I, I, I, I) where I is S^1 with the point $[0]$ removed, or the canonical quadruple $Q^C_{S^1}$ (see equation (4.12)).*

Proof. Let (j^+, j^-, i_1, i_2) be a quadruple of S^1. We first consider the case where i_1 and i_2 belong to the same connected component j of $j^+ \cap j^-$. In this case we we get $(j^+, j^-, i_1, i_2) \sim (j, j, j, j)$ since $j^\pm \supset j$ and $i_1 \cup i_2 \subset j$, and (j, j, j, j) is related to the trivial quadruple by using the following chain of embeddings

$$j = \lfloor c, d \rfloor \subset \lfloor c, 1 \rfloor \supset \widetilde{j} \subset I$$

where $\widetilde{j}$ is a connected component of $\lfloor c, 1 \rfloor \cap I$.

Now we consider the case where i_1 and i_2 belong to different connected components of $j^+ \cap j^-$. This implies that at least one of j^+ and j^- must have width $> 1/2$, since otherwise their intersection would only have one connected component. We can therefore without loss of generality assume that j^+ has width $> 1/2$.

Each segment $\lfloor c, d \rfloor$ of S^1 can be related to its translation by $\epsilon < (1-d)/2$ by the following chain of embeddings $\lfloor c, d \rfloor \subset \lfloor c+\epsilon, d+2\epsilon \rfloor \supset \lfloor c+\epsilon, d \rfloor$. This gives us a means of relating two quadruples of S^1 differing (in their component regions) by a translation, successively relating quadruples differing by a translation by an amount $\leq \epsilon$ by expanding and shrinking the component regions as outlined above. We can therefore relate (j^+, j^-, i_1, i_2) to a quadruple $(\ \lfloor 1/4, d_+ \rfloor, \lfloor c_-, d_- \rfloor, \lfloor c_1, d_1 \rfloor, \lfloor c_2, d_2 \rfloor\ )$, so that the first component has the same centre as the first component of $Q^C_{S^1}$.

As mentioned above, the intersection of $\lfloor 1/4, d_+ \rfloor$ with $\lfloor c_-, d_- \rfloor$ has two connected components and we can assume without loss of generality that $d_+ > 1/2$, which implies $c_- \in (1/2, 1)$. The region $\lfloor 1/4, 1 \rfloor \cap \lfloor c_-, 1 \rfloor$ consists of both arcs of S^1 that connect the antipodal points of $[1/4]$ and c_-. Since $c_- \in (1/2, 1)$, one of these arcs must contain $[0]$ and the other must contain $[1/2]$, and we refer to these arcs as U_1 and U_2 respectively (see figure 4.2). We construct neighbourhoods N_1 and N_2 of $[0]$ and $[1/2]$ respectively such that $N_1 \subset \lfloor 0, 1/4 \rfloor \cap U_1$ and $N_2 \subset \lfloor 1/2, 1/4 \rfloor \cap U_2$ (again see figure 4.2).

We now construct a chain of equivalences as follows

$$\begin{aligned}
&(j^+, j^-, i_1, i_2) \sim (\ \lfloor 1/4, d_+ \rfloor, \lfloor c_-, d_- \rfloor, \lfloor c_1, d_1 \rfloor, \lfloor c_2, d_2 \rfloor\ ) \\
&\sim (\ \lfloor 1/4, 1 \rfloor, \lfloor c_-, 1 \rfloor, \lfloor c_1, d_1 \rfloor, \lfloor c_2, d_2 \rfloor\ ) \sim (\ \lfloor 1/4, 1 \rfloor, \lfloor c_-, 1 \rfloor, U_1, U_2\ ) \quad ,
\end{aligned}$$

where the first equivalence was shown above, and the last equivalence is due to the fact that $\lfloor c_1, d_1 \rfloor$ must be a subset of either U_1 or U_2, so we use the freedom to swap the third

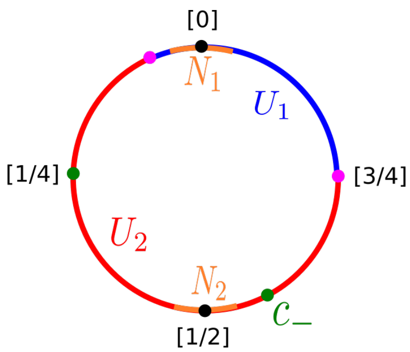

Figure 4.2: Connected components of $\lfloor 1/4, 1\rfloor \cap \lfloor c_-, 1\rfloor$, together with N_1 and N_2.

and fourth components of the quadruple if $\lfloor c_1, d_1\rfloor \subset U_2$. Since $N_1 \subset U_1$ and $N_2 \subset U_2$, we get

$$\sim \big(\ \lfloor 1/4, 1\rfloor, \lfloor c_-, 1\rfloor, N_1, N_2\ \big) \sim \big(\ \lfloor 1/4, 1\rfloor, \widetilde{j}^-, N_1, N_2\ \big) \sim \big(\ \lfloor 1/4, 1\rfloor, \lfloor 3/4, 1\rfloor, N_1, N_2\ \big) \quad ,$$

where $\widetilde{j}^-$ is the component of of $\lfloor 3/4, 1\rfloor \cap \lfloor c_-, 1\rfloor$ containing [0] and [1/2] (see figure 4.3). For the last two equivalences, we use the inclusions $\widetilde{j}^- \subset \lfloor c_-, 1\rfloor$ and $\widetilde{j}^- \subset \lfloor 3/4, 1\rfloor$ respectively. The fact that $\lfloor 1/4, 1\rfloor \cap \widetilde{j}^- \supset N_1 \cup N_2$ ensures that the middle quadruple is a valid quadruple.

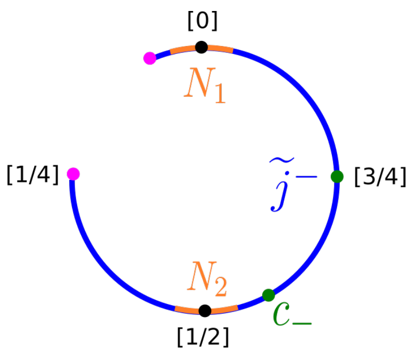

Figure 4.3: Visualising $\widetilde{j}^-$.

We then use $N_1 \subset \lfloor 0, 1/4\rfloor$ and $N_2 \subset \lfloor 1/2, 1/4\rfloor$ to get

$$\begin{aligned} &\sim \big(\ \lfloor 1/4, 1\rfloor, \lfloor 3/4, 1\rfloor, \lfloor 0, 1/4\rfloor, \lfloor 1/2, 1/4\rfloor\ \big) \\ &\sim \big(\ \lfloor 1/4, 3/4\rfloor, \lfloor 3/4, 3/4\rfloor, \lfloor 0, 1/4\rfloor, \lfloor 1/2, 1/4\rfloor\ \big) = Q^C_{S^1} \quad . \end{aligned}$$

We have therefore shown that in the case where i_1 and i_2 belong to different connected components of $j^+ \cap j^-$, the quadruple is equivalent to $Q^C_{S^1}$. Therefore in both cases the quadruple is equivalent to either (I, I, I, I) or $Q^C_{S^1}$. $\square$

We now introduce a definition that formalises the idea of self-intersections of intermediate curves in a homotopy between two curves. We then prove that a homotopy between two embedded curves can always be approximated by one in which only finitely many of the intermediate curves have self-intersections[2].

Definition 4.4.7. *A pair of level double points of a map $H : S^1 \times [0, 1] \to \Sigma$ consists of*

[2] I would like to thank Charles Livingston for a helpful private communication.

a pair $(s_1, t), (s_2, t) \in S^1 \times [0,1]$ such that $H(s_1, t) = H(s_2, t)$, in which case we say the level double point occurs at level t.

Proposition 4.4.8. *Given a homotopy H between two smooth embedded loops in a manifold Σ with $\dim(\Sigma) \geq 3$, there exists another homotopy $\widetilde{H}$ in an ϵ neighbourhood of H (in the C^0 topology defined in definition 1.2.8), such that $\widetilde{H}$ matches H at the boundary and has finitely many level double points.*

Proof. For notational convenience we will use C to denote $S^1 \times [0,1]$. We consider the map $F = \langle H, \mathrm{pr}_2 \rangle : C \to \Sigma \times [0,1]$ where pr_2 is the projection map onto the second component, so the action of F is given by $F(s,t) = \big(H(s,t), t\big)$. We want to apply theorem 1.2.14 to approximate F to a smooth immersion F' that matches F at the boundary, such that $F'|_{\mathrm{int}(C)}$ and $F'|_{\partial C}$ are both self-transverse and transverse to each other, but we must check that the assumptions of the theorem are satisfied. We are applying the theorem to the case $L = \partial C$, and indeed we see that the pair (F, L) satisfy the bullet point conditions of the theorem: The first three are satisfied because H is a smooth self-transverse (see definition 1.2.11) immersion at the boundary, since H is a homotopy between smooth embeddings, the fourth because $L \backslash \partial C = \emptyset$, and the fifth because the component of F that maps to $[0,1]$ is the projection map pr_2. By supposition $\dim(\Sigma) \geq 3$ so $2\dim(C) \leq \dim(\Sigma \times [0,1])$, hence all the conditions of theorem 1.2.14 are satisfied and we can use it to approximate F by a smooth immersion F' that matches F at the boundary, such that $F'|_{\mathrm{int}(C)}$ and $F'|_{\partial C}$ are both self-transverse and transverse to each other.

Any function into a product space can be written in terms of component functions that map into the components of the target space, hence we can write $F' = \langle H', \mathrm{pr}_2' \rangle$ with H' and pr_2' belonging to neighbourhoods of H and pr_2 respectively. Let $G = \langle \mathrm{pr}_1, \mathrm{pr}_2' \rangle$ be a map from C to itself. Since we can arrange our perturbations to be as small as we like, we can make pr_2' arbitrarily close to pr_2, and hence make G arbitrarily close to the identity. The set of diffeomorphisms form an open subset of the set of maps that map the boundary of the domain space to the boundary of the target space (see comments below theorem 2.1.7 in [Hir76]), and because F' matches F at the boundary this implies G must be the identity on the boundary, hence we can arrange the approximations to be sufficiently small so that G is a diffeomorphism on C.

We now define $\widetilde{F} = F' \circ G^{-1} = \langle \widetilde{H}, \mathrm{pr}_2 \rangle$, where $\widetilde{H} = H' \circ G^{-1}$ is well defined and smooth since G is a diffeomorphism. Diffeomorphisms on C preserve self-transversality and transversality between functions defined on C, hence $\widetilde{F}|_{\mathrm{int}(C)}$ and $\widetilde{F}|_{\partial C}$ are both self-transverse and transverse to each other. This implies that $\widetilde{F}$ has finitely many double points by proposition 1.2.15, and since double points of $\widetilde{F}$ are level double points of $\widetilde{H}$, we have shown that $\widetilde{H}$ has finitely many level double points. Since H' matches H at the boundary and G is the identity on the boundary, we see that $\widetilde{H}$ matches H at the boundary, hence $\widetilde{H}$ has the required properties for the proposition.

□

We now prove injectivity.

Proposition 4.4.9. *The map $\mathcal{L} : \mathcal{Q} \to \widetilde{\pi}_1(\mathcal{M})$ is injective if* $\dim(\mathcal{M}) \geq 4$.

Proof. We consider equivalence classes $[q]$ and $[Q]$ such that $\mathcal{L}([q]) = \mathcal{L}([Q])$, and fix a choice of Cauchy surface Σ. We then use theorem 4.3.9 to find representatives $q \in [q]$

and $Q \in [Q]$, whose components are Cauchy developments of Cauchy balls of Σ. To prove injectivity, we must find a sequence of equivalent quadruples relating q and Q. We consider two embedded curves γ_0 and γ_1 in Σ, such that γ_0 is a q-curve and γ_1 is a Q-curve. Since we assume that $\mathcal{L}([q]) = \mathcal{L}([Q])$, it follows that (by reversing the orientation of γ_0 if necessary) there is a homotopy between γ_0 and γ_1. We then use proposition 4.4.8 to approximate this homotopy by a smooth homotopy H which has finitely many level double points, which occur at the levels $\{t_1, \dots t_N\}$. This means that $H(\cdot, t)$ is an embedded loop unless $t \in \{t_1, \dots t_N\}$, in which case it is a loop with finitely many crossings.

If $[q]$ is trivial, in the sense that it has a representative of the form (I, I, I, I) for some $\mathsf{Loc}_C^{\mathcal{M}}$ region I, we choose q to be a representative of this form. Then the q-curve γ_0 has a tubular neighbourhood contained in I, hence for any quadruple q_{S^1} of S^1 there is a quadruple $\mathscr{Q}_0 \in S_{\mathcal{M}}(\gamma_0, q_{S^1})$ such that $\mathscr{Q}_0 \subset q$. For convenience we choose q_{S^1} to be the canonical quadruple $Q^C_{S^1}$ of S^1. If $[q]$ is non-trivial, let q_{S^1} be the quadruple of S^1 whose components are given by the pre-image under γ_0 of the corresponding components of q. Since $[q]$ is non-trivial, the resulting quadruple q_{S^1} of S^1 must be non-trivial, and therefore by lemma 4.4.6 q_{S^1} is equivalent to $Q^C_{S^1}$. There is a quadruple $\mathscr{Q}_0 \in S_{\mathcal{M}}(\gamma_0, q_{S^1})$ such that $\mathscr{Q}_0 \subset q$, since we can choose the radii of the tubular neighbourhoods used to construct the component regions of $\mathscr{Q}_0$ to be arbitrarily small. We therefore see that in both cases, we can construct an element $\mathscr{Q}_0 \in S_{\mathcal{M}}(\gamma_0, q_{S^1})$ such that $\mathscr{Q}_0 \subset q$ for some quadruple q_{S^1} of S^1 which is equivalent to $Q^C_{S^1}$. We can similarly construct $\mathscr{Q}_1 \in S_{\mathcal{M}}(\gamma_1, Q_{S^1})$ such that $\mathscr{Q}_1 \subset Q$ for some quadruple Q_{S^1} of S^1 which is equivalent to $Q^C_{S^1}$.

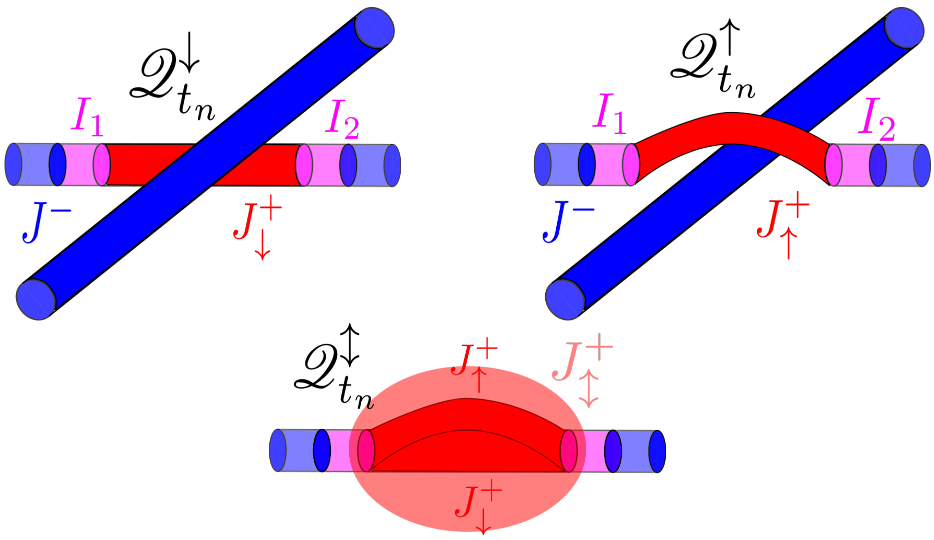

Figure 4.4: Resolving double points at t_n.

For sufficiently small $\epsilon_1 > 0$, there exists $S_{\mathcal{M}}(\gamma_0, q_{S^1}) \ni \widetilde{\mathscr{Q}}_0 \supset \mathscr{Q}_0$ such that $H(S^1, \epsilon_1)$ is an embedding whose image is contained in $\widetilde{\mathscr{Q}}_0$. The quadruple $\widetilde{\mathscr{Q}}_0$ also contains an element $\mathscr{Q}_{\epsilon_1} \in S_{\mathcal{M}}(H(\cdot, \epsilon_1), q_{S^1})$. Similarly, there exists $\epsilon_2 > 0$ sufficiently small such that $S_{\mathcal{M}}(\gamma_{\epsilon_1}, q_{S^1}) \ni \widetilde{\mathscr{Q}}_{\epsilon_1} \supset \mathscr{Q}_{\epsilon_1}$ is a quadruple that contains the embedding $H(S^1, \epsilon_1 + \epsilon_2)$. We can repeat this process to get a chain of quadruples

$$q \supset \mathscr{Q}_0 \subset \widetilde{\mathscr{Q}}_0 \supset \mathscr{Q}_{\epsilon_1} \subset \widetilde{\mathscr{Q}}_{\epsilon_1} \supset \mathscr{Q}_{\epsilon_1+\epsilon_2} \subset \dots \subset \widetilde{\mathscr{Q}}_{t_1-\delta} \supset \mathscr{Q}_{t_1-\delta} \tag{4.13}$$

where $\delta > 0$ is arbitrarily small. This procedure must be modified to relate $\mathscr{Q}_{t_1-\delta}$ to a quadruple containing loops $H(S^1, t)$ with $t > t_1$, since H is no longer an embedding of S^1 at t_1. We therefore need to resolve the crossing points at t_1 which we do as follows: Let (s_1, t_1) and (s_2, t_1) be a pair of double points of $H(\cdot, t_1)$ and $\widetilde{q}_{S^1} = (j^+, j^-, i_1, i_2)$ be a quadruple of S^1 such that j^+ is neighbourhood of s_1 which is disjoint from all the other double points of $H(\cdot, t_1)$ (which is possible since there are only finitely many). We can also arrange for $\widetilde{q}_{S^1}$ to be equivalent to q_{S^1}, since only the first component of $\widetilde{q}_{S^1}$ is constrained so we can arrange its other three regions to make it equivalent to the canonical quadruple $Q^C_{S^1}$ of S^1. We then pick a quadruple $\mathscr{Q}^{\downarrow}_{t_1} = (J^+_{\downarrow}, J^-, I_1, I_2) \in S_{\mathcal{M}}(H(\cdot, t_1 - \delta), \widetilde{q}_{S^1})$ and let $J^+_{\uparrow}$ be the Cauchy development of a tubular neighbourhood of $H(\cdot, t_1 + \delta)|_{j^+}$ (see figure 4.4).

We can choose δ sufficiently small so that the Cauchy surfaces of $J^+_{\downarrow}$ and $J^+_{\uparrow}$ are contained in a common contractible subregion of Σ (the t derivatives of H are bounded), whose Cauchy development is therefore a $\mathsf{Loc}^{\mathcal{M}}_C$ region $J^+_{\updownarrow}$ containing $J^+_{\downarrow}$ and $J^+_{\uparrow}$ (see bottom diagram in figure 4.4). By lemma 4.4.5, we can relate $\mathscr{Q}_{t_1-\delta}$ to $\mathscr{Q}^{\downarrow}_{t_1}$ since they are both constructed from the same embedded curve $H(\cdot, t_1 - \delta)$ using equivalent quadruples of S^1. Let $\mathscr{Q}^{\uparrow}_{t_1} = (J^+_{\uparrow}, J^-, I_1, I_2)$ and $\mathscr{Q}^{\updownarrow}_{t_1} = (J^+_{\updownarrow}, J^-, I_1, I_2)$. We illustrate the quadruples $\mathscr{Q}^{\downarrow}_{t_1}$, $\mathscr{Q}^{\uparrow}_{t_1}$ and $\mathscr{Q}^{\updownarrow}_{t_1}$ at a neighbourhood of the double point in figure 4.4. Since $J^+_{\downarrow}, J^+_{\uparrow} \subset J^+_{\updownarrow}$, we get the following equivalences of quadruples

$$\mathscr{Q}_{t_1-\delta} \sim \mathscr{Q}^{\downarrow}_{t_1} \sim \mathscr{Q}^{\updownarrow}_{t_1} \sim \mathscr{Q}^{\uparrow}_{t_1} \quad .$$

We continue this procedure until we get a quadruple $\mathscr{Q}_{t_1}$ that has resolved all the crossing points of $H(S^1, t_1)$. The quadruple $\mathscr{Q}_{t_1}$ therefore contains a loop that is homotopic through embeddings to $H(S^1, t_1+\delta)$, so we can repeat the procedure outlined above equation (4.13) to show that $\mathscr{Q}_{t_1}$ is equivalent to some $\mathscr{Q}_{t_1+\delta} \in S_{\mathcal{M}}(H(\cdot, t_1 + \delta), q_{S^1})$. We then continue as before to relate quadruples containing $H(S^1, t)$ to q for increasing values of t, resolving crossing points as they occur with the procedure above, until we have related q to a quadruple $\mathscr{Q}_1 \in S_{\mathcal{M}}(H(\cdot, 1), q_{S^1}) = S_{\mathcal{M}}(\gamma_1, q_{S^1})$. Since q_{S^1} and Q_{S^1} are equivalent to $Q^C_{S^1}$, we can use lemma 4.4.5 to get $\mathscr{Q}_1 \sim Q$ and we therefore have

$$q \sim \mathscr{Q}_0 \sim \ldots \sim \mathscr{Q}_1 \sim Q$$

which proves that $[q] = [Q]$, and therefore the map $\mathcal{L} : \mathcal{Q} \to \widetilde{\pi}_1(\mathcal{M})$ is injective. □

4.5 Defining groups and establishing an isomorphism

In this section we first define a group related to $\widetilde{\pi}_1(\mathcal{M})$, we then define a group structure on $\mathcal{Y}$ and use the fact that $\mathcal{L}$ is bijective to construct a map m between these groups which we then prove is an isomorphism.

Definition 4.5.1. *We define the group $\widehat{\pi}_1(\mathcal{M}, p)$ as*

$$\pi_1(\mathcal{M}, p)/\pi_1(\mathcal{M}, p)^2 \tag{4.14}$$

where $\pi_1(\mathcal{M}, p)^2$ denotes the subgroup of $\pi_1(\mathcal{M}, p)$ generated by squares of elements in $\pi_1(\mathcal{M}, p)$.

We note that $\widehat{\pi}_1(\mathcal{M}, p)$ depends on a choice of base-point, but because $\mathcal{M}$ is assumed to be connected, $\widehat{\pi}_1(\mathcal{M}, p) \cong \widehat{\pi}_1(\mathcal{M}, q)$ for all $p, q \in \mathcal{M}$. The quotient that defines $\widehat{\pi}_1(\mathcal{M}, p)$ is well defined because $\pi_1(\mathcal{M}, p)^2$ is a normal subgroup of $\pi_1(\mathcal{M}, p)$

$$\forall g \in \pi_1(\mathcal{M}, p) \text{ and } \forall\, h^2 \in \pi_1(\mathcal{M}, p)^2 \; : \; gh^2g^{-1} = (ghg^{-1})^2 \;\Rightarrow\; gh^2g^{-1} \in \pi_1(\mathcal{M}, p)^2 \quad .$$

The group $\widehat{\pi}_1(\mathcal{M}, p)$ is also Abelian, which can be seen from the following calculation

$$g_1g_2 = g_1^{-1}g_2^{-1} = (g_2g_1)^{-1} = g_2g_1$$

where we have used the fact that every element is equal to its inverse in $\widehat{\pi}_1(\mathcal{M}, p)$. The cosets that form the elements of $\widehat{\pi}_1(\mathcal{M}, p)$ can instead be regarded as equivalence classes consisting of all elements of all equivalences classes in a given coset. We now introduce notation to distinguish between equivalence classes in $\widehat{\pi}_1(\mathcal{M}, p)$ and $\widetilde{\pi_1}(\mathcal{M})$.

Definition 4.5.2. *We use $[\cdot]_\wedge$ to denote the equivalence class that defines elements of $\widehat{\pi}_1(\mathcal{M}, p)$ and use $[\cdot]_\sim$ to denote the equivalence class that defines elements of $\widetilde{\pi}_1(\mathcal{M})$.*

We now define a semi-group structure on $\mathcal{Y}$ as follows.

Definition 4.5.3. *Let F be the forgetful functor from Alg to $\mathsf{SemiGroup}$, such that the semi-group operation is algebra multiplication. The semi-group $\mathcal{Y}_G$ is the sub semi-group of $F\left(\mathscr{U}_E^+(\mathbf{G})\right)$ generated by the elements of $\mathcal{Y}$.*

We will see later that this semi-group structure is in fact a group structure. We now want to define a map from $\widehat{\pi}_1(\mathcal{M}, p)$ to $\mathcal{Y}_G$ using the maps $\mathfrak{q}$ (see definition 4.2.2) and $\mathcal{L}$. The issue is that $\mathcal{L}$ is defined on $\widetilde{\pi}_1(\mathcal{M})$ whose equivalence classes will differ from those of $\widehat{\pi}_1(\mathcal{M}, p)$, since for any $\gamma \in C^0(S^1, \mathcal{M})$ the loop $\gamma*\gamma*\gamma$ belongs to $[\gamma*\gamma*\gamma]_\wedge = [\gamma]_\wedge^3 = [\gamma]_\wedge$ but does not necessarily belong to $[\gamma]_\sim$. We therefore prove the following fact before defining a map from $\widehat{\pi}_1(\mathcal{M}, p)$ to $\mathcal{Y}_G$.

Proposition 4.5.4. *If* $\dim(\mathcal{M}) \geq 4$*, then for all* $\gamma_1, \gamma_2 \in C^0(S^1, \mathcal{M})$

$$(\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1]_\sim)\ (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_2]_\sim) = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1 * \gamma_2]_\sim) \quad . \tag{4.15}$$

Proof. We first show that there is a $\mathsf{Loc}_C^{\mathcal{M}}$ region J, such that for any $[\gamma]_\sim \in \widetilde{\pi}_1(\mathcal{M})$, there is a quadruple associated to it whose first or second component is J. This will be used when we evaluate the product on the left-hand side of equation (4.15), since it will allow us to use lemma 4.1.8 to simplify the expression.

Let Σ be a Cauchy surface of $\mathcal{M}$, then for any fixed choice of smooth embedded curve $\mathcal{P} : [0,1] \to \Sigma$ and any $[\gamma]_\sim \in \widetilde{\pi}_1(\mathcal{M})$, there is a representative $[\gamma]_\sim \ni \widehat{\gamma} : [0,1] \to \Sigma$ which we use to define $\gamma' := (\overline{\mathcal{P}} * \widehat{\gamma}) * \mathcal{P} \in [\gamma]_\sim$. Given how we have grouped the multiplications (see definition 1.1.9), we see that γ' applied to $\mathrm{mod}_1\left([0, 1/2]\right)$ yields the submanifold given by the image of $\mathcal{P}$. Let $j = \mathrm{mod}_1\left((0, 1/2)\right)$, $\widetilde{j} = \mathrm{mod}_1\left((3/5, 4/5)\right)$, and i_1 and i_2 be subsets of distinct components of $j \cap \widetilde{j}$. We then use q_{S^1} and $\overline{q_{S^1}}$ to denote the quadruples of S^1 given by $(\widetilde{j}, j, i_1, i_2)$ and $(j, \widetilde{j}, i_1, i_2)$ respectively.

Since $\mathcal{P}$ is a smooth embedding, γ' restricted to $\overline{j}$ is a smooth self-transverse immersion. By supposition we also have $\dim(\Sigma) \geq 2\dim(S^1) + 1$, hence we can use corollary 1.2.16

to find $\widetilde{\gamma} \in \mathrm{Emb}^\infty(S^1, \Sigma)_{\gamma'|_{\overline{J}}}$ (see definition 1.2.13) which is arbitrarily close to γ'. This implies $\widetilde{\gamma}$ is homotopic to γ', and therefore $\widetilde{\gamma} \in [\gamma]_\sim$. Since $\widetilde{\gamma}|_{\overline{J}}$ matches $\mathcal{P}$, we have $\widetilde{\gamma} = h * \mathcal{P}$ where $h : [0,1] \to \Sigma$ is a smooth embedding. Therefore, $\forall\, \gamma_1, \gamma_2 \in C^0(S^1, \mathcal{M})$, there exist embeddings $h_1, h_2 : [0,1] \to \Sigma$ such that $h_1 * \mathcal{P}$ is homotopic to γ_1 and $h_2 * \mathcal{P}$ is homotopic to $\overline{\gamma_2}$. This implies that for any tubular neighbourhood J of $\mathcal{P}$, we can find $(J^+, J, I_1, I_2) \in S_\mathcal{M}(h_1 * \mathcal{P}, q_{S^1}) \subset \mathcal{L}([\gamma_1]_\sim)$ and $(J, J^-, I_1, I_2) \in S_\mathcal{M}(h_2 * \mathcal{P}, \overline{q_{S^1}}) \subset \mathcal{L}([\gamma_2]_\sim)$.

We can now use lemma 4.1.8 to simplify the multiplication of the elements of $\mathcal{Y}_G$ corresponding to these equivalence classes of quadruples

$$\begin{aligned}(\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1]_\sim)\ (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_2]_\sim) &= \mathfrak{q}\left([(J^+, J, I_1, I_2)]\right)\ \mathfrak{q}\left([(J, J^-, I_1, I_2)]\right) \\ &= \mathfrak{q}\left([(J^+, J^-, I_1, I_2)]\right) \quad .\end{aligned}$$

Since $h_1 * \mathcal{P}$ is homotopic to γ_1 and $h_2 * \mathcal{P}$ is homotopic to $\overline{\gamma_2}$, we find that $h_1 * \overline{h_2} \in [\gamma_1 * \gamma_2]_\sim$. We therefore find that $\mathfrak{q}\left([(J^+, J^-, I_1, I_2)]\right) = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1 * \gamma_2]_\sim)$, because $h_1 * \overline{h_2}$ is a (J^+, J^-, I_1, I_2)-curve, and plugging this back into the above equation proves the proposition. ☐

With this result we are now in a position to define the map from $\widehat{\pi}_1(\mathcal{M}, p)$ to $\mathcal{Y}_G$ which we will go on to show is an isomorphism.

Definition 4.5.5. *Let $m : \widehat{\pi}_1(\mathcal{M}, p) \to \mathcal{Y}_G$ be the map whose action on a representative $\gamma \in [\gamma]_\wedge \in \widehat{\pi}_1(\mathcal{M}, p)$ is given by $(\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma]_\sim)$.*

Proposition 4.5.6. *The map m is well-defined and is a homomorphism.*

Proof. For any two representatives $\gamma_1, \gamma_2 \in [\gamma]_\wedge$, the product $[\gamma_1]_\wedge [\gamma_2]_\wedge^{-1} = [\gamma_1 * \overline{\gamma_2}]_\wedge = \mathbb{1}$. Given how $\widehat{\pi}_1(\mathcal{M}, p)$ is defined, it follows that $\gamma_1 * \overline{\gamma_2}$ must be an element of some equivalence class in $\pi_1(\mathcal{M}, p)^2$. Therefore there exists some $\gamma_3 \in C^0(S^1, \mathcal{M})$ with base-point p such that $\gamma_1 * \overline{\gamma_2}$ is homotopic to $\gamma_3 * \gamma_3$, hence $[\gamma_1 * \overline{\gamma_2}]_\sim = [\gamma_3 * \gamma_3]_\sim$. We can then use proposition 4.5.4 to perform the following calculation

$$\begin{aligned}&(\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1 * \overline{\gamma_2}]_\sim) = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_3 * \gamma_3]_\sim) \\ &= (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_3]_\sim)\ (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_3]_\sim) = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_3]_\sim)\ (\mathfrak{q} \circ \mathcal{L}^{-1})([\overline{\gamma_3}]_\sim) \\ &= (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_3 * \overline{\gamma_3}]_\sim) = \mathbb{1}\end{aligned}$$

where the last equality follows from the fact that $\mathfrak{q} \circ \mathcal{L}^{-1}$ applied to a contractible loop yields the identity. We therefore find

$$\begin{aligned}&(\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1]_\sim) = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1 * (\overline{\gamma_2} * \gamma_2)]_\sim) \\ &= (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_1 * \overline{\gamma_2}]_\sim)\ (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_2]_\sim) = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_2]_\sim) \quad ,\end{aligned}$$

thus we see that m applied to any two representatives of an element of $\widehat{\pi}_1(\mathcal{M}, p)$ yields the same result, hence m is well-defined. The fact that m is a homomorphism clearly follows from proposition 4.5.4. ☐

Proposition 4.5.7. *The map $m : \widehat{\pi}_1(\mathcal{M}, p) \to \mathcal{Y}_G$ is an isomorphism.*

Proof. To show that m is surjective we must first show that all elements of $\mathcal{Y}_G$ are contained in $\mathcal{Y}$, since then surjectivity of m follows from the fact that $\mathfrak{q}$ is surjective (as remarked below definition 4.2.2) and $\mathcal{L}$ is bijective. For all $Y_1, Y_2 \in \mathcal{Y}$ there exists $[\gamma_1]_\sim, [\gamma_2]_\sim \in \widetilde{\pi}_1(\mathcal{M})$ such that $Y_i = (\mathfrak{q} \circ \mathcal{L}^{-1})([\gamma_i]_\sim$ for $i \in \{1, 2\}$. We can therefore use proposition 4.5.4 to see that the product of Y_1 with Y_2 belongs to the image of $\mathfrak{q} \circ \mathcal{L}^{-1}$ and is therefore an element of $\mathcal{Y}$. Therefore all elements of $\mathcal{Y}_G$ are contained in $\mathcal{Y}$ thus m is surjective.

Since m is a homomorphism, to prove it is injective it is sufficient to prove that its kernel is trivial. If $m([\gamma]_\wedge) = \mathbb{1}$ and $\mathcal{L}^{-1}([\gamma]_\sim) = [(J^+, J^-, I_1, I_2)]$ we get

$$m([\gamma]_\wedge) = b_{J^+}(f, g) b_{J^-}(g, f) = \mathbb{1} \quad \Rightarrow \quad b_{J^+}(f, g) = b_{J^-}(f, g) \quad .$$

We see from the relations imposed on our model for $\mathscr{U}_E^+(\mathbf{G})$ in proposition 3.1.12 that for the latter equality to hold we must have $J^+ = J^-$ or the local algebra elements $b(i_\pm^* f, i_\pm^* g) = \mathbb{1}_{J^\pm}$ where $i_\pm : J^\pm \to \mathcal{M}$ are inclusion morphisms and $\mathbb{1}_{J^\pm}$ is the identity element of $\mathscr{A}_{E\mathbf{G}}^+(J^\pm)$, which implies $i_\pm^* f = i_\pm^* g$ hence $I_1 = I_2$. Therefore either $J^+ = J^-$ or $I_1 = I_2$, and in either case all (J^+, J^-, I_1, I_2)-curves are contractible, hence $\mathcal{L}([(J^+, J^-, I_1, I_2)])$ is the identity element of $\widehat{\pi}_1(\mathcal{M}, p)$. We therefore see that m has trivial kernel and is therefore injective. Thus we have shown m is an injective and surjective homomorphism, and therefore is is an isomorphism. □

4.6 Decomposing the universal algebra

Now that we have established an isomorphism between $\widehat{\pi}_1(\mathcal{M}, p)$ and $\mathcal{Y}_G$, we will explore some further useful properties of the algebra elements of $\mathcal{Y}$.

Lemma 4.6.1. *For all $g \in \widehat{\pi}_1(\mathcal{M}, p)$, the elements $Y_g := m(g) \in \mathcal{Y}$ are self-adjoint and square to the identity.*

Proof. A simple calculation shows

$$Y_g^2 = m(g)^2 = m(g^2) = m(g) = \mathbb{1} \quad .$$

Using lemma 4.1.4, we see that taking the adjoint has the effect of swapping the first and second component regions of the associated quadruple. This means the loop associated to the adjoint is the same as before just with the orientation reversed, we therefore get

$$Y_g^* = m(g^{-1}) = m(g) = Y_g \quad .$$

□

Next we prove that the elements of $\mathcal{Y}$ are central, which means they commute with everything in $\mathscr{U}_E^+(\mathbf{G})$. In this proof we use the Einstein causality result for universal algebras proven in the previous chapter, showing one of many possible applications of that result.

Proposition 4.6.2. *Every element of $\mathcal{Y}$ belongs to the centre of $\mathscr{U}_E^+(\mathbf{G})$ if $\mathbf{G}$ has a base space of dimension ≥ 4.*

Proof. We consider a generic generator $\phi_O(\mathcal{A}(f)\mathcal{A}(h))$ (where f and h have compact support in O, recalling that ϕ_O maps elements of the local algebra of $O \in \mathsf{Loc}_C^{\mathcal{M}}$ into the universal algebra) of $\mathscr{U}_E^+(\mathbf{G})$ and show that it commutes with any $Y_g \in \mathcal{Y}$. To do this we fix a choice of Cauchy surface Σ, and find a representative $\gamma \in [\widetilde{\gamma}] = g$ with image contained in Σ that passes through $\widetilde{\Sigma} := \Sigma \backslash \big(J(\operatorname{supp} f \cup \operatorname{supp} h) \cap \Sigma\big)$. We see that $\widetilde{\Sigma}$ is an open subset of Σ because it is the complement of a closed subset of Σ ($J(\operatorname{supp} f \cup \operatorname{supp} h)$ is closed by virtue of 1.3.12).

By supposition, the dimension of Σ is ≥ 3, so we can use [Hir76, Theorem 2.2.13] (if necessary) to perform an arbitrarily small perturbation to γ such that it is an embedded curve. Moreover, since $\widetilde{\Sigma}$ is an open subset of Σ, this perturbation can be made sufficiently small so that γ still passes through $\widetilde{\Sigma}$. Let $q_{S^1} = (j^+, j^-, i_1, i_2)$ where j^+ is a connected component of $\gamma^{-1}(\widetilde{\Sigma})$, j^- is the complement of a connected closed subset of j^+, and i_1 and i_2 are open connected subsets of the components of $j^+ \cap j^-$. Since $\gamma|_{i_1}$ is contained in $\widetilde{\Sigma}$ which is open, it has a tubular neighbourhood contained in $\widetilde{\Sigma}$ (and similarly for $\gamma|_{i_2}$). We can therefore pick a quadruple $(J^+, J^-, I_1, I_2) \in S_{\mathcal{M}}(\gamma, q_{S^1})$ such that I_1 and I_2 are Cauchy developments of subsets of $\widetilde{\Sigma}$, and therefore are causally disjoint from $\operatorname{supp} f \cup \operatorname{supp} h$.

We thus see that Y_g is equivalent to the element of $\mathcal{Y}$ constructed from the quadruple (J^+, J^-, I_1, I_2) (see definition 4.1.3), which is composed of generators which are spacelike separated from $\phi_O(\mathcal{A}(f)\mathcal{A}(h))$. We can therefore apply the Einstein causality result for universal algebras, theorem 3.4.2 which we proved in the previous chapter, to show that Y_g commutes with $\phi_O(\mathcal{A}(f)\mathcal{A}(h))$. $\square$

We now have all the results we need to decompose $\mathscr{U}_E^+(\mathbf{G})$. For each $Y_g \in \mathcal{Y}$ we can define the following algebra element

$$P_g^{\pm} = \frac{1}{2}(\mathbb{1} \pm Y_g) \quad . \tag{4.16}$$

This is algebra element is central and idempotent, hence elements of $P_g^{\pm}\mathscr{U}_E^+(\mathbf{G})$ form a closed subspace under addition and multiplication. The fact that Y_g is also self-adjoint means $P_g^{\pm}\mathscr{U}_E^+(\mathbf{G})$ is also closed under taking adjoints, hence $P_g^{\pm}$ defines a projection (by left multiplication) onto a subalgebra.

To decompose the universal algebra, we take a minimal set of generators of $\widehat{\pi}_1(\mathcal{M}, p)$ given by the set $\{g_i\}_{i\in I}$ where I is an index set. This means any $g \in \widehat{\pi}_1(\mathcal{M}, p)$ can be written in terms of these generators. The choice of $\{g_i\}_{i\in I}$ is of course not unique. For each element in $\{g_i\}_{i\in I}$, we assign a value $c_i = \pm 1$ and denote the collection $\{c_i\}_{i\in I}$ as c_I. If the indext set I is finite (which is the case if $\mathcal{M}$ has finite first Betti number), we can construct the following algebra elements

$$P(c_I) = \prod_{i\in I} P_{g_i}^{c_i} \quad . \tag{4.17}$$

Proposition 4.6.3. *The algebra elements $P(c_I)$ (assuming $|I| < \infty$ so that they can constructed) have the following properties*

i) $P(c_I)P(\widetilde{c}_I) = \delta_{c_I\widetilde{c}_I} P(c_I)$ *where* $c_I, \widetilde{c}_I \in \mathbb{Z}_2^{|I|}$.

ii) $\sum_{c_I} P(c_I) = \mathbb{1}$.

iii) For any $c \in \{1, -1\}$ *and any* $g \in \widehat{\pi}_1(\mathcal{M}, p)$*, either* $P(c_I)\underset{g}{P}^{c} = P(c_I)$ *or* $P(c_I)\underset{g}{P}^{c} = 0$*.*

Proof. Clearly $P(c_I)^2 = P(c_I)$ since each of its factors is central and idempotent. If c_I and $\widetilde{c}_I$ are distinct, then there exists j such that $c_j \neq \widetilde{c}_j$. We know that $\underset{g_j}{P}^{c_j}\underset{g_j}{P}^{\widetilde{c}_j}$ is a factor of $P(c_I)P(\widetilde{c}_I)$ and that it must vanish since

$$c_j \neq \widetilde{c}_j \quad \Rightarrow \quad c_j + \widetilde{c}_j = 0 \quad c_j\widetilde{c}_j = -1 \quad ,$$

and therefore

$$\underset{g_j}{P}^{c_j}\underset{g_j}{P}^{\widetilde{c}_j} = \frac{1}{4}(\mathbb{1} + c_j Y_{g_j} + \widetilde{c}_j Y_{g_j} + c_j\widetilde{c}_j Y_{g_j}^2) = \frac{1}{4}(\mathbb{1} - \mathbb{1}) = 0 \quad ,$$

hence $P(c_I)P(\widetilde{c}_I)$ must also vanish. The sum over all possible choices of signs c_I can be rewritten using centrality of the factors of each $P(c_I)$ to obtain

$$\sum_{c_I} P(c_I) = \sum_{c_I}\prod_{i\in I}\underset{g_i}{P}^{c_i} = \prod_{i\in I}(\underset{g_i}{P}^{+} + \underset{g_i}{P}^{-}) = \prod_{i\in I}\mathbb{1} = \mathbb{1} \quad .$$

Finally given any $c \in \{1, -1\}$ and any $g \in \widehat{\pi}_1(\mathcal{M}, p)$ we can decompose g in terms of the generators as $g = g_{i(1)}\cdots g_{i(n)}$ so that we get

$$\begin{aligned}
\underset{g}{P}^{c} Y_g &= \frac{1}{2}(Y_g + c\mathbb{1}) = c\mathbb{1}\underset{g}{P}^{c} \qquad \Rightarrow \\
P(c_I)\underset{g}{P}^{c} &= P(c_I)\frac{1}{2}(\mathbb{1} + cY_g) = P(c_I)\frac{1}{2}(\mathbb{1} + cY_{g_{i(1)}}\cdots Y_{g_{i(n)}}) \\
&= P(c_I)\frac{1}{2}(\mathbb{1} + c\,c_{i(1)}\cdots c_{i(n)}\mathbb{1})
\end{aligned}$$

and $c\,c_{i(1)}\cdots c_{i(n)} = \pm 1$. Hence we get property iii). □

We now use these algebra elements to obtain the main result of this chapter.

Theorem 4.6.4. *Consider* $\mathbf{G} = (\mathcal{M}, \mathcal{E}, P) \in$ *GlobHypGreen such that* $\dim(\mathcal{M}) \geq 4$ *and* $\mathcal{M}$ *has finite first Betti number so that* $\pi_1(\mathcal{M})$ *is finitely generated. The universal algebra* $\mathscr{U}_E^+(\mathbf{G})$ *decomposes into a product of subalgebras*

$$\mathscr{U}_E^+(\mathbf{G}) \cong \mathscr{A}_E^{s_1}(\mathcal{M}) \times \cdots \times \mathscr{A}_E^{s_n}(\mathcal{M}) \tag{4.18}$$

where each s_i *corresponds to a distinct choice of the signs* c_I *for the algebra elements in equation (4.17). Moreover, the number of subalgebras in this decomposition is given by* $H^1(\mathcal{M}; \mathbb{Z}_2)$ *which counts the number of spin structures that* $\mathcal{M}$ *admits.*

Proof. Each $P(c_I)$ is composed of endomorphisms and hence also defines an endomorphism of the algebra. Properties i) and ii) of proposition 4.6.3 therefore show that the algebra decomposes as a product of subalgebras. Property iii) of proposition 4.6.3 shows that any other projection defined by (4.16) must project down to one of these subalgebras, hence these are all the subalgebras we get from elements of $\mathcal{Y}$.

A choice of signs for these projections can be thought of as a choice of a homomorphism from $\widehat{\pi}_1(\mathcal{M}, p)$ to $\mathbb{Z}_2$, since a homomorphism is defined by its action on generators. Any element of $\pi_1(\mathcal{M}, p)$ in the subgroup $\pi_1(\mathcal{M}, p)^2$ will belong to the kernel of any homomorphism from $\pi_1(\mathcal{M}, p)$ to $\mathbb{Z}_2$. This means there is a one to one correspondence between

$\mathrm{Hom}(\widehat{\pi}_1(\mathcal{M}, p), \mathbb{Z}_2)$ and $\mathrm{Hom}(\pi_1(\mathcal{M}, p), \mathbb{Z}_2)$. The latter corresponds to the number of elements in $H^1(\mathcal{M}; \mathbb{Z}_2)$ (see corollary 1.1.15). This means each choice of signs, and hence each of the subalgebras in (4.18), corresponds to an element of $H^1(\mathcal{M}; \mathbb{Z}_2)$. Thus we have proven that $\mathscr{U}_E^+(\mathbf{G})$ decomposes into subalgebras which are in one to one correspondence with the elements of $H^1(\mathcal{M}; \mathbb{Z}_2)$. $\square$

We have therefore shown that the universal algebra constructed from "even" Fermionic theories restricted to contractible regions of spacetime contains information about the spin structures that the spacetime admits.

Conclusions

In this thesis have used LCQFT to investigate topological aspects of QFT, and in particular we have focused on local to global constructions in QFT. The local to global construction we investigate in this thesis is the universal algebra construction due to Fredenhagen [Fre90], which can be used to extend a theory defined on a special subclass of spacetimes to a larger class of spacetimes. We have investigated the extended theories, which we call universal theories, obtained by the universal algebra construction for a class of linear theories modelled on the free scalar/Dirac field defined on contractible regions of spacetime. In chapter 3 we proved the universal theories for the full field algebras (see definition 3.1.5) are equivalent to their original theories, and that regardless of whether or not we consider the even subtheory that assigns even algebras (see definition 3.1.6), each of these universal theories satisfies Einstein causality.

In order to prove these results, geometric constructions involving Cauchy balls connected via tubes were introduced. A similar result for connecting disjoint contractible regions of a Cauchy surface within a single contractible region, was proven by Lang [Lan12, Lemma 1.1.6]. However Lang's result does not apply to arbitrary disjoint contractible regions, instead the regions must be part of a *good cover*. The contractible region that connects the two disjoint contractible regions is not necessarily part of the good cover, hence only two contractible regions can be connected using Lang's techniques. With the techniques introduced in this thesis, we can connect any finite number of Cauchy balls within a larger Cauchy ball. This has applications to other local-to-global QFT constructions, where relations that hold in local algebras need to be extended to the global algebra.

Further work could be done to see if these universal theories also satisfy the time slice axiom. This would prove that these theories satisfy the two main axioms that a theory of LCQFT should satisfy. There is a transformation on sections which takes a section to another one localised near a given Cauchy surface, such that they both index the same algebra element (see lemma 3.3.1). The difficulty in proving timeslice comes from the fact that this transformation increases the support of the section it's applied to. This means that the transformed sections that are localised near a given Cauchy surface may have support which is no longer contained in a contractible region. This is problematic because relations in the universal algebra can only be implemented by means of embedding generators into common local algebras, hence we cannot directly relate algebra elements to their corresponding elements localised near a given Cauchy surface if there is no contractible region of spacetime whose corresponding algebra contains both of them.

It may be possible to circumvent this issue by transforming the section that indexes a given algebra element in steps, between Cauchy surfaces that each differ by a small time step Δt, as in proposition 4.3.7 for example, and at each time step using linearity to break up the sections into parts with smaller support so that their support does not get too large to be contained in a contractible region. We could then implement local algebra relations at each time step to get a chain of relations relating a generator to its corresponding localisation near a given Cauchy surface. The potential problem with this method is that in the even algebra case the algebra elements are indexed by pairs of sections, so the algebra elements cannot be localised to arbitrarily small regions of spacetime. The even algebra elements can however be localised to arbitrarily thin tubular regions which connect the supports of the pair of sections; perhaps this would be enough to ensure that at each time step there is a local algebra containing a the supports of a section localised near the Cauchy surface at t and its transformation to a section localised near the Cauchy surface at $t + \Delta t$. One possible issue with this method would be if the tubes connecting pairs of sections got closer and closer to forming a closed loop, so that at some time step it would no longer be possible to get from the Cauchy surface at t to the Cauchy surface at $t + \Delta t$ without expanding the tube in such a way that it formed a closed loop, thus making it a non-contractible region.

In chapter 4 we then focused on the universal theories corresponding to linear "even" Fermionic theories. We proved that the corresponding universal theory assigns a universal algebra to each spacetime $\mathcal{M}$ which decomposes into a product of subalgebras, and that these subalgebras were in one to one correspondence with $H^1(\mathcal{M}, \mathbb{Z}_2)$, which counts the number of distinct spin structures that $\mathcal{M}$ admits. We did this by constructing a set of elements $\mathcal{Y}$ of the universal algebra that were eventually used to decompose the universal algebra. These elements were constructed using generators $b(f, g)$ of the local algebras satisfying some relations, among them being the relation

$$b(f, g)b(g, h) = b(f, h) \quad .$$

This then allowed us to relate the elements of $\mathcal{Y}$ to equivalence classes of quadruples (see definitions 4.1.2 and 4.2.1), which we were then able to show were in bijective correspondence to the set of loops identified up to orientation and free homotopy (see definition 4.2.3). This required us to establish further geometrical tools, including results in Lorentzian geometry and differential topology. This equivalence then allowed us to establish some further properties of the elements of $\mathcal{Y}$, in particular that they are central and square to the identity, which we could then use to construct projection operators to subalgebras.

The above equation seems to resemble a form of cocycle condition in a cohomology theory. This is interesting because it was also noted at the end of section 4.1 that in the case of a universal algebra built from local field algebras, the corresponding central elements of the universal algebra all collapse to the identity. This was because the local algebra generators $b(f, g)$ were of the form

$$b(f, g) = F(f)F(g)^{-1} \quad .$$

This seems to indicate that to construct non-trivial central elements in the universal algebra, it is sufficient that there be elements $b(f,g)$ in each local algebra that belong to some sort of non-trivial cohomology class. The only way to obtain a non-trivial decomposition of the universal algebra is to have non-trivial central elements, so it is the existence of a non-trivial cohomology class in the even subalgebras which allowed us to obtain spin structure information from the universal algebra. This suggests that this analysis could be generalised to other types of theory, if it can be demonstrated that the theory in question assigns local algebras which have a non-trivial cohomology class in the above sense. It may also be possible to construct higher non-trivial cohomology classes for the algebras assigned by linear even Fermionic theories, and therefore encode further topological information in the universal algebra.

This may be related the the net cohomology construction outlined in [BR09]. In this framework, given a poset one can construct a cohomology associated to it, and from this products of 1-cycles can be defined to give a notion of paths, which then leads to a definition of a fundamental group associated to any poset. For a poset given by a basis of arcwise simply-connected subsets of a manifold $\mathcal{M}$, the associated fundamental group coincides with the fundamental group of $\mathcal{M}$. Similarly, a cohomology with coefficients that are unitary operators in a Hilbert space can be defined on a poset. Given a net of C^* algebras over a poset K, a unitary net representation (see the paper for details on the definition) of this net defines a 1-cocycle in the cohomology associated to K. Moreover, [BR09, Lemma 2.1] shows that any two unitarily equivalent net representations define equivalent 1-cocycles. For the net of local C^* algebras associated to the poset of Cauchy diamonds of a fixed spacetime $\mathcal{M}$, each unitary net representation of this net corresponds to a unitary representation of the fundamental group of $\mathcal{M}$ in a fixed Hilbert space. This gives a useful means of analysing the topological content of unitary net representations. The key difference between this net cohomology and the algebra cohomology outlined in the previous paragraph is that the former pertains to the unitary transformations that relate the associated Hilbert space representations of C^* algebras associated to different poset elements, while the latter is inherent to the structure of the local algebras and independent of the mappings between them.

Another direction for future work will be in analysing the subalgebras that the universal algebra decomposes into. We conjecture that the subalgebra corresponding to the homomorphism that assigns $+1$ to each loop, is isomorphic to the global even subalgebra whose net of local algebras was used to construct the universal algebra. This subalgebra corresponds to the subalgebra where all the elements of $\mathcal{Y}$ get projected to the identity element, which means that algebra elements that differ only by which local algebra they are mapped into the universal algebra from, will be equivalent. One must still show that given two generators of the universal algebra, one can (by splitting into pieces if necessary) show that the commutation relations are equivalent to those in the original even subalgebra.

We also conjecture that for the Dirac field (or any other theory defined over spin bundles), each factor will be isomorphic to the even subalgebra equipped with a different spin structure. In particular, the signs that specify each factor in (4.18) will also specify the spin structure of the even subalgebra that the factor is isomorphic to. Recall that the signs that specify a subalgebra determine whether the projection of Y_g to that subalgebra is plus or minus the identity for each loop g. Loops for which $Y_g = -\mathbb{1}$ will also be loops where the spin structures of the original global algebra and the spin structure of the subalgebra disagree, in the sense that Fermions that undergo parallel transport around the loop g will differ by a sign. It is interesting to note that we have not assumed that the second Stiefel-Whitney class $w_2(T\mathcal{M}) \in H^2(\mathcal{M}, \mathbb{Z}_2)$ is trivial for the spacetimes that we consider, hence the results proven in this thesis hold even without this assumption. This condition is required for $\mathcal{M}$ to admit a spin structures.

Bibliography

[AMR88] R. Abraham, J.E. Marsden, and T. Ratiu. Manifolds, Tensor Analysis, and Applications. Springer-Verlag New York, 1988.

[BFIR14] R. Brunetti, K. Fredenhagen, P. Imani, and K. Rejzner. The locality axiom in quantum field theory and tensor products of C*-algebras. *Rev. Math. Phys.*, **26**(06):1450010, (2014).

[BFM09] R. Brunetti, L. Franceschini, and V. Moretti. *Topological Features of Massive Bosons on Two-Dimensional Einstein Space-Time*. *Annales Henri Poincaré*, **10**:1027, (2009).

[BFV03] R. Brunetti, K. Fredenhagen, and R. Verch. *The generally covariant locality principle — a new paradigm for local quantum field theory*. *Commun. Math. Phys.*, **237**:31–68, (2003).

[BG11] C. Bär and N. Ginoux. *Classical and Quantum Fields on Lorentzian Manifolds*. *Springer Proc. Math.*, **17**:359–400, (2011).

[BH59] A. Borel and F. Hirzebruch. *Characteristic Classes and Homogeneous Spaces, II*. *American Journal of Mathematics*, **81**(2):315–382, (1959).

[BR09] R. Brunetti and G. Ruzzi. *Quantum charges and spacetime topology: The emergence of new superselection sectors*. *Commun. Math. Phys.*, **287**:523–563, (2009).

[BS05] A. Bernal and M. Sánchez. *Smoothness of Time Functions and the Metric Splitting of Globally Hyperbolic Spacetimes. Commun. Math. Phys.*, **257**:43–50, (2005).

[BS06] A. N. Bernal and M. Sanchez. *Further Results on the Smoothability of Cauchy Hypersurfaces and Cauchy Time Functions*. *Letters in Mathematical Physics*, **77**(2):183–197, (2006).

[BS07] A. Bernal and M. Sánchez. *Globally hyperbolic spacetimes can be defined as 'causal' instead of 'strongly causal'. Classical and Quantum Gravity*, **24**(3):745–749, (2007).

[BSW17] M. Benini, A. Schenkel, and L. Woike. *Operads for algebraic quantum field theory*. *arXiv:1709.08657 [math-ph]*, (2017).

[BW15] C. Bär and R. Tagne Wafo. *Initial Value Problems for Wave Equations on Manifolds. Mathematical Physics, Analysis and Geometry*, **18**, 2015.

[DHR69a] S. Doplicher, R. Haag, and J. E. Roberts. *Fields, observables and gauge transformations I. Commun. Math. Phys.*, **13**(1):1–23, (1969).

[DHR69b] S. Doplicher, R. Haag, and J. E. Roberts. *Fields, observables and gauge transformations II. Commun. Math. Phys.*, **15**(3):173–200, (1969).

[DHR71] S. Doplicher, R. Haag, and J. E. Roberts. *Local observables and particle statistics I. Commun. Math. Phys.*, **23**(3):199–230, (1971).

[DHR74] S. Doplicher, R. Haag, and J. E. Roberts. *Local observables and particle statistics II. Commun. Math. Phys.*, **35**(1):49–85, (1974).

[DHS14] C. Dappiaggi, T. Hack, and K. Sanders. *Electromagnetism, local covariance, the Aharonov-Bohm effect and Gauss' law*. *Commun. Math. Phys.*, **328**:625–667, (2014).

[Dim82] J. Dimock. *Dirac quantum fields on a manifold*. *Trans. Am. Math. Soc.*, **269**:133, (1982).

[DL12] C. Dappiaggi and B. Lang. *Quantization of Maxwell's Equations on Curved Backgrounds and General Local Covariance. Letters in Mathematical Physics*, **101**(3):265–287, (2012).

[DM04] M. Davier and W. J. Marciano. *The theoretical prediction for the muon anomalous magnetic moment. Annual Review of Nuclear and Particle Science*, **54**(1):115–140, (2004).

[Few15] C. J. Fewster. *The split property for locally covariant quantum field theories in curved spacetime*. *Letters in Mathematical Physics*, **105**(12):1633–1661, (2015).

[FR19] C. J. Fewster and K. Rejzner. *Algebraic Quantum Field Theory – an introduction. arXiv:1904.04051 [hep-th]*, (2019).

[Fre90] K. Fredenhagen. Generalizations of the theory of superselection sectors. The Algebraic Theory of Superselection Sectors, World Scientific Publishing, (1990).

[FV12] C. J. Fewster and R. Verch. *Dynamical Locality and Covariance: What Makes a Physical Theory the Same in all Spacetimes? Annales Henri Poincaré*, **13**:1613–1674, (2012).

[FV15] C. J. Fewster and R. Verch. Algebraic Quantum Field Theory in Curved Spacetimes, pages 125–189. Springer International Publishing, Cham, (2015).

[Ger70] R. Geroch. *Spinor Structure of Space-Times in General Relativity, II*. *Journal of Mathematical Physics*, **11**:343, (1970).

[Haa96] R. Haag. Local quantum physics. Springer-Verlag Berlin Heidelberg, second edition, 1996.

[Haw74] S. Hawking. *Black hole explosions? Nature*, **248**:30–31, (1974).

[Hir76] M. W. Hirsch. Differential topology. Springer-Verlag New York, 1976.

[HK64] R. Haag and D. Kastler. *An Algebraic Approach to Quantum Field Theory. Journal of Mathematical Physics*, **5**(7):848–861, 1964.

[HM19] R. A. Hounnonkpe and E. Minguzzi. *Globally hyperbolic spacetimes can be defined without the 'causal' condition. Classical and Quantum Gravity*, **36**(19), 2019.

[Hov52] Léon Van Hove. *Les difficultés de divergences pour un modèle particulier de champ quantifié. Physica*, **18**:145–159, (1952).

[Ish78] C. J. Isham. *Spinor Fields in Four Dimensional Space-Time. Proc. R. Soc. Lond. A*, **364**:591–599, (1978).

[Kli11] Wilhelm P.A. Klingenberg. Riemannian Geometry. De Gruyter, 2011.

[Kos80] C. Kosniowski. A First Course in Algebraic Topology. Cambridge University Press, 1980.

[Kos93] A. A. Kosinski. Differential manifolds. Academic Press, 1993.

[Lan12] Benjamin Lang. *Universal constructions in algebraic and locally covariant quantum field theory.* PhD thesis, University of York, 2012. (doctoral thesis) `http://etheses.whiterose.ac.uk/id/eprint/8019`.

[Lee03] J. M. Lee. Introduction to Smooth Manifolds. Springer-Verlag New York, 2003.

[Mac78] S. MacLane. Categories for the working mathematician. Springer-Verlag New York, second edition, 1978.

[Mil63a] J. Milnor. Morse theory. Princeton University Press, 1963.

[Mil63b] J. Milnor. *Spin structures on manifolds. L'Enseignement Mathématique*, **9**:198–203, (1963).

[O'N83] B. O'Neill. Semi-Riemannian Geometry With Applications to Relativity. Academic Press, 1983.

[Rot88] J. Rotman. An Introduction to Algebraic Topology. Springer-Verlag New York, 1988.

[RS81] M. Reed and B. Simon. Methods of Modern Mathematical Physics, volume 1 of *Methods of Modern Mathematical Physics.* Academic Press, 1981.

[Rud86] W. Rudin. Principles of Mathematical Analysis. International Series in Pure & Applied Mathematics. McGraw - Hill, third edition, 1986.

[San10] K. Sanders. *The locally covariant Dirac field. Rev. Math. Phys.*, **22**:381, (2010).

[Spa82] E. Spanier. Algebraic Topology. Springer-Verlag New York, 1982.

[Ste51] N. Steenrod. The topology of fibre bundles. Princeton University Press, 1951.

[Unr76] W. G. Unruh. *Notes on black-hole evaporation. Phys. Rev. D*, **14**:870–892, (1976).

[Vok] L. Vokřínek. Differential topology [lecture notes]. URL: `http://www.math.muni.cz/~koren/DifferentialTopology.pdf`. Last visited on 01/09/2020.

[Wal84] R. M. Wald. General Relativityy. University of Chicago Press, 1984.

[WG65] A.S. Wightman and L. Garding. *fields as operator-valued distributions in relativistic quantum field theory. Arkiv Fys.*, **28**, 1965.