Quantum Field Theory in Curved Spacetime

Lecture notes for a mini-course given at the
Poznań Spring School in Mathematical Physics
held from 07.04.-16.04. 2025 at the AMU in Poznań, Poland

Ko Sanders

April 16, 2025

Preface

Quantum field theories (QFTs) in curved spacetimes allow the description of a vast spectrum of physical phenomena, ranging from the standard model of elementary particle physics to gravitational effects on quantum fields, such as the particle creation due to the expansion of the universe and the Hawking radiation of black holes. They offer a reliable setting for studying some of the most advanced theories of fundamental physics and they serve as a stepping stone for studying quantum gravity in a semi-classical setting, where the classical metric is no longer fixed, but responds to the presence of the quantum fields. The properties of the quantum stress-energ-momentum tensor are particularly interesting, because they intermediate between (perturbative) quantum gravity and classical physics. This can be seen especially in the case of quantum energy inequalities (QEIs), which have been verified for a range of QFTs and roughly express the idea that energy densities cannot be too negative for too long. By taking a suitable classical limit, these quantum energy inequalities could potentially allow a first-principles derivation of the classical energy conditions (ECs) that appear in general relativity (GR) to express typical behaviour of large classes of classical matter theories. These ECs have notable consequences for the mathematical structure of GR and for its physical predictions, e.g. the celebrated singularity theorems of Hawking and Penrose, for which Penrose received the 2020 Nobel prize in physics.

These lecture notes aim to provide an up-to-date introduction to the mathematical structure of Lagrangian QFTs in curved spacetimes. For the algebraic structure of Lagrangian QFTs we will follow the recent ideas of Buchholz and Fredenhagen [7] and we will take a modern prespective on the categorical structure of localisation regions also for classical GR.

Throughout these notes we assume that the reader has some familiarity with differential geometry, partial differential equations and, ideally, some distribution theory, as well as a good working knowledge of linear algebra and functional analysis on Banach and Hilbert spaces. A general interest in physics is expected and some basic knowledge of quantum physics and special or general relativity theory would be beneficial. For backgrounud material and further reading in this area of physics I recommend the following book and review papers: [37] (GR), [11, 18] (algebraic and perturbative QFT), [1] (QFT in curved spacetimes) and [23, 15, 10] (QEIs).

We will use Planck units throughout, i.e. $c = \hbar = G = 1$.

Contents

Preface **1**

1 Classical field theory **5**
1.1 Classical Lagrangian field theory . . . 6
1.2 Energy conditions in general relativity . . . 10
1.3 More about spacetime . . . 15
1.3.1 Well-posedness and global hyperbolicity . . . 15
1.3.2 Locality and general covariance . . . 16

2 Quantum field theory **20**
2.1 Algebraic quantum theory . . . 20
2.1.1 Algebras and states . . . 20
2.1.2 Algebras with additional properties . . . 23
2.2 Lagrangian QFTs . . . 25
2.3 Toy model: the free scalar field . . . 30
2.3.1 The classical scalar field . . . 30
2.3.2 The quantized free scalar field . . . 33
2.4 Locally covariant QFT . . . 35
2.4.1 Lagrangian QFTs as LCQFTs . . . 37

3 The stress tensor in QFT **40**
3.1 Metric perturbations in LCQFT . . . 40
3.1.1 Example: the free scalar field . . . 42
3.1.2 The stress tensor . . . 44
3.2 The renormalized quantum stress tensor for free scalar fields . . . 45
3.3 Energy conditions and QEIs in QFT . . . 51
3.3.1 Violation of the NEC in QFT . . . 51
3.3.2 Quantum energy inequalities . . . 53

3.3.3 Sketch of proof of theorem 3.3.2 55

Bibliography **57**

Chapter 1

Classical field theory

In physics, fields are spacetime dependent quantities that are used to describe disturbances of systems that propagate through spacetime from point to neighbouring point. Maxwell's theory of electromagnetism is a prime example and a milestone in theoretical physics, although its modern formulation in terms of vector potentials is complicated by the presence of gauge symmetries. Einstein's theory of general relativity (GR) describes gravity in terms of a Lorentzian metric field, which overcomes some drawbacks of Newton's theory of gravity. When we want to simplify matters, however, we will focus on much simpler scalar field models. The propagation of disturbances of the field is described by a partial differential equation (PDE), the equation of motion, which the fields must satisfy. A common feature of all theories that we will consider, is that these equations can be derived from a Lagrange density.

Before we proceed, let us first summarise the basic setting of GR. Let $\mathcal{M}$ be a smooth[1] manifold of dimension $d \geq 2$ (where $d = 4$ is of special interest in physics) and let g_{ab} be a smooth Lorentzian metric field.[2] I.e., at

[1]The assumption that $\mathcal{M}$ and the Lorentzian metric g_{ab} are smooth can be weakened considerably, but the assumption is convenient, because it obviates the cumbersome task of keeping track of the minimal amount of regularity required for various arguments in the remainder of these notes.

[2]We use the abstract index notation conventions of [37]. The placement of the indices indicates the type of a tensor. A vector carries an upper index, X^a, whereas a dual vector carries a lower index, ω_a. One may think of the indices as referring to components in some coordinate basis, but the basis may be chosen arbitrarily and all equations with abstract indices are valid regardless of this choice. We use the summation convention that any index that appears both as an upper and as a lower index is summed over. E.g. $\omega_a X^a$

every point $x \in \mathcal{M}$, $(g_x)_{ab}$ determines a non-degenerate symmetric quadratic form on the tangent space $T_x\mathcal{M}$, which has the signature $(- + \ldots +)$. This generalises Minkowski space, where $\mathcal{M} = \mathbb{R}^d$ and g_{ab} is given in inertial coordinates by the constant diagonal matrix $\mathrm{diag}(-1, +1, \ldots, +1)$. We will call a vector X^a timelike when $g_{ab}X^aX^b < 0$, spacelike when $g_{ab}X^aX^b > 0$ and null when $g_{ab}X^aX^b = 0$ but $X^a \neq 0$.

The Lorentzian metric field g_{ab} admits a unique compatible and torsion free covariant derivative, ∇_a (corresponding to the Levi-Civita connection). This covariant derivative gives rise to a Riemann curvature tensor ${R_{abc}}^d$, such that

$$ {R_{abc}}^d X^c = -(\nabla_a \nabla_b - \nabla_b \nabla_a) X^d \tag{1.1} $$

for any smooth vector field X^a on $\mathcal{M}$. Related to the Riemann curvature tensor are the Ricci curvature tensor $R_{ac} = {R_{abc}}^b$ and the scalar curvature $R = g^{ac}R_{ac}$, where g^{ac} denotes the inverse metric on $T^*\mathcal{M}$.

From now on we will assume that $\mathcal{M}$ is oriented and we will always use oriented coordinate charts, so changes of coordinates are always orientation preserving (their Jacobi-matrix has a positive determinant). In this case the Lorentzian metric field g_{ab} also determines a natural volume form on $\mathcal{M}$, i.e. a non-vanishing differential form of the highest rank d, which is given in local coordinates by

$$ \mathrm{d}vol_g := \sqrt{|\det(g_{\mu\nu})|}\, \mathrm{d}\, x_0 \wedge \ldots \wedge \mathrm{d}\, x_{d-1}\,. \tag{1.2} $$

One can check that this formula is invariant under orientation preserving changes of coordinates, whereas it gets a sign under orientation reversing changes of coordinates.

1.1 Classical Lagrangian field theory

A real scalar field is a spacetime dependent quantity on $\mathcal{M}$ with values in $\mathbb{R}$. Any function $\phi \in C^\infty(\mathcal{M}, \mathbb{R})$ is a configuration of the field and we call $C^\infty(\mathcal{M}, \mathbb{R})$ its *configuration space*. Typically ϕ is unknown and should be found by solving a suitable equation of motion with appropriate initial data

denotes the action of the dual vector ω_a on the vector X^a, whereas $\omega_a X^b$ is a tensor product (leading to a tensor of mixed type). For components in a particular choice of coordinate basis we use lower case Greek letters, e.g. X^μ with $\mu = 0, \ldots, d-1$.

or boundary values. More generally we can consider tensor fields, such as the metric g_{ab}, or sections of a suitable (vector) bundle over $\mathcal{M}$, such as spinor fields.

The dynamics can be obtained from an action principle. In general we consider a *Lagrange density* with several terms that describe the dynamics and interactions of the fields. E.g., when only the metric g_{ab} is present we will take the *Einstein-Hilbert Lagrange density*

$$\mathcal{L}_{\mathrm{EH}}[g] = \frac{1}{16\pi}(R - 2\Lambda)\ \mathrm{d}vol_g\,, \tag{1.3}$$

where $\Lambda \in \mathbb{R}$ is called the cosmological constant, R is the scalar curvature of g and we suppress the index structure of the metric as an argument of the functional. If further fields are present (some kind of "matter"), which we generically denote by Φ, we consider the total Lagrange density

$$\mathcal{L}[g,\Phi] = \mathcal{L}_{\mathrm{M}}[g,\Phi] - \mathcal{L}_{\mathrm{EH}}[g]\,, \tag{1.4}$$

where $\mathcal{L}_{\mathrm{M}}[g,\Phi]$ is the Lagrange density that describes the matter.

The Lagrange densities that we will allow should be local functionals of the fields, so at any point $x \in \mathcal{M}$ they only depend on the fields g_{ab} and Φ and their derivatives at that point. In particular, if Φ denotes all fields, including the metric field g_{ab}, then they should be of the form

$$\mathcal{L}[\Phi](x) = F(x, \Phi(x), \partial_\mu\Phi(x), \ldots, \partial_{\mu_1}\cdots\partial_{\mu_n}\Phi(x))\ \mathrm{d}vol_g(x) \tag{1.5}$$

for some $n \geq 0$ and some smooth function F.

The equations of motion are the *Euler-Lagrange equations* determined by the Lagrange density $\mathcal{L}[g,\Phi]$. To find these equations in a bounded open region $O \subset \mathcal{M}$ we consider the *local action*

$$S(f)[g,\Phi] = \int_{\mathcal{M}} f \cdot \mathcal{L}[g,\Phi]\,, \qquad f|_O \equiv 1\,, \tag{1.6}$$

where the test-function $f \in C_0^\infty(\mathcal{M})$ ensures that the integral converges. For any field we find its equation of motion in O by requiring that the local action $S(f)[g,\Phi]$ is a stationary point for compactly supported perturbations of the field in O. This amounts to setting the functional derivatives $\frac{\delta}{\delta g^{ab}}S(f)[g,\Phi]$ and $\frac{\delta}{\delta\Phi}S(f)[g,\Phi]$ equal to zero in O. Here Φ can have multiple components and hence so can the corresponding Euler-Lagrange equation. Physically, all these equations should hold simultaneously.

Remark 1.1.1. Classically one would consider the action (1.6) with $f \equiv 1$, which does not have a compact support. The localized action $S(f)[g]$ will do equally well to determine the equations of motion, because we can adapt f to the region O of interest. Note that the equations of motion in O are independent of the choice of f, as long as $f|_O \equiv 1$, due to the locality of the Lagrange density.

As an example we consider a real scalar field ϕ with Lagrange density

$$\mathcal{L}_{\mathrm{sf}}[g,\phi] = \left(\frac{1}{2}g^{ab}\partial_a\phi \cdot \partial_b\phi + \frac{1}{2}m^2\phi^2 + \frac{1}{2}\xi R\phi^2 + V(\phi)\right) \mathrm{d}vol_g\,, \qquad (1.7)$$

where $m \geq 0$ denotes the mass of the field, ξ is the scalar curvature coupling constant and V is some (smooth) potential energy function, e.g. $V(\phi) = \lambda\phi^4$ for some coupling constant λ. If no other matter fields are present, then $\mathcal{L}_{\mathrm{M}} = \mathcal{L}_{\mathrm{sf}}$ and the Euler-Lagrange equation can be computed to be the modified Klein-Gordon equation

$$-\Box\phi + m^2\phi + \xi R\phi + V'(\phi) = 0\,, \qquad (1.8)$$

where $\Box = \nabla^a\nabla_a$ is the *Laplace-Beltrami operator* (which generalises the d'Alembert wave operator in Minkowski space). Indeed, writing $\tilde{V}(\phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{2}\xi R\phi^2 + V(\phi)$ we have for any compactly supported and smooth perturbation ϕ_0 in O

$$\begin{aligned}
&\lim_{\epsilon\to 0}\frac{S_{\mathrm{sf}}(f)[g,\phi+\epsilon\phi_0] - S_{\mathrm{sf}}(f)[g,\phi]}{\epsilon} \\
&= \lim_{\epsilon\to 0}\int_{\mathcal{M}} f\cdot\left(\frac{1}{2}\nabla^a\phi_0\cdot\nabla_a\phi + \frac{1}{2}\nabla^a\phi\cdot\nabla_a\phi_0 + \frac{1}{2}\epsilon\nabla^a\phi_0\cdot\nabla_a\phi_0\right. \\
&\qquad\qquad \left. + \frac{\tilde{V}(\phi+\epsilon\phi_0) - \tilde{V}(\phi)}{\epsilon}\right)\mathrm{d}vol_g \\
&= \int_{\mathcal{M}} f\cdot\left(\frac{1}{2}\nabla^a\phi_0\cdot\nabla_a\phi + \frac{1}{2}\nabla^a\phi\cdot\nabla_a\phi_0 + \phi_0\tilde{V}'(\phi)\right)\mathrm{d}vol_g \\
&= \int_{\mathcal{M}} \phi_0\left(-\Box\phi + m^2\phi + \xi R\phi + V'(\phi)\right)\mathrm{d}vol_g
\end{aligned}$$

by partial integration, because $f \equiv 1$ on the support of ϕ_0. The *variational principle* tells us that this vanishes for all ϕ_0 iff

$$\frac{\delta}{\delta\phi}S_{\mathrm{sf}}(f)[g,\phi] = \left(-\Box\phi + m^2\phi + \xi R\phi + V'(\phi)\right)\mathrm{d}vol_g$$

vanishes on O and hence also (1.8) holds. Varying the region and the function f we see that the equation must hold on $\mathcal{M}$.

The metric appears not only in $\mathcal{L}_{\mathrm{EH}}[g]$ as defined in (1.3), but also in the Lagrange density of the matter fields, e.g. through the volume form $\mathrm{d}vol_g$, covariant derivatives or through inner products of vector or tensor fields. The Euler-Lagrange equation for the metric can therefore be expressed as

$$\frac{\delta}{\delta g^{ab}} S_{\mathrm{EH}}(f)[g] = \frac{\delta}{\delta g^{ab}} S_{\mathrm{M}}(f)[g, \Phi]$$

where $S_{\mathrm{EH}}(f)[g] = \int_{\mathcal{M}} f \cdot \mathcal{L}_{\mathrm{EH}}[g]$ and similarly for $S_{\mathrm{M}}(f)[g, \Phi]$. Note that the right-hand side depends on the matter theory at hand, but the left-hand side can be computed directly by an explicit (and rather lengthy) computation. Together this leads to *Einstein's equations*

$$R_{ab} - \frac{1}{2} g_{ab} R + \Lambda g_{ab} = 8\pi T_{ab}[g, \Phi] \tag{1.9}$$

where the stress-energy-momentum tensor $T_{ab}[g, \Phi]$ (stress tensor for short) of the matter is defined in local coordinates on O by

$$T_{\mu\nu}[g, \Phi] \; \mathrm{d}vol_g := 2 \frac{\delta}{\delta g^{\mu\nu}} S_{\mathrm{M}}(f)[g, \Phi] \,, \tag{1.10}$$

where the choice of f is irrelevant, as long as $f \equiv 1$ on O.

E.g., for the scalar field Lagrange density (1.7) we find

$$\begin{aligned} T_{ab}[g, \phi] = \partial_a \phi \cdot \partial_b \phi - g_{ab} \left(\frac{1}{2} \nabla^c \phi \cdot \nabla_c \phi + \frac{1}{2} m^2 \phi^2 + \frac{1}{2} \xi R \phi^2 + V(\phi) \right) \\ + \xi (R_{ab} - \nabla_a \nabla_b + g_{ab} \Box) \phi^2 \,. \end{aligned} \tag{1.11}$$

The second term on the first line is $-g_{ab} \mathcal{L}_{\mathrm{sf}}$, except that we omitted the metric volume form. This term comes from the variation of the volume form. When $\xi = 0$ we call the field minimally coupled and the stress tensor simplifies to

$$T_{ab}[g, \phi] = \partial_a \phi \cdot \partial_b \phi - g_{ab} \left(\frac{1}{2} \nabla^c \phi \cdot \nabla_c \phi + \frac{1}{2} m^2 \phi^2 + V(\phi) \right) \,. \tag{1.12}$$

1.2 Energy conditions in general relativity

In this section we will consider a number of properties of the classical stress tensor as it appears in Einstein's equations (1.9). One can show that the left-hand side of that equation is *symmetric* under the exchange of the indices a, b and that it is *conserved*,

$$\nabla^a \left(R_{ab} - \frac{1}{2} g_{ab} R + \Lambda g_{ab} \right) = 0 \,.$$

This follows from the symmetries of the curvature tensor and the contracted Bianchi identity. Consequently, in order to have solutions, the stress tensor $T_{ab}[g, \Phi]$ must have these same properties. The symmetry of $T_{ab}[g, \Phi]$ is clear from its definition as a functional derivative in (1.10), because $g^{ab} = g^{ba}$. Under suitable circumstances that same definition implies that the stress tensor is conserved. To see this we pick O and f as before and a compactly supported vector field χ^a with support in O. The vector field generates a flow, which is a one-parameter group of diffeomorphisms $\Psi_s : \mathcal{M} \to \mathcal{M}$ such that Ψ_0 is the identity and $\partial_s \Psi_s^\mu(x)|_{s=0} = \chi^\mu(x)$ in any local coordinates. Now suppose that the diffeomorphisms act in a natural way on the fields Φ, so the pullback $\Psi_s^* \Phi$ is well-defined. (For scalar and tensor fields this is always the case, but for Dirac fields, e.g., we need to lift Ψ_s to a bundle isomorphism of the spinor bundle of which the field is a section.) Suppose furthermore, that $S_M(f)[g, \Phi]$ is invariant under the diffeomorphisms Ψ_s, i.e. $S_M(f)[\Psi_s^* g, \Psi_s^* \Phi] = S_M(f)[g, \Phi]$. (Note that $\Psi_s^* f = f$.) Then we find

$$\begin{aligned} 0 &= \partial_s S_M(f)[\Psi_s^* g, \Psi_s^* \Phi]|_{s=0} \\ &= \int_{\mathcal{M}} \frac{1}{2} (\partial_s \Psi_s^* g|_{s=0})^{ab} T_{ab}[g, \Phi] \mathrm{d}vol_g + (\partial_s \Psi_s^* \Phi|_{s=0}) \frac{\delta}{\delta \Phi} S_M(f)[g, \Phi] \,. \end{aligned}$$

If the field configuration (g, Φ) is a solution to the equations of motion for Φ, then the second term in the integral vanishes. Using the fact that $(\partial_s \Psi_s^* g|_{s=0})^{ab} = -\nabla^a \chi^b - \nabla^b \chi^a$, which is the Lie-derivative $\mathcal{L}_\chi g^{ab}$, we find

$$\begin{aligned} 0 &= \int_{\mathcal{M}} -2(\nabla^a \chi^b) T_{ab}[g, \Phi] \mathrm{d}vol_g \\ &= \int_{\mathcal{M}} 2\chi^b \nabla^a T_{ab}[g, \Phi] \mathrm{d}vol_g \end{aligned}$$

using the symmetry of the stress tensor and an integration by parts. By the variational lemma we must then have $\nabla^a T_{ab}[g, \Phi] \equiv 0$.

The stress tensor $T_{ab}[g,\Phi]$ can be interpreted in terms of the energy, momentum and stresses of the matter, cf. [37]. In particular, if we pick at some point $x \in \mathcal{M}$ a timelike unit vector t^a and any spacelike unit vector v^a in $T_x\mathcal{M}$ with $t^a v_a = 0$, then an observer whose worldline has a velocity vector t^a at x will measure the energy density $t^a t^b T_{ab}[g,\Phi]$ and the momentum density $t^a v^b T_{ab}[g,\Phi]$. These interpretations also allow us to express some general properties of matter in terms of the stress tensor components. We will consider here the main *energy conditions* that one finds in GR textbooks.

Weak energy condition: $T_{ab}[g,\Phi]$ satisfies the weak energy condition (WEC) iff for all timelike unit vectors t^a we have

$$t^a t^b T_{ab}[g,\Phi] \geq 0\,. \tag{1.13}$$

This says that the energy density is non-negative for all observers at all points. Boundedness from below is expected for any stable system and since the energy in classical mechanics can be shifted by an arbitrary amount, it seems natural in GR to set the minimum to 0.

Strong energy condition: $T_{ab}[g,\Phi]$ satisfies the strong energy condition (SEC) iff the dimension of $\mathcal{M}$ is $d > 2$ and for all timelike unit vectors t^a

$$t^a t^b \left(T_{ab}[g,\Phi] - \frac{1}{d-2} g_{ab} T^c{}_c[g,\Phi] \right) \geq 0\,. \tag{1.14}$$

The quantity on the left has been interpreted as an effective energy density or a relativistic analogue of the Newtonian potential. Physically, the SEC is not very compelling, but it has nice mathematical consequences, because in conjuction with Einstein's equations it implies $t^a t^b R_{ab} \geq 0$. Note that there is no implication between WEC and SEC in either direction.

Dominant energy condition: $T_{ab}[g,\Phi]$ satisfies the dominant energy condition (DEC) iff for all timelike unit vectors t^a and u^a with $t^a u_a < 0$ we have

$$t^a u^b T_{ab}[g,\Phi] \geq 0\,. \tag{1.15}$$

This says that the energy-momentum flow $-t^a T_{ab}[g,\Phi]$ should be a causal vector (i.e. timelike or null) which is future pointing exactly when t^a is. Note that DEC$\Rightarrow$WEC.

Null energy condition: $T_{ab}[g, \Phi]$ satisfies the null energy condition (NEC) iff for all null vectors n^a we have

$$n^a n^b T_{ab}[g, \Phi] \geq 0 \,. \tag{1.16}$$

NEC is weaker than WEC, SEC and DEC.

These energy conditions can be verified on a case by case basis. E.g., for a minimally coupled ($\xi = 0$) scalar field with Lagrange density (1.7) and potential $V \geq 0$ the stress tensor is given by (1.12). One may verify that it satisfies DEC and hence also WEC and NEC, but the SEC can be violated unless $m = 0$ and $V \equiv 0$. When $\xi \neq 0$, even the NEC can be violated (cf. [23]).

The energy conditions are linear in $T_{ab}[g, \Phi]$, so if an energy condition is satisfied for two Lagrange densities, then their sum will also satisfy it.

Note that Einstein's equations depend on the matter theory only through the stress tensor. The importance of the energy conditions is, that they allow us to analyse Einstein's equations for a wide range of matter models simultaneously. They were introduced for this reason by Penrose in the derivation of his singularity theorem [28], which won him the 2020 Nobel prize in physics. Many other results in mathematical relativity, concerning e.g. cosmology, singularities, black hole horizons and spacetime topology, assume one or several of the energy conditions, together with other assumptions, in order to prove a result.

In the remainder of this section we present an application of the energy conditions, namely Wald's "cosmic no hair theorem" [36]. For this we will consider a Lorentzian manifold of the form $\mathcal{M} = \mathbb{R} \times \Sigma$ with the metric

$$g = -\mathrm{d}\, t^2 + a(t)^2 h \,,$$

where (Σ, h) is some Riemannian manifold of dimension $d - 1 \geq 2$, $t \in \mathbb{R}$ and $a : \mathbb{R} \to \mathbb{R}_{>0}$ is a (smooth) scale factor. We will show that Einstein's equations (1.9) with a positive cosmological constant $\Lambda > 0$ and a stress tensor that satisfies the WEC and the SEC, together with suitable initial conditions, imply that the expansion rate function $\alpha(t) := \frac{a'(t)}{a(t)}$ converges exponentially to $\sqrt{\frac{2}{(d-1)(d-2)}\Lambda} > 0$ at late times. (The original result of [36] assumed in addition that the Lorentzian manifolds are homogeneous, giving an asymptotic approach to the de Sitter metric at late times.)

We will first use the energy conditions to deduce two inequalities concerning the geometry of $M = (\mathcal{M}, g)$. Using Einstein's equations we have

$$\begin{aligned} 8\pi \left(T_{ab} - \frac{1}{d-2} g_{ab} T^c{}_c \right) &= R_{ab} - \frac{1}{2} g_{ab} R + \Lambda g_{ab} - \frac{1}{d-2} g_{ab} \left(R - \frac{d}{2} R + d\Lambda \right) \\ &= R_{ab} - \frac{2}{d-2} \Lambda g_{ab} \,. \end{aligned}$$

Using this identity, the SEC imposes a restriction on the Ricci curvature,

$$0 \leq t^a t^b R_{ab} + \frac{2}{d-2} \Lambda \,,$$

for any future pointing timelike unit vector field t^a. Similarly, the WEC combined with Einstein's equations implies

$$0 \leq t^a t^b R_{ab} + \frac{1}{2} R - \Lambda \,.$$

Taken together, the SEC, WEC and Einstein's equations then yield

$$-t^a t^b R_{ab} \leq \frac{2}{d-2} \Lambda \leq \frac{1}{d-2} \left(2 t^a t^b R_{ab} + R \right) \,. \tag{1.17}$$

Now note that for the class of metrics that we consider here, the curvature of M can be expressed in terms of the curvature of (Σ, h), the scale factor $a(t)$ and the metric h. Near an arbitrary point x we choose coordinates such that $x_0 = t$ is the time coordinate and x^i for $i = 1, \ldots, d-1$ are coordinates on Σ. Writing the components of R_{ab} in these coordinates we find, after some algebra, for the Ricci tensor

$$\begin{aligned} R_{ij} &= \tilde{R}_{ij} + (a a'' + (d-2)(a')^2) h_{ij} \\ R_{00} &= -(d-1) \frac{a''}{a} \\ R_{0i} &= R_{i0} = 0 \,, \end{aligned}$$

where $\tilde{R}_{ij}$ denotes the Ricci curvature tensor of the Riemannian manifold (Σ, h) and a prime denotes a time derivative of a. Consequently,

$$R = \frac{1}{a^2} \tilde{R} + 2(d-1) \frac{a''}{a} + (d-2)(d-1) \left(\frac{a'}{a} \right)^2 \,,$$

where $\tilde{R}$ is the scalar curvature of (Σ, h). Using these identities we can rewrite the inequalities (1.17), taking for t^a the unit vector field along $\mathbb{R}$ so we get the (00)-components:

$$(d-1)\frac{a''}{a} \leq \frac{2}{d-2}\Lambda \leq \frac{\tilde{R}}{(d-2)a^2} + (d-1)\left(\frac{a'}{a}\right)^2 . \tag{1.18}$$

We now assume that $\Lambda > 0$ and we impose the initial conditions $\tilde{R} \leq 0$ and $a'(0) > 0$. We will rewrite the inequalities above in terms of the constant $C := \sqrt{\frac{2\Lambda}{(d-1)(d-2)}} > 0$ and the expansion rate function $\alpha(t) := \frac{a'(t)}{a(t)}$ with $\alpha(0) > 0$. Note that $\alpha' = \frac{a''}{a} - \alpha^2$, so we find

$$\alpha' \leq C^2 - \alpha^2 \leq 0\,.$$

The second inequality shows that the continuous function α can never cross 0, so $\alpha > 0$ for all time and hence $\alpha(t) \geq C$. Note that α is a decreasing function, so if $\alpha(0) = C$, then $\alpha(t) = C$ for all $t \geq 0$. If $\alpha(0) > C$, then we can write the first inequality as $2C \leq \frac{\alpha'}{\alpha+C} - \frac{\alpha'}{\alpha-C}$. Integrating this leads to

$$\begin{aligned} 2Ct &\leq \log\left(\frac{\alpha(t)+C}{\alpha(t)-C}\right) - 2C' \\ e^{2Ct+2C'} &\leq \frac{\alpha(t)+C}{\alpha(t)-C} = 1 + \frac{2C}{\alpha(t)-C} \\ \alpha(t) &\leq C + \frac{2C}{e^{2Ct+2C'}-1} \end{aligned}$$

for some $C' > 0$ depending on the initial value $\alpha(0)$ and C. It follows that $\alpha(t)$ approaches the value C exponentially.

Integrating the inequalities $C \leq \frac{a'}{a} \leq C + \frac{2C}{e^{2Ct+C'}-1}$ we similarly find

$$a(0)e^{Ct} \leq a(t) \leq a(0)\frac{\sinh(C'+Ct)}{\sinh(C')}\,.$$

With this and Einstein's equations one then shows that

$$8\pi T_{00} = R_{00} + \frac{1}{2}R - \Lambda = \frac{\tilde{R}}{2a^2} + \frac{1}{2}(d-2)(d-1)\left(\frac{a'}{a}\right)^2 - \Lambda$$

converges exponentially fast to 0. Note that $T_{0i} = T_{i0} = 0$ follows from the form of the Ricci tensor and the metric. If also the DEC holds, then any other component of the stress tensor can be bounded by $|T_{ij}| \leq T_{00}$ by choosing suitable timelike vectors. Hence, all components of the stress tensor vanish exponentially.

1.3 More about spacetime

1.3.1 Well-posedness and global hyperbolicity

The equations of motion that we derive from a Lagrange density are typically hyperbolic equations, which describe a wave-like propagation of disturbances through spacetime. Here the maximum speed is encoded in the Lorentzian metric g_{ab}: at every point disturbances must propagate within the future light-cone. For hyperbolic equations it is natural to consider the Cauchy problem, which poses initial data at some "time", that should uniquely determine a solution throughout the manifold. Typically, the solution depends continuously on the specified data (in some suitable topology) and we then speak of a *well-posed Cauchy problem*.

The Cauchy problem for GR coupled to other fields satisfying hyperbolic equations can be shown to be well-posed under fairly general circumstances, cf. [30, 9, 19]. Here one prescribes initial data on some smooth manifold S, satisfying certain constraints related to the diffeomorphism invariance of GR and the gauge symmetries of the fields Φ, and one finds a manifold $\mathcal{M}$ of one dimension higher, together with solutions g_{ab} and Φ for the field equations and a natural embedding of S into $\mathcal{M}$, such that the solution reproduces the prescribed initial data on S. It is interesting to note that the proof of well-posedness of such Cauchy problems often exploits energy estimates, which in turn are related to the DEC for a (auxiliary) stress tensor [9] Appendix III.

Moreover, in GR there is a unique maximal solution $(\mathcal{M}, g_{ab})$ up to diffeomorphism invariance, if we insist that solutions should be *globally hyperbolic*. There are several equivalent definitions of this concept, but a convenient one is that there must exist a *Cauchy surface* $\Sigma \subset \mathcal{M}$, which is a set Σ such that every inextendible timelike curve γ in $\mathcal{M}$ intersects Σ exactly once. Such a Cauchy surface is automatically an embedded continuous hypersurface. (It is even an embedded Lipschitz-continuous hypersurface, cf. [27] Ch.14 Prop.25 and its proof.) Naturally, the initial data surface S is a Cauchy surface.

Very many globally hyperbolic Lorentzian manifolds exist. Indeed, in every Lorentzian manifold $(\mathcal{M}, g_{ab})$ and for every $x \in \mathcal{M}$ we can find an open neighbourhood $U \subset \mathcal{M}$ of x, such that $(U, g_{ab}|_U)$ is globally hyperbolic (cf. [27] Ch.14 Lemma 43). In a globally hyperbolic Lorentzian manifold there also exist many Cauchy surfaces. One can show that $\mathcal{M}$ can be foliated by smooth and spacelike Cauchy surfaces [2, 3]. Indeed, there are very many ways to foliate $\mathcal{M}$ by such Cauchy surfaces and there is usually no preferred way to do this. Each such Cauchy surface can be called a "time-slice" and is a suitable set to specify initial data on.

A consequence of global hyperbolicity is that the Lorentzian manifold $(\mathcal{M}, g_{ab})$ is *time-orientable*, i.e. that there exists a timelike vector field t^a on $\mathcal{M}$ which is everywhere timelike for g_{ab}, $t^a t_a < 0$. In particular, t^a must be non-zero everywhere. At every point $x \in \mathcal{M}$ the timelike vectors in $T_x\mathcal{M}$ form two disjoint open cones and the existence of t^a shows that we can choose which of these are the future pointing vectors at every point $x \in \mathcal{M}$ in a way that depends smoothly on x. Mathematically, a *time-orientation* is an equivalence class of such timelike vector fields, where t^a and s^a are equivalent iff the vectors lie in the same cone at every point $x \in \mathcal{M}$ (i.e. $t^a s_a < 0$). We call the Lorentzian manifold *time-oriented* when a time-orientation has been chosen and we sometimes denote the equivalence class by $\mathfrak{t}$. For any set $A \subset \mathcal{M}$ we can then denote by $J^+(A)$, resp. $J^-(A)$, the causal future, resp. past, of A, i.e. the set of points that can be reached with a (piecewise) C^1 curve from A whose tangent vector is always a future, resp. past pointing causal vector.

Note that the description of time-orientability parallels that of orientability. $\mathcal{M}$ is orientable if there exists a *volume form* $\omega_{a_1 \cdots a_d}$, i.e. a nowhere vanishing differential form of maximal degree d (where we recall that d is the dimension of $\mathcal{M}$). $\omega_{a_1 \cdots a_d}$ is a section of a line bundle over $\mathcal{M}$ and picks out one half of that bundle at every point x in a way that depends smoothly on x. An orientation is an equivalence class of such volume forms, where two volume forms are equivalent if they are positive multiples of each other at every point. We call $\mathcal{M}$ oriented, when an orientation has been chosen and we sometimes denote the equivalence class by $\mathfrak{o}$. We have assumed from the outset that $\mathcal{M}$ is oriented.

1.3.2 Locality and general covariance

At this point we will change our perspective in two ways. Firstly, we will assume from now on that the Lorentzian metric g_{ab} is fixed, rather than a dynamical variable. This is a reasonable approximation when variations in the metric are negligible compared to the variations of the other fields. This is also the approximation that we will make later on, when we will consider QFTs in fixed background metrics. It is important to note that many hyperbolic field equations, especially linear ones, still have a well-posed

Cauchy problem in a fixed globally hyperbolic Lorentzian manifold $(\mathcal{M}, g_{ab})$, where we can prescribe initial data on any smooth, spacelike Cauchy surface.

Secondly, we will introduce a systematic way to handle the fact that we do not know in advance exactly what manifold $\mathcal{M}$ or what Lorentzian metric g_{ab} is the right one to describe, say, the universe that we live in. Although it seems clear that the dimension is four, we are less certain about, e.g., the topology of $\mathcal{M}$. Unfortunately cosmological observations only provide limited evidence on such questions, but, on the positive side, we do not expect the global shape or topology of the universe to play any rôle in the description of local physical processes taking place, e.g., on Earth or in our solar system. It shouldn't matter which spacetime region we choose to describe a system, as long as it contains all the times and places of interest and everything that can influence those.

The differential geometric formalism of GR already allows us to describe physics without committing to a particular choice of coordinates, treating all choices on an equal footing. (After all, physical processes take place regardless of the coordinates that we use to describe them.) In a completely analogous way we now want to treat all globally hyperbolic Lorentzian manifolds on an equal footing. A convenient way of doing so invokes some basic category theory. We refer to [25] (and also [39]) for further details.

Recall that a *category* C consists of a class of *objects*, $\mathsf{Obj}(\mathsf{C})$ and a class $\mathsf{Mor}(\mathsf{C})$ of *morphisms*, which can be depicted as arrows between objects. Each morphism f has a *domain* $\mathsf{dom}(f) \in \mathsf{Obj}(\mathsf{C})$ and a *codomain* $\mathsf{cod}(f) \in \mathsf{Obj}(\mathsf{C})$ and we write $\mathsf{Hom}(a, b)$ for the class of all morphisms with domain a and codomain b. Furthermore, for each object $a \in \mathsf{Obj}(\mathsf{C})$ there is an *identity morphism* $\mathsf{id}_a \in \mathsf{Hom}(a, a)$ and any two morphisms $f \in \mathsf{Hom}(a, b)$ and $g \in \mathsf{Hom}(b, c)$, where $\mathsf{cod}(f) = \mathsf{dom}(g)$, have a *composition* $g \circ f \in \mathsf{Hom}(a, c)$. The objects and morphisms of a category are required to satisfy the following two assumptions for any $f \in \mathsf{Hom}(a, b)$, $g \in \mathsf{Hom}(b, c)$ and $h \in \mathsf{Hom}(c, d)$:

$$h \circ (g \circ f) = (h \circ g) \circ f$$
$$f \circ \mathsf{id}_a = f = \mathsf{id}_b \circ f\,.$$

We will make use of the following category of localisation regions:

Definition 1.3.1. Fix $d \in \mathbb{N}$, $d \geq 2$. The category Loc of localisation regions has objects $M = (\mathcal{M}, g_{ab}, \mathfrak{o}, \mathfrak{t})$, which are all smooth connected manifolds $\mathcal{M}$ of dimension d, with a smooth Lorentzian metric g_{ab}, which makes the

Lorentzian manifold globally hyperbolic, and with an orientation $\mathfrak{o}$ and a time-orientation $\mathfrak{t}$. The morphisms $\psi : M_1 \to M_2$ in Loc are all smooth embeddings $\psi : \mathcal{M}_1 \to \mathcal{M}_2$ which preserve the metric, $\psi^*(g_2) = g_1$, the orientation, $\psi^*(\mathfrak{o}_2) = \mathfrak{o}_1$, the time-orientation, $\psi^*(\mathfrak{t}_2) = \mathfrak{t}_1$[3] and also the causal structure, in the sense that for any causal curve γ in M_2 between two points $\psi(x), \psi(y)$ in $\psi(\mathcal{M}_1)$, the entire curve γ must already lie in $\psi(\mathcal{M}_1)$.

Preservation of the causal structure means, roughly speaking, that $\mathcal{M}_2$ does not contain any new causal pathways for an event $\psi(x)$ to influence the event $\psi(y)$ that were not already included in $\mathcal{M}_1$ (or its diffeomorphic immage under ψ).

To describe physical theories we will use *functors*. A *covariant functor* $F : \mathsf{C} \to \mathsf{D}$ between two categories consists of two mappings, $F : \mathsf{Obj}(\mathsf{C}) \to \mathsf{Obj}(\mathsf{D})$ and $F : \mathsf{Mor}(\mathsf{C}) \to \mathsf{Mor}(\mathsf{D})$, which we both denote by the same name F, such that a morphism $f : a \to b$ in C gets mapped to a morphism $F(f) : F(a) \to F(b)$ with $F(\mathsf{id}_a) = \mathsf{id}_{F(a)}$ for all $a \in \mathsf{Obj}(\mathsf{C})$ and $F(f \circ g) = F(f) \circ F(g)$ for all morphisms $f, g \in \mathsf{Mor}(\mathsf{C})$ whose composition is defined.

A *contravariant functor* $F : \mathsf{C} \to \mathsf{D}$ is defined analogously, except that it reverses the direction of morphisms. It consists of the mappings $F : \mathsf{Obj}(\mathsf{C}) \to \mathsf{Obj}(\mathsf{D})$ and $F : \mathsf{Mor}(\mathsf{D}) \to \mathsf{Mor}(\mathsf{C})$, such that $f : a \to b$ gets mapped to $F(f) : F(b) \to F(a)$ and we have $F(\mathsf{id}_a) = \mathsf{id}_{F(a)}$ and $F(f \circ g) = F(g) \circ F(f)$.

We will use functors in the following way. Suppose that C is some category, whose objects we use to describe physical systems. A local and covariant physical theory should provide us with a functor $F : \mathsf{Loc} \to \mathsf{C}$. In particular, it associates an object $F(M) \in \mathsf{C}$ for every globally hyperbolic Lorentzian manifold $M \in \mathsf{Obj}(\mathsf{Loc})$, where we think of $F(M)$ as describing the system in region M. The fact that F is a functor now expresses the locality and covariance of the theory. Indeed, every morphism $\psi : M_1 \to M_2$ in Loc can be written as a composition of an isomorphism $\psi' : M_1 \to M_1'$ with $\mathcal{M}_1' := \psi(\mathcal{M}_1)$ and a canonical inclusion $\iota : M_1' \to M_2$, which is a morphism in Loc such that $\iota(x) = x$ on the subset $\mathcal{M}_1 \subset \mathcal{M}_2$. Covariance arises, because any isomorphism ψ' in Loc gives rise to an isomorphism $F(\psi)$ between the objects in C that describe the physical systems. Locality arises from the morphisms $F(\iota)$ corresponding to inclusions in Loc, which show that the physical system $F(M_1)$ is a subsystem of $F(M_2)$.

[3] Here the pull-back of a timelike vector field can be defined, because ψ is a diffeomorphism onto its range, so we can push-forward by ψ^{-1}.

Using the ideas of category theory one can consider e.g. natural transformations and isomorphisms between functors to compare different locally covariant theories, or to replace categories with equivalent ones. One can also consider limits and more advanced constructions, but we will not pursue these ideas here (see e.g. [39, 14] and references therein).

Remark 1.3.2. (a) All globally hyperbolic Lorentzian manifolds of dimension d can be embedded in a Minkowski space $\mathbb{M}^{N(d)}$ of suitably large dimension $N(d)$, see [26]. It therefore suffices to consider only those embedded globally hyperbolic manifolds. More precisely, the category Loc is equivalent to the full subcategory Loc' whose objects are embedded hypersurfaces in $\mathbb{M}^{N(d)}$. The latter is a small category in the sense that $\mathsf{Obj}(\mathsf{Loc}')$ is a set, rather than a more general class, cf. [39] Sec.2.2.

(b) For any object $M \in \mathsf{Obj}(\mathsf{Loc})$ any smooth isometric diffeomorphism of M defines a morphism in Loc. In this way, Loc contains the symmetry group of M consisting of all isometries. The use of the category Loc, rather than groups of isometries, is a quite useful generalisation, because it expresses both locality and general covariance, which are in some sense the symmetries of the theories we consider. It can therefore be argued that the term "spacetime" should apply to the category Loc, rather than its individual objects [33]. For further interesting comments on the concept of general covariance we refer to [17].

(c) One might wonder if the assumption of global hyperbolicity for objects in Loc is too restrictive. Apart from the fact that Loc already has many physically useful objects, one usually encounters difficulties when attempting to generalise Loc. Moreover, in the context of quantum gravity, attempts to enlarge the category Loc are beside the point, because the assumption that there is a smooth manifold underlying our physical theories is unjustified: we cannot localise events with perfect precision. Instead, spacetime is only expected to emerge as an effective description at low energy scales. In the language of category theory this could perhaps be expressed as follows: A physical theory is a category C of systems with their subsystem relation. It is a local and covariant theory if there exists a functor $F : \mathsf{Loc} \to \mathsf{C}$. It admits an emergent spacetimes if there are parameter-dependent families Loc_λ and functors $F_\lambda : \mathsf{Loc}_\lambda \to \mathsf{C}$, where λ denotes the energy scale and Loc_λ some kind of "quantum spacetime", such that Loc_λ approaches Loc as $\lambda \to 0^+$ in an appropriate sense.

Chapter 2

Quantum field theory

2.1 Algebraic quantum theory

A general physical system can be described in terms of the states that it can be in and the observables that we can measure on it. In classical physics, the set of pure states forms the phase space and the observables can be described by real-valued functions on this phase space. In quantum physics, the observables are given by self-djoint elements in a suitable complex *-algebra that characterises the system. In this section we will review this algebraic point of view and see how it applies to QFTs. Following [6] we will introduce locally covariant QFTs (LCQFTs), which generalise the algebraic QFTs (AQFTs) of Haag-Kastler [18] to curved spacetimes.

2.1.1 Algebras and states

Let $\mathcal{A}$ be an associative complex algebra, i.e. $\mathcal{A}$ is a complex vector space with a bilinear product map $m : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$, written as $m(A, B) = A \cdot B$ (or shorter: AB), such that $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ for all $A, B, C \in \mathcal{A}$. A *-algebra is an associative complex algebra $\mathcal{A}$ with a $\star$-operation, i.e., a complex antilinear map $^\star : \mathcal{A} \to \mathcal{A}$, which is an involution, $(A^\star)^\star = A$ for all $A \in \mathcal{A}$, and such that $(A \cdot B)^\star = B^\star \cdot A^\star$. Unless stated otherwise, we will assume that a *-algebra $\mathcal{A}$ has a unit I, i.e. an element $I \in \mathcal{A}$ such that $I \cdot A = A = A \cdot I$ for all $A \in \mathcal{A}$. This unit is always unique and in particular we have $I^\star = I$. If $\mathcal{A}$ does not have a unit, one can show that one can add a unit by enlarging $\mathcal{A}$.

Example 2.1.1. (i) An important example of *-algebra is $\mathcal{B}(\mathcal{H})$, the algebra of all bounded linear operators on a Hilbert space $\mathcal{H}$. Here $A^\star = A^*$ is the adjoint operator.

(ii) If we also want to consider unbounded operators we can choose a dense linear subspace $\mathcal{D} \subset \mathcal{H}$ and consider $\mathcal{L}(\mathcal{H}, \mathcal{D})$, the algebra of linear maps $A : \mathcal{D} \to \mathcal{D}$ such that the domain of A^* again contains $\mathcal{D}$. The *-operation on $\mathcal{L}(\mathcal{H}, \mathcal{D})$ is then defined by $A^\star = A^*|_{\mathcal{D}}$ and the unit is $I|_{\mathcal{D}}$. Note that $\mathcal{L}(\mathcal{H}, \mathcal{H}) = \mathcal{B}(\mathcal{H})$, where the boundedness op the operators follows from the Hellinger-Toeplitz theorem, which is a consequence of the closed graph theorem, cf. [29] Sec.III.5.

A *-homomorphism between *-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ is a complex linear map $\alpha : \mathcal{A}_1 \to \mathcal{A}_2$ that preserves products and the *-operation. Unless stated otherwise, we will also assume that α preserves the unit, $\alpha(I_1) = I_2$. If α is bijective, then its inverse is also a *-homomorphism and the two algebras are isomorphic. By a *-representation of a *-algebra $\mathcal{A}$ we will mean a *-homomorphism $\pi : \mathcal{A} \to \mathcal{L}(\mathcal{H}, \mathcal{D})$ for some Hilbert space $\mathcal{H}$ and some dense domain $\mathcal{D}$.

A state on $\mathcal{A}$ is a linear functional $\omega : \mathcal{A} \to \mathbb{C}$ which is normalised, $\omega(I) = 1$, and positive, $\omega(A^\star A) \geq 0$ for all $A \in \mathcal{A}$. This generalises the expectation value $\langle v, Av \rangle$ for operators $A \in \mathcal{B}(\mathcal{H})$ for a unit vector $v \in \mathcal{H}$, which is commonly encountered in quantum mechanics to describe pure states. It also generalises the notion of mixed states, given by $\mathrm{tr}(\rho \cdot A)$ for all $A \in \mathcal{B}(\mathcal{H})$, where ρ is a density matrix (or better: density operator), i.e. $\rho \geq 0$ is a positive trace class operator with $\mathrm{tr}(\rho) = 1$. (Here positivity of ρ means $\langle v, \rho v \rangle \geq 0$ for all $v \in \mathcal{H}$.) Diagonalising ρ we have $\mathrm{tr}(\rho \cdot A) = \sum_{j \in \mathbb{N}} p_j \langle v_j, Av_j \rangle$ for unit vectors $v_j \in \mathcal{H}$ and numbers $p_j \in [0, 1]$ such that $\sum_{j \in \mathbb{N}} p_j = 1$. One can think of ρ as describing a situation where the system is in the state given by v_j with probability p_j.

The following result and its constructive proof allow us to express abstract *-algebras and states in terms of the more concrete language of Hilbert spaces.

Theorem 2.1.2 (GNS-construction)**.** Given a state ω on a *-algebra $\mathcal{A}$, there exists a GNS-qudrauple $(\mathcal{H}_\omega, \Omega_\omega, \mathcal{D}_\omega, \pi_\omega)$, where $\mathcal{H}_\omega$ is a complex Hilbert space, Ω_ω a unit vector in $\mathcal{H}_\omega$, $\mathcal{D}_\omega \subset \mathcal{H}_\omega$ a dense linear subspace and $\pi_\omega : \mathcal{A} \to \mathcal{L}(\mathcal{H}_\omega, \mathcal{D}_\omega)$ a *-representation such that $\mathcal{D}_\omega = \pi_\omega(\mathcal{A})\Omega_\omega$ and

$$\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle_{\mathcal{H}_\omega}$$

for all $A \in \mathcal{A}$. The GNS-quadruple is unique up to unitary equivalence, i.e., for any other quadruple $(\mathcal{H}'_\omega, \Omega'_\omega, \mathcal{D}'_\omega, \pi'_\omega)$ with the stated properties there is a unitary isomorphism $U : \mathcal{H}_\omega \to \mathcal{H}'_\omega$ such that $\Omega'_\omega = U\Omega_\omega$, $\mathcal{D}'_\omega = U\mathcal{D}_\omega$ and $\pi'_\omega(A) = U\pi_\omega(A)U^*$ for all $A \in \mathcal{A}$.

Proof: The construction works as follows. ω defines a semi-definite inner product on $\mathcal{A}$ by $(A, B) \mapsto \omega(A^\star B)$. Let $\mathcal{N} = \{A \in \mathcal{A} \mid \omega(A^\star A) = 0\}$ be the set of vectors in $\mathcal{A}$ with 0 "norm". Note that $\mathcal{N} = A\{A \in \mathcal{A} \mid \omega(B^\star A) = 0\ \forall B \in \mathcal{A}\}$ (by the Cauchy-Schwarz inequality), so N is a linear space. We let $D_\omega = \mathcal{A}/\mathcal{N}$ be the quotient space, on which we define an inner product by $([A], [B]) := \omega(A^\star B)$ and we let $\mathcal{H}_\omega$ be its Hilbert space completion. We let $\Omega_\omega = [I] \in \mathcal{D}_\omega$ and we define the representation by setting $\pi_\omega(A)[B] = [AB]$, which is well-defined, because $\omega(C^\star AB) = \omega((A^\star C)^\star B)$ entails that $A\mathcal{N} \subset \mathcal{N}$. We also have $\pi_\omega(A)^*|_{\mathcal{D}_\omega} = \pi_\omega(A^\star)$ and $\pi_\omega(A \cdot B) = \pi_\omega(A)\pi_\omega(B)$, so π_ω is a *-representation, as desired. Given any other quadruple with the stated properties one can define a linear map $U : \mathcal{D}_\omega \to \mathcal{D}'_\omega$ by setting $U\pi_\omega(A)\Omega_\omega := \pi'_\omega(A)\Omega'_\omega$ for all $A \in \mathcal{A}$. One can show that this is a well-defined isometry, which extends uniquely to the desired unitary U. □

Remark 2.1.3. (a) The fact that the state ω is represented by a vector Ω_ω in the GNS-representation does not show that ω is a pure state. Indeed, starting with a mixed state on $\mathcal{B}(\mathcal{H})$ which is given concretely by a density matrix ρ in the Hilbert space $\mathcal{H}$, the GNS-construction provides us with a different Hilbert space representation on some Hilbert space $\mathcal{H}_\omega$ and a vector Ω_ω that represents the same mixed state. There are ways to characterise pure and mixed states on abstract *-algebras, but they do not refer to vectors in a Hilbert space. (Cf. Def.3.4.5 in [22].)

(b) The Hilbert space $\mathcal{H}_\omega$ and the representation π_ω may depend on the choice of state. E.g., we cannot expect a thermal state of a QFT in Minkowski space at non-zero temperature to be representable in the GNS-Hilbert space of the vacuum. Physically, this may be attributed to an infinite difference in the amount of particles or energy.

We will denote the set of all states on a *-algebra $\mathcal{A}$ by $\mathcal{A}^{*,+,1}$, where the three superscripts indicate, respectively, that we are dealing with linear functionals ($*$) on $\mathcal{A}$ which are positive ($+$) and normalised (1). In physical applications the space $\mathcal{A}^{*,+,1}$ is often much too large, because it contains many states that are physically unacceptable due to bad behaviour, such as infinite energy densities. To avoid these, one can try to select a subset $\mathcal{S} \subset \mathcal{A}^{*,+,1}$ to act as the state space of a physical system described by $\mathcal{A}$. Not every subset is equally suitable for this purpose, however. In particular, one would like to have the following properties.

Definition 2.1.4. We will call a set $\mathcal{S} \subset \mathcal{A}^{*,+,1}$ a *well-behaved state space* for a *-algebra $\mathcal{A}$ iff

1. $\mathcal{S}$ **separates** $\mathcal{A}$**:** for every $A \in \mathcal{A} \neq 0$ there is a $\omega \in \mathcal{S}$ with $\omega(A) \neq 0$,[1]
2. $\mathcal{S}$ **is convex**: $\lambda\omega_1 + (1-\lambda)\omega_2 \in \mathcal{S}$ for every $\omega_1, \omega_2 \in \mathcal{S}$ and $\lambda \in [0,1]$,
3. $\mathcal{S}$ **is preserved under operations from** $\mathcal{A}$: For every $\omega \in \mathcal{S}$ and $A \in \mathcal{A}$ with $\omega(A^\star A) = 1$ the state $\omega_A : B \mapsto \omega(A^\star BA)$ is in $\mathcal{S}$.

If (1) fails we can divide $\mathcal{A}$ by the linear subspace of unobservable operations. (2) allows for statistical mixing of states.

If $\mathcal{S}$ is a well-behaved state space for $\mathcal{A}$, then we can use the pair $(\mathcal{A}, \mathcal{S})$ to describe the quantum system and we can exploit the GNS-construction to recover the conventional Hilbert space formulation of quantum physics. Note that $\mathcal{S} = \mathcal{A}^{*,+,1}$ is always a well-behaved state space.

2.1.2 Algebras with additional properties

It is often useful to consider *-algebras $\mathcal{A}$ with additional properties. We will mention a few important cases here and refer to the literature on operator algebras for further details, e.g. [22].

Definition 2.1.5. $U \in \mathcal{A}$ is called a *unitary* iff $U^\star U = UU^\star = I$. A *$U^\star$-algebra* is a *-algebra $\mathcal{A}$ such that every element in $\mathcal{A}$ can be written as a finite linear combination of unitaries.

This definition follows Ch.2 of [12]. For every *-representation π of $\mathcal{A}$ (with $\pi(I) = I$) and every unitary $U \in \mathcal{A}$ we have $\|\pi(U)x\|^2 = \langle x, \pi(U^\star U)x\rangle = \|x\|^2$ and also $\|\pi(U^\star)x\|^2 = \|x\|^2$. Thus $\pi(U)$ is a unitary on $\mathcal{H}$ (restricted to a dense domain $\mathcal{D}$) and hence bounded. Any *-representation represents all operators of a $U^\star$-algebra by bounded operators.

[1] Because $\mathcal{S} \subset \mathcal{A}^{*,+,1}$ we also automatically have that $\mathcal{A}$ separates $\mathcal{S}$: for every $\omega_1, \omega_2 \in \mathcal{S}$ with $\omega_1 \neq \omega_2$ there is an $A \in \mathcal{A}$ such that $\omega_1(A) \neq \omega_2(A)$.

If we are interested in bounded operators, we can often introduce an operator norm already at an abstract level.

Definition 2.1.6. A $C^\star$*-algebra* is a *-algebra $\mathcal{A}$ with a norm $\|.\|$ such that (i) $\mathcal{A}$ is a Banach space, (ii) $\|A \cdot B\| \leq \|A\| \cdot \|B\|$, and (iii) $\|A^\star A\| = \|A\|^2$ for all $A, B \in \mathcal{A}$.

$C^\star$-algebras generalise a number of key properties of the algebra $\mathcal{B}(\mathcal{H})$ in Example 2.1.1(i). A key structural result in the theory of operator algebras is that every $C^\star$-algebra is isomorphic to a norm-closed *-sub-algebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$ ([22] Thm.4.5.6.). Furthermore, every $C^\star$-algebra is also a $U^\star$-algebra (see [22] Thm.4.1.7.).

In order to apply spectral calculus results to self-adjoint operators one would like *-algebras with sufficiently many orthogonal projections, i.e. operators $E \in \mathcal{A}$ such that $E = E^\star = E^2$. The following classes of $C^\star$-algebras serve this purpose.

Definition 2.1.7. A $W^\star$*-algebra* is a $C^\star$-algebra $\mathcal{A}$ such that $\mathcal{A} = \mathcal{B}'$ as Banach spaces, where $\mathcal{B}'$ is the space of continuous linear functionals on a Banach space $\mathcal{B}$. A *von Neumann algebra* $\mathcal{A}$ on a Hilbert space $\mathcal{H}$ is a *-subalgebra of $\mathcal{B}(\mathcal{H})$ which is closed in the weak operator topology.

If $\omega \in \mathcal{B}$ defines a state on $\mathcal{A}$ by $\omega(A) := A(\omega)$, then it is called a normal state. For any normal state, $\pi_\omega(\mathcal{A})$ is a von Neumann algebra on $\mathcal{H}_\omega$.

Example 2.1.8. To exemplify the various algebras we consider a simple harmonic oscillator with classical phase space $\mathbb{R}^2$, variables q, p and symplectic form $\sigma((q,p),(q',p')) = qp' - pq'$. The corresponding quantum system can be described by the *-algebra $\mathcal{A}_1$ generated by two self-adjoint elements Q, P satisfying $[Q,P] = i\hbar I$. $\mathcal{A}_1$ is not a $U^\star$-algebra.

One can describe the same system by a $U^\star$-algebra $\mathcal{A}_2$ generated by two families of unitaries $U(q), V(p)$ such that $U(q)V(p) = e^{-i\hbar qp}V(p)U(q)$. Formally, $U(q) = e^{iqQ}$ and $V(p) = e^{ipP}$.

By the Stone-von Neumann theorem there is a unique irreducible regular representation π of $\mathcal{A}_2$, up to unitary equivalence ([4] Cor.5.2.15). We can define a norm on $\mathcal{A}_2$ by setting $\|A\| := \|\pi(A)\|$. Completing $\mathcal{A}_2$ in this norm yields a $C^\star$-algebra $\mathcal{A}_3$. Extending π by continuity to $\mathcal{A}_3$ we find an isomorphism $\mathcal{A}_3 \to \mathcal{B}(\mathcal{H})$, where $\mathcal{B}(\mathcal{H})$ is a von Neumann algebra. The fact that we do not get more interesting von Neumann algebras is related to the finite dimension of the classical phase space in this example.

2.2 Lagrangian QFTs

In this section we will define an algebra that describes the observables of a real scalar Lagrangian QFT. We follow the basic framework of [7], extended to an arbitrary globally hyperbolic Lorentzian manifold $M = (\mathcal{M}, g_{ab})$. Examples with states and representations will be discussed in later sections.

A classical real scalar field has configuration space $\mathcal{E} = C^\infty(\mathcal{M}, \mathbb{R})$. Which configurations are allowed by the dynamics of a theory depends on the equations of motion, which, in our case, will be derived from a Lagrange density. We will consider Lagrange densities of the form

$$\mathcal{L}[g', \phi] = \left(\frac{1}{2} (g')^{ab} \nabla_a \phi \cdot \nabla_b \phi + \frac{1}{2} m^2 \phi^2 + \frac{1}{2} \xi R' \phi^2 + \sum_{n=0}^{N} c_n(x) \phi(x)^n \right) \mathrm{d}vol_{g'} , \tag{2.1}$$

where $N \in \mathbb{N}$ and $c_n \in C^\infty(\mathcal{M}, \mathbb{R})$ are arbitrary and we assume that the difference $\delta g_{ab} := g'_{ab} - g_{ab}$ of the Lorentzian metrics is compactly supported and smooth. These Lagrange densities are a special case of $\mathcal{L}_{\mathrm{sf}}[g', \phi]$ as given in (1.7) with potential function $V(\phi) = \sum_{n=0}^{N} c_n(x) \phi(x)^n$. We denote the corresponding action functionals by $S(f)[g', \phi]$ as in Section 1.1.

Note that $\mathrm{d}vol_{g'}(x) = \mu_{g'} \mathrm{d}vol_g(x)$ for some $\mu_{g'} \in C_0^\infty(\mathcal{M}, \mathbb{R}_{>0})$ with $\mu_{g'} - 1$ compactly supported. We can therefore express S in terms of $\mathrm{d}vol_g$ as

$$S(f)[g', \phi] = \int_{\mathcal{M}} f \mu_{g'} \left(\frac{1}{2} (g')^{ab} \partial_a \phi \partial_b \phi + \frac{1}{2} m^2 \phi^2 + \frac{1}{2} \xi R \phi^2 + V(\phi) \right) \mathrm{d}vol_g .$$

The algebra that we will define is meant to encode the response of a QFT to compactly supported variations of the Lagrange density $\mathcal{L}[g, \phi]$ (with the reference metric g_{ab}) within the class (2.1). We are therefore interested in differences $\mathcal{L}'[g', \phi] - \mathcal{L}[g, \phi]$, where the coefficients vary only in a compact set. Because $(\mu_{g'} - 1) m^2$ and $\mu_{g'} R' - R$ are in $C_0^\infty(\mathcal{M}, \mathbb{R})$ (with R the scalar curvature of the metric g_{ab} and R' that of g'_{ab}), we can write such a difference $\lambda[\phi] = \mathcal{L}'[g', \phi] - \mathcal{L}[g, \phi]$ in the form

$$\lambda[\phi] = \left(\frac{1}{2} \gamma^{ab} \partial_a \phi \cdot \partial_b \phi + \sum_{n=0}^{N} c_n(x) \phi(x)^n \right) \mathrm{d}vol_g \tag{2.2}$$

for some (new) coefficient functions $c_n \in C_0^\infty(\mathcal{M}, \mathbb{R})$ and with

$$\gamma^{ab} = \mu_{g'} (g')^{ab} - g^{ab} \tag{2.3}$$

a compactly supported, smooth and symmetric tensor field. Correspondingly we define the functionals $F[\phi] := \int_{\mathcal{M}} \lambda$, i.e.

$$F[\phi] = c + \int_{\mathcal{M}} \left(\frac{1}{2}\gamma^{ab}(x)\partial_a\phi(x)\partial_b\phi(x) + \sum_{n=1}^{N} c_n(x)\phi(x)^n \right) \mathrm{d}vol_g(x) \,. \quad (2.4)$$

Note that adding an $n = 0$ term to the sum amounts to changing the constant c by $\int_{\mathcal{M}} c_0(x)\mathrm{d}vol_g(x)$. The advantage of excluding $n = 0$ from the sum is, that the coefficient c, the tensor γ^{ab} and the functions c_n for $n \geq 1$ are then uniquely determined by F, as one can show using the variational lemma.

When $d = 2$, all Lorentzian metrics are locally conformally equivalent, i.e. we locally have $g'_{ab} = \nu^2 g_{ab}$ for some smooth $\nu > 0$ and hence $\mu_{g'} = \nu^2$, $(g')^{ab} = \nu^{-2}g^{ab}$ and $\gamma^{ab} = 0$. This means that perturbations of the kinetic term of $\mathcal{L}[g,\phi]$ are not described by any local degrees of freedom. To avoid this issue we will assume $d > 2$ from now on.

When $d > 2$ and $\lambda = \mathcal{L}'[g',\phi] - \mathcal{L}[g,\phi]$, then we can recover g'_{ab} from g_{ab} and the formula (2.3) for γ^{ab} by setting

$$g'_{ab} = \mu_{(g^{-1}+\gamma)^{-1}}^{\frac{-2}{d-2}} (g^{-1} + \gamma)^{-1}_{ab} \,,$$

where $(g^{-1} + \gamma)^{-1}_{ab}$ denotes the Lorentzian metric whose inverse is $g^{ab} + \gamma^{ab}$. Note that a functional F of the form (2.4) arises from a perturbation of the Lagrange density $\mathcal{L}$ only if $g^{ab} + \gamma^{ab}$ defines an inverse Lorentzian metric, which is not true for all test tensor fields γ^{ab}.

Definition 2.2.1. We will write $\mathcal{F}(M,\mathcal{L})$, or $\mathcal{F}$, for the set of all functionals F of the form (2.4) with the property that $g^{ab} + \gamma^{ab}$ defines an inverse Lorentzian metric making $\mathcal{M}$ globally hyperbolic. We will write $(g_F)_{ab}$ for the Lorentzian metric determined by F in this way, i.e $g_F^{ab} = g^{ab} + \gamma^{ab}$.

$\mathcal{F}(M,\mathcal{L})$ is not a vector space, but it contains the vector space of elements with $\gamma^{ab} = 0$. Every $F \in \mathcal{F}(M,\mathcal{L})$ determines a unique γ^{ab}, so $(g_F)_{ab}$ is well-defined and it equals g_{ab} outside a compact set. By assumption, $(\mathcal{M}, g_F)$ is again globally hyperbolic.

When we modify S by adding a term $F \in \mathcal{F}(M,\mathcal{L})$, then we modify $\mathcal{L}$ by adding a term λ of the form (2.2). We consider how this changes the theory to the future of F. For this purpose we define the spacetime support of F by

$$\mathrm{supp}(F) := \mathrm{supp}(\gamma^{ab}) \cup \bigcup_{n=1}^{N} \mathrm{supp}(c_n) \,.$$

In particular, when $N = 0$ and $\gamma^{ab} = 0$, then the functional $F[\phi] = c$ is constant and we set $\mathrm{supp}(F) = \emptyset$.

Remark 2.2.2. With the same motivation one could also define the support of densities λ of the form (2.2) by the same formula. Note, however, that the support of the coefficient function c_0 does not contribute to the support of λ, essentially because it does not change the equations of motion.

Definition 2.2.3. We define the *relative action* for all $\phi_0 \in C_0^\infty(\mathcal{M}, \mathbb{R})$ by

$$\delta S(\phi_0)[g, \phi] := S(f)[g, \phi + \phi_0] - S(f)[g, \phi]\,,$$

where $f \in C_0^\infty(M)$ with $f \equiv 1$ on $\mathrm{supp}(\phi_0)$.

Note that δS only depends on the perturbation ϕ_0 of the field configuration, not on f, as long as $f \equiv 1$ on $\mathrm{supp}(\phi_0)$. Moreover, $\delta S(\phi_0)$ is a functional in $\mathcal{F}(M, \mathcal{L})$ with $g_{\delta S} = g$, because the quadratic terms in $\partial_a\phi$ stemming from the kinetic term of $\mathcal{L}$ cancel out and the linear terms in $\partial_a\phi$ can be rewritten using an integration by parts. By a similar argument, if $F \in \mathcal{F}(M, \mathcal{L})$, then the functional $\phi \mapsto F[\phi + \phi_0]$ is also in $\mathcal{F}(M, \mathcal{L})$ with the same γ^{ab}.

Definition 2.2.4. Given the Lagrange density $\mathcal{L}$ on M, the *dynamical algebra* $\mathcal{A}_M^{(\mathcal{L})}$ is generated algebraically by an identity I and operators $U(F)$ and their adjoints $U(F)^*$ for each $F \in \mathcal{F}(M, \mathcal{L})$, subject to the following relations:

(1) **unitarity:** $U(F)^* = U(F)^{-1}$ for all $F \in \mathcal{F}(M, \mathcal{L})$,

(2) **normalisation:** $U(c) = e^{ic}I$ for constant functionals $c[\phi] = c$, $c \in \mathbb{R}$,

(3) **dynamics:** for each $F \in \mathcal{F}(M, \mathcal{L})$ and $\phi_0 \in C_0^\infty(\mathcal{M}, \mathbb{R})$,

$$U(F) = U(F[. + \phi_0] + \delta S(\phi_0))\,,$$

where $F[. + \phi_0]$ is the functional in $\mathcal{F}(M, \mathcal{L})$ defined by $\phi \mapsto F[\phi + \phi_0]$,[2]

[2]Alternatively there is an off-shell formalism, where the equations of motion are not implemented in this strict sense, but one requires $U(F)U(\delta S(\phi_0)) = U(F[. + \phi_0] + \delta S(\phi_0)) = U(\delta S(\phi_0))U(F)$ instead. In this case the operators $U(\delta S(\phi_0))$ are in the centre of the algebra. For clarity we sometimes call $\mathcal{A}^{(\mathcal{L})}$ the on-shell dynamical algebra.

(4) **causal factorisation:** for all $F_1, F_2, F_3 \in \mathcal{F}(M, \mathcal{L})$ such that $F_1 + F_2 + F_3, F_1 + F_3, F_2 + F_3 \in \mathcal{F}(M, \mathcal{L})$ and such that $\mathrm{supp}(F_1)$ lies to the future of $\mathrm{supp}(F_2)$ in $(\mathcal{M}, g_{F_3})$, i.e. $\mathrm{supp}(F_1)$ does not intersect[3] $J^-_{(\mathcal{M}, g_{F_3})}(\mathrm{supp}(F_2))$,

$$U(F_1 + F_3 + F_2) = U(F_1 + F_3)U(F_3)^{-1}U(F_3 + F_2)\,.$$

A few comments on this definition are in order.

In quantum physics we can think of unitary operators as operations acting on the system. Relation (1) in Definition 2.2.4 posits that a change of $\mathcal{L}$ is an operation that changes the dynamics of the system and should be described by a unitary operator. A general element $A \in \mathcal{A}_M^{(\mathcal{L})}$ is of the form

$$A = z_0 + \sum_{n=1}^{N} z_n U(F_1^{(n)})^{s_1^{(n)}} \cdots U(F_N^{(n)})^{s_N^{(n)}}$$

for some $N \in \mathbb{N}$, $z_n \in \mathbb{C}$, $F_j^{(n)} \in \mathcal{F}(M, \mathcal{L})$ and $s_j^{(n)} \in \{\pm 1\}$. Here, each term is a multiple of a unitary, so $\mathcal{A}_M^{(\mathcal{L})}$ is a $U^\star$-algebra. By a standard construction one may define a norm on $\mathcal{A}_M^{(\mathcal{L})}$, namely $\|A\| := \sup_{\omega \in \left(\mathcal{A}_M^{(\mathcal{L})}\right)^{*,+,1}} \|\pi_\omega(A)\|_{\mathcal{H}_\omega}$. Taking the completion in this norm yields a $C^\star$-algebra [7].

Relation (2) is a kind of normalisation. Even without (2) one can derive from (4) that the map $c \mapsto U(c)$ is a group homomorphism from $(\mathbb{R}, +)$ to the group of unitaries in the centre of $\mathcal{A}_M^{(\mathcal{L})}$ (i.e. commuting with all $A \in \mathcal{A}_M^{(\mathcal{L})}$).

In relation (3) we first recall that the functional $F[. + \phi_0]$ is indeed in $\mathcal{F}(M, \mathcal{L})$ and that the sum of the functionals is always well-defined, because δS does not change the reference metric. This dynamical axiom is an integrated form of the Schwinger-Dyson equation known from perturbative QFT. We will see in examples that relation (3) fixes the dynamics of $U(F)$ for F at most linear in ϕ, but it does not fix the dynamics completely, see the discussion of the time-slice axiom in Section 2.4.1 or [5].

The causal factorisation in Relation (4) refers to the time ordering of F_1 and F_2 w.r.t. the deformed metric g_{F_3}. When $F_3 = 0$, it simply states that $U(F_1 + F_2) = U(F_1)U(F_2)$ when the support of F_1 is later than that of F_2 in the reference metric g. In particular, if these supports are spacelike

[3] Equivalently one may require that there exists a smooth spacelike Cauchy surface Σ in $(\mathcal{M}, g_{F_3})$ such that $\mathrm{supp}(F_1)$ lies to the future (w.r.t. g_{F_3}) of Σ and $\mathrm{supp}(F_2)$ to its past.

separated, then $U(F_1)$ and $U(F_2)$ commute. More generally the equality $U(F_1+F_2) = U(F_1)U(F_2)$ shows that $U(F_1+F_2)$ behaves like a time ordered exponential that is often encountered in perturbative QFT. In fact, relation (4) is motivated by identities that were first developed in the perturbative setting.

For all $F, H \in \mathcal{F}(M, \mathcal{L})$ such that $F + H \in \mathcal{F}(M, \mathcal{L})$ we define

$$U_H(F) := U(H)^{-1}U(H + F)\,. \tag{2.5}$$

Then relation (4) is equivalent to the validity of the time ordering formula $U_{F_3}(F_1 + F_2) = U_{F_3}(F_1)U_{F_3}(F_2)$ in $(\mathcal{M}, g_{F_3})$ for all F_1, F_2, F_3 for which the sums are defined.

Relation (4) almost follows from the time ordering formula for U itself. If Σ is a Cauchy surface in $(\mathcal{M}, g_{F_3})$ that separates the supports of F_1 and F_2 and if we can write $F3 = F_+ + F_-$ with $\mathrm{supp}(F_\pm)$ to the future, resp. past of Σ, then the time ordering formula for U gives

$$\begin{aligned} U(F_1 + F_3 + F_2) &= U(F_1 + F_+)U(F_- + F_2) \\ &= U(F_1 + F_+)U(F_-)(U(F_+)U(F_-))^{-1}U(F_+)U(F_- + F_2) \\ &= U(F_1 + F_3)U(F_3)^{-1}U(F_3 + F_2)\,. \end{aligned}$$

In general, however, the splitting of F_3 destroys the smoothness of the coefficients.

If $\mathcal{L}'[g, \phi]$ is a Lagrange density with the same reference metric, then $\mathcal{F}(M, \mathcal{L}') = \mathcal{F}(M, \mathcal{L})$. Setting $H[\phi] := \int_{\mathcal{M}} \mathcal{L}'[g, \phi] - \mathcal{L}[g, \phi]$ we have $H \in \mathcal{F}(M, \mathcal{L})$ and the operators $U_H(F)$ of (2.5) are well-defined for all $F \in \mathcal{F}(M, \mathcal{L}')$. It is easy to verify that these operators satisfy the defining relations for the algebra $\mathcal{A}_M^{(\mathcal{L}')}$. We can recover $U(F) = U_H(-H)^{-1}U_H(-H+F)$, so the algebras $\mathcal{A}_M^{(\mathcal{L}')}$ and $\mathcal{A}_M^{(\mathcal{L})}$ are isomorphic.

This isomorphism fails when the difference of the Lagrange densities is not compactly supported, or when the kinetic term is modified, because $\mathcal{F}(M, \mathcal{L}') \neq \mathcal{F}(M, \mathcal{L})$ in those cases. However, if the change in the Lagrange density has coefficients that are not compactly supported, then the argument above can be applied at least locally by using suitable cut-off functions [7].

2.3 Toy model: the free scalar field

As an example we consider a free scalar field given by the Lagrange density

$$\mathcal{L}_{\mathrm{fsf}}[g,\phi] := \frac{1}{2}\left(g^{ab}\partial_a\phi\cdot\partial_b\phi + m^2\phi^2 + \xi R\phi^2\right)\mathrm{d}vol_g\,.$$

Before we study the quantum theory, we first review the classical theory.

2.3.1 The classical scalar field

For any $\phi_0, f \in C_0^\infty(\mathcal{M},\mathbb{R})$ with $f \equiv 1$ on $\mathrm{supp}(\phi_0)$ we compute the relative action

$$\begin{aligned}\delta S_{\mathrm{fsf}}(\phi_0)[\phi] &= S_{\mathrm{fsf}}(f)[\phi+\phi_0] - S_{\mathrm{fsf}}(f)[\phi]\\ &= \int_{\mathcal{M}} f(\mathcal{L}_{\mathrm{fsf}}[g,\phi+\phi_0] - \mathcal{L}_{\mathrm{fsf}}[g,\phi])\\ &= \int_{\mathcal{M}} f(g^{ab}\partial_a\phi_0\cdot\partial_b\phi + m^2\phi_0\phi + \xi R\phi_0\phi)\;\mathrm{d}vol_g + \int_{\mathcal{M}} f\mathcal{L}_{\mathrm{fsf}}[\phi_0]\\ &= \int_{\mathcal{M}} \phi_0 P\phi\;\mathrm{d}vol_g + \int_{\mathcal{M}} \mathcal{L}_{\mathrm{fsf}}[\phi_0]\,, \end{aligned} \tag{2.6}$$

where $P = -\Box + m^2 + \xi R$. The Euler-Lagrange equation is therefore the modified Klein-Gordon equation (1.8) with $V' \equiv 0$.

It is well-known that the Klein-Gordon equation has a well-posed initial value formulation on any globally hyperbolic Lorentzian manifold $M = (\mathcal{M}, g)$, cf. Ch.3 of [1]. I.e., for any smooth spacelike Cauchy surface Σ and any functions $f_0, f_1 \in C^\infty(\Sigma)$ there is a unique solution $\phi \in C^\infty(\mathcal{M},\mathbb{R})$ to $P\phi = 0$ with $\phi|_\Sigma = f_0$ and $\nabla_n\phi|_\Sigma = f_1$, where ∇_n denotes the future pointing normal derivative to Σ. Furthermore, if the initial data are supported in $A \subset \Sigma$, then $\mathrm{supp}(\phi) \subset J^+(A)\cup J^-(A)$.

Definition 2.3.1. A closed set $A \subset M$ is called *spacelike compact*, resp. *past spacelike compact*, resp. *future spacelike compact* when $A \subset J^+(K)\cup J^-(K)$, resp. $A \subset J^+(K)$, resp. $A \subset J^-(K)$ for some compact $K \subset \mathcal{M}$.

A function is called *spacelike compact*, resp. *past spacelike compact*, resp. *future spacelike compact* iff its support is. We will denote the spaces of all smooth functions with such supports by $C^\infty_{\mathrm{sc/psc/fsc}}(\mathcal{M},\mathbb{R})$, respectively.

When the initial data are compactly supported, the solution ϕ is spacelike compact. One can show that ϕ is spacelike compact iff its intersection with every smooth, spacelike Cauchy surface in M is compact and hence its initial data on any other smooth spacelike Cauchy surface is also compact [31]. We will write $\mathcal{S}_{\mathrm{sc}}^{\infty}(\mathcal{M},\mathbb{R})$ for the space of all real-valued spacelike compact solutions ϕ to $P\phi = 0$ on $\mathcal{M}$.

Related to the well-posedness of the initial value formulation is the fact that the Klein-Gordon equation has unique advanced ($-$) and retarded ($+$) fundamental solutions $E^{\pm}$. These are linear maps $E^{-} : C_{\mathrm{fsc}}^{\infty}(\mathcal{M},\mathbb{R}) \to C_{\mathrm{fsc}}^{\infty}(\mathcal{M},\mathbb{R})$ and $E^{+} : C_{\mathrm{psc}}^{\infty}(\mathcal{M},\mathbb{R}) \to C_{\mathrm{psc}}^{\infty}(\mathcal{M},\mathbb{R})$ such that $PE^{\pm}f = f$ for all $f \in C_{\mathrm{fsc/psc}}^{\infty}(\mathcal{M},\mathbb{R})$ and $\mathrm{supp}(E^{\pm}f) \subset J^{\pm}(\mathrm{supp}(f))$. The operators $E^{\pm}$ are dual to each other under the inner product of $L^2(\mathcal{M},\mathbb{R},\mathrm{d}vol_g)$, because for any $f_{\pm} \in C_{\mathrm{psc/fsc}}^{\infty}(\mathcal{M},\mathbb{R})$

$$\langle f_-, E^+ f_+\rangle = \langle PE^- f_-, E^+ f_+\rangle = \langle E^- f_-, PE^+ f_+\rangle = \langle E^- f_-, f_+\rangle\,,$$

where the integrations by parts needed to move P to the right are allowed, because the intersection of the supports of $E^- f_-$ and $E^+ f_+$ is compact. It follows from further integrations by parts that

$$\langle f_-, E^+ P f_+\rangle = \langle E^- f_- h, P f_+\rangle = \langle PE^- f_-, f_+\rangle = \langle f_-, f_+\rangle$$

and hence

$$E^+ P f_+ = f_+$$

for all $f_+ \in C_{\mathrm{psc}}^{\infty}(\mathcal{M},\mathbb{R})$. Similarly,

$$E^- P f_- = f_-$$

for all $f_- \in C_{\mathrm{fsc}}^{\infty}(\mathcal{M},\mathbb{R})$.

The operator $E := E^+ - E^-$, defined on $C_0^{\infty}(\mathcal{M},\mathbb{R})$, is a convenient tool to generate solutions to the Klein-Gordon equation, because $PEf = PE^- f - PE^+ f = f - f = 0$ for all $f \in C_0^{\infty}(\mathcal{M},\mathbb{R})$. The following result summarises some of the basic facts of the classical theory.

Proposition 2.3.2. The following sequence is exact

$$\{0\} \to C_0^{\infty}(\mathcal{M},\mathbb{R}) \xrightarrow{P} C_0^{\infty}(\mathcal{M},\mathbb{R}) \xrightarrow{E} \mathcal{S}_{\mathrm{sc}}^{\infty}(\mathcal{M},\mathbb{R}) \to \{0\}\,. \qquad (2.7)$$

This shows in particular that all spacelike compact solutions to $P\phi = 0$ can be written as $\phi = Ef$ for some $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$.

Proof: $\ker(P) = \{0\}$, because a compactly supported f with $Pf = 0$ has vanishing initial data on some Cauchy surface to the past of its support. Also, $\mathrm{ran}(P) \subset \ker(E)$, because $EPf = E^+Pf - E^-Pf = f - f = 0$. Conversely, if $Ef = 0$, then $\chi := E^+f = E^-f$ must be compactly supported in $J^+(\mathrm{supp}(f)) \cap J^-(\mathrm{supp}(f))$ and therefore $f = P\chi$, so $\ker(E) \subset \mathrm{ran}(P)$.

To see why E is surjective we take any spacelike compact solution ϕ and two Cauchy surfaces $\Sigma_\pm$ in M such that Σ_+ lies to the future of Σ_-. We then choose $\chi \in C^\infty(\mathcal{M}, \mathbb{R})$ such that $\chi \equiv 0$ on $J^-(\Sigma_-)$ and $\chi \equiv 1$ on $J^+(\Sigma_+)$. Then $\chi\phi$ has past spacelike compact support and hence $E^+P\chi\phi = \chi\phi$. Similarly, $E^-P(\chi - 1)\phi = (\chi - 1)\phi$. Because $P(\chi - 1)\phi = P\chi\phi$ we then find $EP\chi\phi = E^+P\chi\phi - E^-P(\chi - 1)\phi = \chi\phi - (\chi - 1)\phi = \phi$. □

Remark 2.3.3. Note that the handling of the supports in the proof of Proposition 2.3.2 is a bit delicate. We do not have $E^-P\chi\phi = \chi\phi$, because the support of $\chi\phi$ is not future spacelike compact. Instead we have $E^-P\chi\phi = E^-P(\chi - 1)\phi = (\chi - 1)\phi$.

The last argument in the proof of Proposition 2.3.2 can be refined further, because we still have a lot of freedom to choose χ and, hence, to control the support of $P\chi\phi$. In particular we will use the following result.

Lemma 2.3.4. Given $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ and a compact set $K \subset \mathcal{M}$ we can find $f_\pm, \phi_\pm \in C_0^\infty(\mathcal{M}, \mathbb{R})$ such that $f = f_+ + P\phi_+ = f_- + P\phi_-$ and $\mathrm{supp}(f_\pm) \cap J^\mp(K) = \emptyset$.

Proof. In the last paragraph of the proof of Proposition 2.3.2 we can choose $\Sigma_\pm$ in such a way that Σ_- lies to the future of K. Taking $\phi = Ef$, $f_+ = P\chi\phi$ and $\phi_+ = E^+(f - f_+)$ it then follows that $\mathrm{supp}(f_+) \cap J^-(K) = \emptyset$ and $\phi = Ef_+$, so $E(f - f_+) = 0$ and $\phi_+ = E^+(f - f_+) = E^-(f - f_+)$ is compactly supported with $P\phi_+ = f - f_+$. The proof for the other sign is analogous. □

As a last thing before we consider the quantum theory we note that E defines a symplectic form on the quotient vector space $C_0^\infty(\mathcal{M}, \mathbb{R})/PC_0^\infty(\mathcal{M}, \mathbb{R})$ by $\sigma([f], [h]) = \langle f, Eh \rangle$. Equivalently this may be viewed as a symplectic form on the solution space $\mathcal{S}_{\mathrm{sc}}^\infty(\mathcal{M}, \mathbb{R})$. The latter can be expressed in terms of the initial data of such solutions on any Cauchy surface. Writing

(f_0, f_1) for the initial data of Ef on a smooth spacelike Cauchy surface Σ with $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ and similarly for h we have by Stokes' theorem

$$
\begin{aligned}
\sigma([f],[h]) &= \langle f, Eh \rangle \\
&= \sum_{\pm} \int_{J^\pm(\Sigma)} (PE^\mp f) Eh \; \mathrm{d}vol_g \\
&= \sum_{\pm} \int_{J^\pm(\Sigma)} \nabla^a (E^\pm f \nabla_a Eh - (\nabla_a E^\pm f) Eh) + E^\pm f K Eh \; \mathrm{d}vol_g \\
&= \sum_{\pm} \int_{\Sigma} \pm (E^\pm f \nabla_n Eh - (\nabla_n E^\pm f) Eh) \; \mathrm{d}vol_\Sigma \\
&= \int_{\Sigma} (Ef \nabla_n Eh - (\nabla_n Ef) Eh) \; \mathrm{d}vol_\Sigma \\
&= \int_{\Sigma} (f_0 h_1 - f_1 h_0) \; \mathrm{d}vol_\Sigma
\end{aligned}
\tag{2.8}
$$

where $\mathrm{d}vol_\Sigma$ is the volume form on Σ determined by the Riemannian metric induced by g. This equality is analogous to the canonical commutation relations $[q,p] = i\hbar$ in quantum mechanics, if we think of f_0 and h_0 as positions and f_1, h_1 as momenta. Remarkably, this formula is independent of the choice of the smooth spacelike Cauchy surface Σ.

2.3.2 The quantized free scalar field

Now we turn to the algebra $\mathcal{A}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ and we will focus on the operators $U(F)$ for the simple class of functionals $F = c + L_f$, where $c \in \mathbb{R}$ is a constant functional and

$$
L_f[\phi] := \int_{\mathcal{M}} f\phi \; \mathrm{d}vol_g
$$

with $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$. Note that $U(L_f + c) = U(L_f)U(c) = e^{ic}U(L_f)$ by normalisation and causal factorisation, because c has empty support.

When $\phi_0 \in C_0^\infty(\mathcal{M}, \mathbb{R})$, then $\delta S_{\mathrm{fsf}}(\phi_0)$ is also of the form $L_f + c$, because

$$
\delta S_{\mathrm{fsf}}(\phi_0) = L_{P\phi_0} + S_{\mathrm{fsf}}(h)[\phi_0] \tag{2.9}
$$

by equation (2.6), for any $h \in C_0^\infty(\mathcal{M}, \mathbb{R})$ with $h \equiv 1$ on $\mathrm{supp}(\phi_0)$. In $\mathcal{A}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ we have $U(\delta S_{\mathrm{fsf}}(\phi_0)) = I$ by relations (3) and (2) and hence

$$
U(L_{P\phi_0}) = e^{-iS_{\mathrm{fsf}}(h)[\phi_0]} I \qquad (h \equiv 1 \, \mathrm{on} \; \mathrm{supp}(\phi_0)) \,. \tag{2.10}
$$

Let $f, f_+, \phi_+ \in C_0^\infty(\mathcal{M}, \mathbb{R})$ such that $f = f_+ + P\phi_+$. Since $L_f = L_{f_+} + L_{P\phi_+}$ we can use (2.9) and relation (3) to compute (with $h \equiv 1$ on $\mathrm{supp}(\phi_+)$)

$$\begin{aligned} U(L_f) &= e^{-iS_{\mathrm{fsf}}(h)[\phi_+]} U(L_{f_+} + \delta S_{\mathrm{fsf}}(\phi_+)) \\ &= e^{-iS_{\mathrm{fsf}}(h)[\phi_+]} U(L_{f_+}[. - \phi_+]) \\ &= e^{-iL_{f_+}[\phi_+] - iS_{\mathrm{fsf}}(h)[\phi_+]} U(L_{f_+}) , \end{aligned} \tag{2.11}$$

because $L_{f_+}[. - \phi_+] = L_{f_+} - L_{f_+}[\phi_+]$. This allows us to move the support of f to the support of f_+, at the cost of a phase factor $e^{-iL_{f_+}[\phi_+] - iS_{\mathrm{fsf}}(h)[\phi_+]}$.

If $F \in \mathcal{F}(M, \mathcal{L}_{\mathrm{fsf}})$ is supported in the compact set K we can choose f_+ supported to the future of K. Noting that $\phi_+ = E^+ P\phi_+ = E^+ f - E^+ f_+$ and using the fact that the supports of $E^+ f_+$ and F do not intersect we then have $F[. + \phi_+] = F[. + E^+ f]$. With this identity we find

$$\begin{aligned} U(L_f)U(F) &= e^{-iL_{f_+}[\phi_+] - iS_{\mathrm{fsf}}(h)[\phi_+]} U(L_{f_+})U(F) \\ &= e^{-iL_{f_+}[\phi_+] - iS_{\mathrm{fsf}}(h)[\phi_+]} U(L_{f_+} + F) \\ &= e^{-iL_{f_+}[\phi_+] - iS_{\mathrm{fsf}}(h)[\phi_+]} U(L_{f_+}[. + \phi_+] + F[. + \phi_+] + \delta S_{\mathrm{fsf}}(\phi_+)) \\ &= U(L_f + F[. + E^+ f]) , \end{aligned} \tag{2.12}$$

using again (2.9) in the last line. Similarly, for $f = f_- + P\phi_-$ with f_- supported to the past of K,

$$\begin{aligned} U(F)U(L_f) &= U(L_f + F[. + E^- f]) , \\ U(L_f)U(F) &= U(F[. + Ef])U(L_f) . \end{aligned} \tag{2.13}$$

If $F = L_h$ with $h \in C_0^\infty(\mathcal{M}, \mathbb{R})$, then (2.12,2.13) yields

$$U(L_h)U(L_f) = e^{i\langle h, E^- f\rangle} U(L_{f+h}) = e^{-i\langle h, Ef\rangle} U(L_f)U(L_h)$$

in terms of the L^2 inner product, because $L_f + L_h[. + E^\pm f] = L_{f+h} + \langle h, E^\pm f\rangle$. If $\mathrm{supp}(h) \cap J^+(\mathrm{supp}(f)) = \emptyset$ the phase factor in the middle term vanishes. This is in line with the time ordering in $\mathcal{A}_M^{(\mathcal{L}_{\mathrm{fsf}})}$, but it is possible to get a more symmetric relation, or rather, an anti-symmetric one under reversing the time-orientation, as follows. We define the Weyl operators[4]

$$W(f) = e^{\frac{i}{2}\langle f, E^- f\rangle} U(L_f) \tag{2.14}$$

[4] A more symmetric expression for the phase factor in (2.14) can be obtained by using $\langle f, E^- f\rangle = \langle E^+ f, f\rangle = \langle f, E^+ f\rangle$ and hence $\langle f, E^- f\rangle = \langle f, E_D f\rangle$ with $E_D := \frac{1}{2}(E^+ + E^-)$ the Dirac propagator.

for all $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$. For all $f, h \in C_0^\infty(\mathcal{M}, \mathbb{R})$ we then find

$$W(h)W(f) = e^{-\frac{i}{2}\langle h, Ef\rangle} W(f+h) = e^{-i\langle h, Ef\rangle} W(f)W(h)\,, \tag{2.15}$$

using $\langle f, E^- h\rangle = \langle h, E^+ f\rangle$. Because $W(0) = I$ we find in particular that $W(-f) = W(f)^{-1} = W(f)^*$.

(2.15) are the Weyl relations, which arise quite naturally if one interprets $W(f) = e^{i\phi(f)}$ as the complex exponential of a smeared field operator, i.e. $\phi(f) = \int_{\mathcal{M}} f(x)\phi(x)\mathrm{d}vol_g(x)$, and imposes the canonical commutation relations $[\phi(f), \phi(h)] = iE(f, h)$. The Weyl operators and the Weyl relations have been known for about a century. It is nice to see that they are naturally included in the recent dynamical algebra approach.

2.4 Locally covariant QFT

In QFT we want to describe quantum physics while keeping track of the localisation regions of its observables, which are thought to be described by quantum fields. Following the ideas of [6] we therefore associate a quantum system to every localisation region in a coherent way.

We first introduce a category to describe (localised) physical systems and their embeddings.

Definition 2.4.1. The category Phys has objects $(\mathcal{A}, \mathcal{S})$, where $\mathcal{A}$ is a *-algebra and $\mathcal{S}$ a well-behaved state space for $\mathcal{A}$, cf. Definition 2.1.4. A morphism $\alpha : (\mathcal{A}_1, \mathcal{S}_1) \to (\mathcal{A}_2, \mathcal{S}_2)$ in Phys is a *-homomorphism[5] $\alpha : \mathcal{A}_1 \to \mathcal{A}_2$ with the property that the pull-back $\alpha^* : \mathcal{A}_2^{*,+,1} \to \mathcal{A}_1^{*,+,1}$, given by $\alpha^*\omega := \omega \circ \alpha$, maps $\mathcal{S}_2$ into $\mathcal{S}_1$.

A *locally covariant QFT* (LCQFT) is a covariant functor $\mathbf{P} : \mathsf{Loc} \to \mathsf{Phys}$.

[5]This definition differs from the approach of [6] in three minor ways. Firstly, [6] treats the observables and the state spaces separately. Secondly, [6] uses a category of $C^\star$-algebras (at least in first instance). Thirdly, [6] insist that the morphisms should be injective. We will drop the last assumption, because it is already violated in free electromagnetism due to Gauss' law [32]. Indeed, one can consider an observable which determines the electric flux through a surface that encloses a spacetime region O. If O is topologically non-trivial, there is no way of knowing what charge might be hidden there, so the observable is non-trivial. Embedding this into a larger region, however, we might close up the hole, fix the electric charges present and map the observable to a multiple of the identity in line with Gauss' law.

We can think of a morphism α as telling us how $\mathcal{A}_1$ is a subsystem of $\mathcal{A}_2$, and α^* tells us how states on $\mathcal{A}_2$ restrict to the subsystem $\mathcal{A}_1$. Instead of $\mathbf{P}(M)$ and $\mathbf{P}(\psi)$ we will prefer to write $(\mathcal{A}_M, \mathcal{S}_M)$ and α_ψ.

We can describe the observable algebras of a LCQFT by a functor $\mathbf{A}$: $\mathsf{Loc} \to \mathsf{Alg}$, where Alg is, e.g., a category of *-algebras with *-homomorphisms (that preserve the unit, as always) as morphisms. $\mathbf{A}$ can be obtained as the composition of $\mathbf{P}$ with a forgetful functor $F : \mathsf{Phys} \to \mathsf{Alg}$. We call $\mathbf{A}$ the observables functor of $\mathbf{P}$. Given an observables functor $\mathbf{A}$ we can try to define a LCQFT $\mathbf{P}$ by taking $\mathcal{S}_M = \mathcal{A}_M^{*,+,1}$. This is well-behaved iff $\mathcal{A}_M^{*,+,1}$ separates $\mathcal{A}_M$. If $\mathcal{A}_M$ can be given a locally convex topology (e.g. if $\mathcal{A}_M$ is a $U^\star$- or a $C^\star$-algebra) this follows from the Hahn-Banach theorem.

LCQFTs can enjoy many additional properties. Here are the two most basic ones:

Einstein causality: For any two morphisms $\psi_i : M_i \to M$ in Loc, $i \in \{1, 2\}$, such that $\psi_1(M_1)$ and $\psi_2(M_2)$ are spacelike to each other, the corresponding algebras $\alpha_{\psi_1}(\mathcal{A}_{M_1})$ and $\alpha_{\psi_2}(\mathcal{A}_{M_2})$ in $\mathcal{A}_M$ commute. This is motivated by the idea that operations localised at spacelike separation should not be able to influence each other, so the order in which these operations are performed should not matter. Note that this property only depends on the observables functor of the theory.

Time-slice axiom: For any morphism $\psi : M_1 \to M_2$ in Loc such that $\psi(M_1)$ contains a Cauchy surface for M, the correponding morphism α_ψ in Phys is an isomorphism. This means that $\alpha_\psi : \mathcal{A}_{M_1} \to \mathcal{A}_{M_2}$ is a *-isomorphism and $\alpha_\psi^* : \mathcal{S}_{M_2} \to \mathcal{S}_{M_1}$ is a bijection. This property is motivated by the idea that the theory should satisfy some dynamical law, so knowing a state in $\psi(M_1)$ should be enough to predict the state in the entire region M_2. Similarly, any observable in M_2 can be reexpressed using the dynamical law in terms of observables in $\psi(M_1)$. For the default state space $\mathcal{S}_M = \mathcal{A}_M^{*,+,1}$ one only needs to verify that α_ψ is a *-isomorphism.

Remark 2.4.2. [6] point out that their framework, and the variation of it presented in this section, generalises some of the ideas of algebraic QFT (AQFT) as pioneered by Haag and Kastler [18]. Indeed, if we restrict our attention to localisation regions that are subregions of Minkowski space, then we naturally obtain a net of *-algebras and the functorial behavior naturally leads to an appropriate action of the Poincaré group. Einstein causality and the time-slice axiom in AQFT follow from those properties in LCQFT.

Some structures in AQFT, e.g. those referring to the vacuum state, cannot be generalised, however, because there is no local and covariant choice of a preferred and well-behaved state that could replace the vacuum [14].

2.4.1 Lagrangian QFTs as LCQFTs

The dynamical algebras of Section 2.2 can sometimes be pieced together to an observables functor $\mathbf{A}^{(\mathcal{L})} : \mathsf{Loc} \to \mathsf{Alg}$ for a LCQFT. To see this we fix a Lagrange density $\mathcal{L}[g, \phi]$ of the form (2.1) with coefficient functions that depend in a local and covariant way on the background metric. For simplicity we will take the c_n to be constant, but functions of the scalar curvature R or other geometric quantities would also work. The formula for the Lagrange density $\mathcal{L}[g, \phi]$ can be used on every globally hyperbolic Lorentzian manifold $M = (\mathcal{M}, g_{ab})$ (with the same metric as in $\mathcal{L}$) to define a dynamical algebra $\mathcal{A}_M^{(\mathcal{L})}$.

To define morphisms α_ψ associated to any morphism $\psi : M_1 \to M_2$ in Loc we proceed as follows. Any configuration $\phi \in C^\infty(M_2, \mathbb{R})$ can be pulled back to a configuration $\psi^*\phi = \phi \circ \psi \in C^\infty(M_1, \mathbb{R})$ and any functional F on $C^\infty(M_1, \mathbb{R})$ can therefore be pushed forward to a functional $(\psi_*F)[\phi] := F[\psi^*\phi]$ on $C^\infty(M_2, \mathbb{R})$. For $F \in \mathcal{F}(M_1, \mathcal{L})$ we will show that ψ_*F is in $\mathcal{F}(M_2, \mathcal{L})$. F is of the form (2.4) and the compactly supported smooth coefficients c_n and tensor γ^{ab} on M_1 can be pushed forward to $\psi(M_2)$ and then extended by 0 to all of M_2. These pushed forward coefficients define a functional $\tilde{F}$ in $\mathcal{F}(M_2, \mathcal{L})$. Using the fact that ψ locally preserves the volume form one verifies that $\tilde{F}[\phi] = F[\psi^*\phi]$, i.e. $\tilde{F} = \psi_*F$ and hence $\psi_*F \in \mathcal{F}_{M_2}$. We now define a *-homomorphism $\alpha_\psi : \mathcal{A}_{M_1}^{(\mathcal{L})} \to \mathcal{A}_{M_2}^{(\mathcal{L})}$ by setting

$$\alpha_\psi(U_{M_1}(F)) := U_{M_2}(\psi_*F)\,,$$

where we put subscripts on the generators of the algebras in order to distinguish them. To see that this is a well-defined *-homomorphism one needs to verify that the defining relations of the dynamical algebras are preserved. The unitarity and normalisation are easily verified. For the dynamical relation we note that $\psi_*(F[.+\phi_0]) = (\psi_*F)[.+\psi_*\phi_0]$, where $\psi_*\phi_0$ is defined by extension by 0, as for the coefficients c_n above. Putting subscripts on the relative actions to distinguish them we also have $\psi_*(\delta S_{M_1}(\phi_0)) = \delta S_{M_2}(\psi_*\phi_0)$

due to the geometric nature of the Lagrange densities. It follows that

$$\begin{aligned}\alpha_\psi(U_{M_1}(F[.+\phi_0]+\delta S_{M_1}(\phi_0))) &= U_{M_2}(\psi_*(F[.+\phi_0]+\delta S_{M_1}(\phi_0)))\\ &= U_{M_2}((\psi_*F)[.+\psi_*\phi_0]+\delta S_{M_2}(\psi_*\phi_0)))\\ &= U_{M_2}(\psi_*F)\\ &= \alpha_\psi(U_{M_1}(F))\end{aligned}$$

as desired. For the causal factorisation we note that ψ preserves the causal structure of the background metrics. Moreover, if the metric gets perturbed by $F \in \mathcal{F}(M_1, \mathcal{L})$, then $\psi : \mathcal{M}_1 \to \mathcal{M}_2$ still maps the metric $(g_F)_{ab}$ on $\mathcal{M}_1$ to $(g_{\psi_*F})_{ab}$ on $\mathcal{M}_2$ and it preserves the corresponding causal structures. Indeed, if γ is a causal curve in $(\mathcal{M}_2, g_{\psi_*F})$ between two points $x, y \in \psi(\mathcal{M}_1)$, then γ must lie entirely in $\psi(\mathcal{M}_1)$, for otherwise there is a piece γ' of γ that connects two points x', y' in $\psi_1(\mathcal{M}_1)$ in a region where the metric is unperturbed, but such that γ' does not lie entirely in $\psi(\mathcal{M}_1)$. By assumption on the morphisms of Loc this is not possible. Since ψ preserves also the perturbed causal structures, the required support properties of F_1 and F_2 in the causal factorisation are preserved under the push forward by ψ.

Einstein causality of the Lagrangian QFTs $\mathbf{A}^{(\mathcal{L})}$ follows immediately from the causal factorisation (with $F_3 = 0$).

To study the time-slice axiom for $\mathbf{A}^{(\mathcal{L})}$ we consider a morphism $\psi : N \to M$ in Loc such that $\psi(N)$ contains a Cauchy surface for M. Now let $F \in \mathcal{F}(M, \mathcal{L})$ and consider $U(F)$. We want to know if there exists an $A \in \mathcal{A}^{(\mathcal{L})}_{\psi(N)}$ such that $U(F) = A$. For constant functionals $F = c$ we can take $A = e^{ic}I$, because the support of F is empty. For a linear functional $F = L_f$ with $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ in a free scalar field theory, $\mathcal{L} = \mathcal{L}_{\mathrm{fsf}}$, we can use equation (2.11) to shift the support of f to N using essentially the classical well-posedness of the equation of motion. (This would fail in the off-shell algebra.)

For functionals of higher order in ϕ, however, the time-slice axiom fails, even for free fields. The reason is that the dynamical axiom of $\mathcal{A}^{(\mathcal{L})}_M$ does not completely fix the dynamics for such functionals. We will now fix this (but not in the most general setting).

Definition 2.4.3. The *support* $\mathrm{supp}(\chi)$ of a diffeomorphism $\chi : \mathcal{M} \to \mathcal{M}$ is the smallest closed subset $C \subset \mathcal{M}$ such that $\chi(x) = x$ for all $x \notin C$. We will write $\mathrm{Diff}_c(\mathcal{M})$ for the group of diffeomorphisms of $\mathcal{M}$ of compact support.

An easy class of examples of elements of $\mathrm{Diff}_c(\mathcal{M})$ can be obtained by considering the flow under a compactly supported smooth vector field on $\mathcal{M}$.

Given an object $M = (\mathcal{M}, g_{ab})$ in Loc, any element $\chi \in \mathrm{Diff}_c(\mathcal{M})$ defines a morphism $\chi : M \to M'$ with $M' = (\mathcal{M}, (\chi_* g)_{ab})$ and by pushing functionals forward, $(\chi_* F)[\phi] = F[\chi^* \phi]$, we obtain a bijection $\mathcal{F}(M, \mathcal{L}) \to \mathcal{F}(M', \mathcal{L})$. (Here we used the fact that $\chi_* \mathcal{L} = \mathcal{L}$ when $\mathcal{L}$ defines a LCQFT). The functorial behaviour of the theory then yields a *-isomorphism $\beta_\chi : \mathcal{A}_M^{(\mathcal{L})} \to \mathcal{A}_{M'}^{(\mathcal{L})}$.

When $g'_{ab} = g_{ab}$, β_χ is a *-automorphism of the dynamical algebra $\mathcal{A}_M^{(\mathcal{L})}$ itself. Furthermore, for any $F \in \mathcal{F}(M, \mathcal{L})$, if χ does not change the metric in the support of F too much, then $\chi_* F$ is again in $\mathcal{F}(M, \mathcal{L})$ and it defines a *-homomorphism from a subset of $\mathcal{A}_M^{(\mathcal{L})}$ into itself.

The following property is a simplified version of the unitary anomalous master Ward identity.

Definition 2.4.4. $\mathbf{A}^{(\mathcal{L})}$ is *free of diffeomorphism anomalies* iff for all objects $M = (\mathcal{M}, g_{ab})$ and all $\chi \in \mathrm{Diff}_c(\mathcal{M})$ there is a *-automorphism β_g on $\mathcal{A}_M^{(\mathcal{L})}$ such that $\beta_g(U(F)) = U(\chi_*(F))$ for all $F \in \mathcal{F}(M, \mathcal{L})$ for which the right-hand side is well-defined.

Note in particular that $\beta_g(U(F)) = U(F)$ when the supports of χ and F are disjoint.

If $\mathbf{A}^{(\mathcal{L})}$ is free of diffeomorphism anomalies in this sense and $F \in \mathcal{F}(M, \mathcal{L})$, then one can use a suitable (small) χ to move the support of F to that of $\chi_* F$. Repeating this a finite number of times one can find a functional supported in any given neighbourhood N of a Cauchy surface of M and thereby prove the time-slice axiom. We refer to [5] for a fuller discussion of symmetries, the related renormalisation group and unitary master Ward identity.

Chapter 3

The stress tensor in QFT

In classical field theory, the stress tensor encodes the response of the Lagrange density, and therefore the theory, to an infinitesimal variation of the metric, cf. equation (1.10). With the stress tensor in hand we can verify various energy conditions, which help to encode the stability of matter and the attractive nature of gravity. In QFT there should also be a stress tensor that fulfills the same purpose, but which can now be an operator-valued distribution. (Note in particular that the Lagrangian QFTs of Section 2.2 can be defined equally well for Lagrange densities that correspond to clasically unstable systems, e.g. with potential energies that are unbounded from below like $V(\phi) = -\phi^4$.) This means that stabililty of matter has not been built into the framework yet. In this chapter we therefore want to investigate the quantum stress tensor and its properties.

3.1 Metric perturbations in LCQFT

A LCQFT $\mathbf{P} : \mathsf{Loc} \to \mathsf{Phys}$ already tells us how the theory depends on the background metric, so some information about the stress tensor is already encoded by relative Cauchy evolutions [6], a concept which we will now review. We will assume that we are given an observable functor $\mathbf{P}$ satisfying the time-slice axiom.

Let $M = (\mathcal{M}, g_{ab})$ and $M' = (\mathcal{M}, g'_{ab})$ be any two objects in Loc with the same underlying manifold $\mathcal{M}$ and such that $g'_{ab} - g_{ab}$ is supported in a compact set K. (We also assume that M and M' are endowed with the same orientation and with time orientations that coincide outside K.) We want to express the effect that the change of the metric from g_{ab} to g'_{ab} has as an isomorphism on $\mathcal{A}_M$. For this purpose we choose two Cauchy surfaces, $\Sigma_\pm \subset M$, such that K lies to the future of Σ_- and to the past of Σ_+. The past of Σ_-, $\mathcal{M}_- := J^-(\Sigma_-)$, defines an object $M_- = (\mathcal{M}_-, g_{ab}|_{\mathcal{M}_-})$ in Loc and the inclusion $\mathcal{M}_- \subset \mathcal{M}$ defines a morphism ι_- in Loc. Similarly, $\mathcal{M}_+ := J^+(\Sigma_+)$ defines an object $M_+ = (\mathcal{M}_+, g_{ab}|_{\mathcal{M}_+})$ in Loc and the inclusion $\mathcal{M}_+ \subset \mathcal{M}$ defines a morphism ι_+. Note that the regions $\mathcal{M}_-$ and $\mathcal{M}_+$ both contain Cauchy surfaces for M, so by the time-slice axiom there are *-isomorphisms

$$\begin{aligned}\alpha_{\iota_-} &: \mathcal{A}_{M_-} \to \mathcal{A}_M \\ \alpha_{\iota_+} &: \mathcal{A}_{M_+} \to \mathcal{A}_M \,.\end{aligned}$$

Because $g'_{ab} = g_{ab}$ outside K the inclusions $\mathcal{M}_- \subset \mathcal{M}$ and $\mathcal{M}_+ \subset \mathcal{M}$ also define morphisms $\theta_- : M_- \to M'$ and $\theta_+ : M_+ \to M'$. Here too the regions $\mathcal{M}_\pm$ contain Cauchy surfaces for M', so there are *-isomorphisms

$$\begin{aligned}\alpha_{\theta_-} &: \mathcal{A}_{M_-} \to \mathcal{A}_{M'} \\ \alpha_{\theta_+} &: \mathcal{A}_{M_+} \to \mathcal{A}_{M'} \,.\end{aligned}$$

The effect of the change of metric on an operator $A \in \mathsf{Alg}_M$ can now be found by expressing A in the region $\mathcal{M}_-$, propagating it through M' to $\mathcal{M}_+$ and expressing the result again in Alg_M. I.e., we consider the *-isomorphism

$$\alpha_{g',g} := \alpha_{\iota_+} \circ \alpha_{\theta_+}^{-1} \circ \alpha_{\theta_-} \circ \alpha_{\iota_-}^{-1} \tag{3.1}$$

on $\mathcal{A}_M$. This isomorphism is called a *relative Cauchy evolution*.

For the special case of Lagrangian QFTs, $\mathcal{A}_M^{(\mathcal{L})}$, there are is another way in which we can encode the effect of the change in the metric from g_{ab} to g'_{ab}. We can consider the ensuing change in the action, which is given by a functional $F[\phi] = \int_\mathcal{M} \mathcal{L}[g', \phi] - \mathcal{L}[g, \phi]$ of the form (2.4) and the adjoint action of the corresponding unitary $U(F)$ then defines a *-automorphism $\beta_{g',g}$ of $\mathcal{A}_M^{(\mathcal{L})}$. If the theory satisfies the time-slice axiom (e.g. in a representation), then one might expect $\beta_{g',g} = \alpha_{g',g}$.

Alternatively, if $G[\phi] = \int_\mathcal{M} \theta^{ab} T_{ab}[\phi] \mathrm{d}vol_g$ is the classical stress tensor, smeared with a symmetric test tensor θ^{ab}, then G is a functional of the form (2.4) and we can consider the *-automorphism τ_θ of $\mathcal{A}_M^{(\mathcal{L})}$ defined by the adjoint action of $U(G)$, which would encode the response to an infinitesimal variation of the metric.

3.1.1 Example: the free scalar field

As an example we consider the free scalar field of Section 2.3 with Lagrange density $\mathcal{L}_{\mathrm{fsf}}$ and observables functor $\mathbf{A}^{(\mathcal{L}_{\mathrm{fsf}})}$, which is a LCQFT. For the Weyl operators we know from equation (2.11) how the time-slice axioms works, even without requiring that the theory is free of diffeomorphism anomalies. For this reason we now focus on the observables functor $\mathbf{W}^{(\mathcal{L}_{\mathrm{fsf}})}$ that assigns to each object M the subalgebra $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ of $\mathcal{A}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ that is generated by all functionals that are at most of first order in the fields, $F = L_f + c$ for some $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$, $c \in \mathbb{R}$. For constant functionals we trivially have $U(c) = e^{ic}I$ and as *-isomorphisms preserve the unit we have $\alpha_{g',g}(U(c)) = \beta_{g',g}(U(c)) = \tau_\theta(U(c)) = U(c)$.

To clarify the following computations, let us denote the Weyl operators in $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ by $W_M(f)$ with $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$. Given $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ we can use Lemma 2.3.4 to find $f_-, \phi_- \in C_0^\infty(\mathcal{M}, \mathbb{R})$ such that $\mathrm{supp}(f_-) \subset \mathcal{M}_-$ and $f = f_- + P\phi_-$. Similarly we can find $f_+, \phi_+ \in C_0^\infty(\mathcal{M}, \mathbb{R})$ such that $\mathrm{supp}(f_+) \subset \mathcal{M}_+$ and $f_- = f_+ + P'\phi_+$, where P' is the Klein-Gordon operator for the perturbed metric g'_{ab}. Using equation (2.11) and some algebra we then have $W_M(f) = W_M(f_-)$, $W_{M'}(f_-) = W_{M'}(f_+)$ and hence

$$\begin{aligned}
\alpha_{g',g}(W_M(f)) &= \alpha_{\iota_+} \circ \alpha_{\theta_+}^{-1} \circ \alpha_{\theta_-} \circ \alpha_{\iota_-}^{-1}(W_M(f_-)) \\
&= \alpha_{\iota_+} \circ \alpha_{\theta_+}^{-1} \circ \alpha_{\theta_-}(W_{M_-}(f_-)) \\
&= \alpha_{\iota_+} \circ \alpha_{\theta_+}^{-1}(W_{M'}(f_-)) \\
&= \alpha_{\iota_+} \circ \alpha_{\theta_+}^{-1}(W_{M'}(f_+)) \\
&= \alpha_{\iota_+}(W_{M_+}(f_+)) \\
&= W_M(f_+)\,,
\end{aligned}$$

where $f_+ = f - P\phi_- - P'\phi_+$. Note that if $g'_{ab} = g_{ab}$, then $f_+ = f - P(\phi_- + \phi_+)$ and from the dynamical relation we then find $W_M(f_+) = W_M(f)$.

Let us now consider the adjoint action $\beta_{g',g}$ of $U(F)$ in $\mathcal{A}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ acting on Weyl operators, where F is the change in the action. Note that F is then a quadratic functional of the form $F[\phi] = \frac{1}{2}\langle \phi, Q\phi \rangle$, where Q is the second order partial differential operator

$$Q = -\nabla_a \gamma^{ab} \nabla_b + (\mu_{g'} - 1)m^2 + \xi(\mu_{g'} R' - R)\,.$$

One can show that $P + Q = \mu_{g'} P'$ (using $\nabla_a(\mu_{g'} V^a) = \mu_{g'} \nabla'_a V^a$ for any vector field V^a). This operator is normally hyperbolic and has unique advanced and retarded fundamental solutions (w.r.t. the metric $\mu_{g'}^{-1} g'_{ab}$) given by $E'^{\pm} \mu_{g'}^{-1}$.

Given any $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ we note that $f_1 := f + QE^- f = \mu_{g'} P' E^- f$ is also compactly supported and hence so is

$$\tilde{f} := PE'^+ \frac{1}{\mu_{g'}} f_1 = f_1 - QE'^+ \frac{1}{\mu_{g'}} f_1 \,.$$

Moreover, using the fact that $E'^+ \frac{1}{\mu_{g'}} f_1$ has past spacelike compact support,

$$(I + QE^+)\tilde{f} = (P+Q)E^+ PE'^+ \frac{1}{\mu_{g'}} f_1 = (P+Q)E'^+ \frac{1}{\mu_{g'}} f_1 = f_1 \,.$$

With this identity, equations (2.12,2.13) and $F[.+h] = F + L_{Qh} + F[h]$ we then compute in $\mathcal{A}_M^{(\mathcal{L}_{\mathrm{fsf}})}$

$$\begin{aligned}
U(F)U(L_f)U(F)^{-1} &= U(L_f + F[.+E^- f])U(F)^{-1} \\
&= U(L_{f_1} + F + F[E^- f])U(F)^{-1} \\
&= e^{iF[E^- f] - iF[E^+ \tilde{f}]} U(L_{\tilde{f} + QE^+ \tilde{f}} + F + F[E^+ \tilde{f}])U(F)^{-1} \\
&= e^{iF[E^- f] - iF[E^+ \tilde{f}]} U(L_{\tilde{f}} + F[.+E^+ \tilde{f}])U(F)^{-1} \\
&= e^{iF[E^- f] - iF[E^+ \tilde{f}]} U(L_{\tilde{f}}) \,.
\end{aligned}$$

With some further computations one can show that the phase factor works out nicer if we consider Weyl operators:

$$\beta_{g',g}(W(f)) = U(F)W(f)U(F)^{-1} = W(\tilde{f}) \,.$$

Comparing $\tilde{f} = PE'^+ P' E^- f$ and f_+ we can use the supports of $f_\pm$ in the future and past regions to see that $P'E^- f_- = PE^- f_- = f_-$ and $PE'^+ f_+ = P'E'^+ f_+ = f_+$ and hence

$$\begin{aligned}
f_1 &= \mu_{g'} P' E^- f = \mu_{g'} P' E^- (f_- + P\phi_-) = \mu_{g'} (f_- + P'\phi_-) \\
\tilde{f} &= PE'^+ \frac{1}{\mu_{g'}} f_1 = PE'^+ (f_- + P'\phi_-) \\
&= PE'^+ (f_+ + P'\phi_+ + P'\phi_-) = f_+ + P(\phi_+ + \phi_-) \,.
\end{aligned}$$

Thus we see that $W(\tilde{f}) = W(f_+)$, i.e. $\alpha_{g',g}(W(f)) = \beta_{g',g}(W(f))$ and $U(F)$ implements the relative Cauchy evolution on Weyl operators.

Similarly, if $G[\phi]$ is the stress tensor averaged with test-tensor θ^{ab}, then, in the minimally coupled case ($\xi = 0$),

$$\begin{aligned} G[\phi] &= \int_{\mathcal{M}} \theta^{ab}\partial_a\phi \cdot \partial_b\phi - \frac{1}{2}\theta\left(g^{ab}\partial_a\phi \cdot \partial_b\phi + m^2\phi^2\right) \mathrm{d}vol_g \\ &= \int_{\mathcal{M}} \gamma^{ab}\partial_a\phi \cdot \partial_b\phi - \frac{1}{2}\eta m^2\phi^2 \mathrm{d}vol_g\,, \end{aligned}$$

where $\theta = g_{ab}\theta^{ab}$ and we set $\gamma^{ab} = \theta^{ab} - \theta g^{ab}$ in the second line. For suitable (small) θ^{ab}, $\hat{g}^{ab} = g^{ab} + \gamma^{ab} = (1-\theta)g^{ab} + \theta^{ab}$ is an inverse Lorentzian metric and the functional G is again quadratic. Using similar computations as before we find $\tau_\theta(W(f)) = U(G)W(f)U(G)^{-1} = W(\hat{f})$ with $\hat{f} = P\hat{E}^+\hat{P}E^-f$. In general we do not expect $\tau_\theta(W(f)) = \alpha_{g',g}(W(f))$, because $\hat{g}_{ab} \neq g'_{ab}$.

3.1.2 The stress tensor

To obtain a stress tensor we need to consider infinitesimal variations of the metric. We let $(g_\lambda)_{ab}$ with $\lambda \in \mathbb{R}$ be a one-parameter family of Lorentzian metrics with $(g_0)_{ab} = g_{ab}$ and we let $\theta^{ab} := \partial_\lambda g_\lambda^{ab}|_{\lambda=0} = -g^{ac}g^{bd}\partial_\lambda(g_\lambda)_{cd}|_{\lambda=0}$ denote the infinitesimal variation in the metric. For sufficiently small λ we obtain a relative Cauchy evolution $\alpha_\lambda := \alpha_{g_\lambda,g}$ as a *-isomorphism on $\mathbf{A}_M$. In the case of $\mathcal{A}_M^{(\mathcal{L})}$ we silimarly have one-paramater families $\beta_{g_\lambda,g}$, resp. $\tau_{\lambda\theta}$, implemented by unitaries $U(F(\lambda))$, resp. $U(\lambda G)$.

Heuristically, the stress tensor should be a self-adjoint quantum field, which becomes a (possibly unbounded) operator $T = \int_{\mathcal{M}} \theta^{ab}T_{ab}\mathrm{d}vol_g$ when smeared with the test tensor field θ^{ab}. The relation between T and α_λ is that $\alpha_\lambda(A) \simeq e^{\frac{i}{2}\lambda T}Ae^{-\frac{i}{2}\lambda T}$ (at least up to first order in λ), where we took the factor 2 in equation (1.10) into account. Hence, for all $A \in \mathcal{A}_M$

$$\partial_\lambda\alpha_\lambda(A)|_{\lambda=0} = \frac{i}{2}[T, A] \tag{3.2}$$

should depend only on θ^{ab}. Note that T may contain contributions in the centre of the algebra $\mathcal{A}_M$ that cannot be determined from $\partial_\lambda\alpha_\lambda|_{\lambda=0}$ in this way.

In the case of $\mathcal{A}_M^{(\mathcal{L})}$, similar heuristic arguments lead us to expect

$$T = -i\partial_\lambda U(F(\lambda))|_{\lambda=0} = -i\partial_s U(sG)|_{s=0} \tag{3.3}$$

and consequently $\partial_\lambda\beta_{g_\lambda,g}(A)|_{\lambda=0} = \partial_\lambda U(\lambda G)AU(\lambda G)^{-1}|_{\lambda=0} = i[T, A]$.

The derivatives in (3.2,3.3) are problematic, however. They require a suitable choice of topology on the algebra and for Bosons they typically fail to converge in, say, the topology of a $C^\star$-norm. (The Weyl algebra has a unique $C^\star$-norm, in which we have $\|W(f)-W(h)\|=2$ for all $f\neq h$. Hence the limit $\lim_{\lambda\to 0}\frac{W(f_+(\lambda))-W(f)}{\lambda}$ does not exist as a norm limit.) We are often in a better position when we consider representations π of $\mathcal{A}_M$ on a Hilbert space $\mathcal{H}$. In that case one can consider a suitable (dense) set of vectors $\mathcal{V}\subset\mathcal{H}$ and consider the weak limit, e.g.

$$\partial_\lambda\langle\xi,\pi(\alpha_\lambda(W(f)))\psi\rangle|_{\lambda=0}. \tag{3.4}$$

for $\xi,\psi\in\mathcal{V}$. For nice representations and perhaps a nice class of operators in $\mathcal{A}_M$ one then expects this limit to exist, but the details clearly depend on the theory and the classes of states and observables being considered.

For the Weyl operators of a free scalar field in the GNS-representations of quasi-free Hadamard states, which we will discuss in Section 3.2, the weak derivative (3.4) has been determined in [6] and leads to (3.2) for a stress tensor that we will discuss in Section 3.2 below. Moreover, one can show that

$$\partial_\lambda\tilde{f}|_{\lambda=0}=\partial_\lambda\hat{f}|_{\lambda=0}\,,$$

so infinitesimally, $\beta_{g_\lambda,g}$ and $\tau_{\lambda\theta}$ coincide.

3.2 The renormalized quantum stress tensor for free scalar fields

As in Section 3.1.1 we will consider a free scalar quantum field with Lagrange density $\mathcal{L}_{\rm fsf}$ and the $U^\star$-subalgebra $\mathcal{W}_M^{(\mathcal{L}_{\rm fsf})}$ of $\mathcal{A}_M^{(\mathcal{L}_{\rm fsf})}$ generated by the Weyl operators $W(f)$.

One can show that every state on $\mathcal{W}_M^{(\mathcal{L}_{\rm fsf})}$ can be extended to a state on $\mathcal{A}_M^{(\mathcal{L}_{\rm fsf})}$ (taking closures in a suitable $C^\star$-norm and using the Hahn-Banach theorem). It is known that $\mathcal{W}_M^{(\mathcal{L}_{\rm fsf})}$, and hence also $\mathcal{A}_M^{(\mathcal{L}_{\rm fsf})}$, admits many states that exhibit unphysical behaviour. However, there is a good criterion to select a class of physically acceptable states on $\mathcal{W}_M^{(\mathcal{L}_{\rm fsf})}$.

Definition 3.2.1. We will call a state $\omega : \mathcal{W}_M^{(\mathcal{L}_{\rm fsf})} \to \mathbb{C}$ *regular*, if for all $n \in \mathbb{N}$, $f_1, \ldots, f_n \in C_0^\infty(\mathcal{M}, \mathbb{R})$ and $s \in \mathbb{R}^n$ the map $s \mapsto \omega(W(s_1 f_1) \cdots W(s_n f_n))$ is smooth and the maps

$$\omega_n(f_1, \ldots, f_n) := (-i)^n \partial_{s_1} \cdots \partial_{s_n} \omega(W(s_1 f_1) \cdots W(s_n f_n))|_{s_1 = \ldots = s_n = 0}$$

define distribution densities on $\mathcal{M}^n$. We extend the ω_n by complex linearity and we call them the *n-point distributions* of the state ω.

For a general Lagrange density $\mathcal{L}$ it seems reasonable to expect that there exist states for which the map $s \mapsto \omega(U(s_1 F_1) \cdots U(s_n F_n))$ is smooth for all $F_1, \ldots, F_n \in \mathcal{F}(\mathcal{M}, \mathcal{L})$ and we could call such states smooth.

For $\mathcal{L}_{\rm fsf}$ we can heuristically think of $W(f)$ as a complex exponential $e^{i\phi(f)}$ with a quantum field ϕ, viewed as an operator-valued distribution, smeared with a test function f and thus $\omega_n(f_1, \ldots, f_n) = \omega(\phi(f_1) \cdots \phi(f_n))$. From the Weyl relations (2.15) we then see that

$$\begin{aligned} \omega_2(f_1, f_2) &= -\partial_{s_1} \partial_{s_2} e^{-is_1 s_2 \langle f_1, E f_2 \rangle} \omega(W(s_2 f_2) W(s_1 f_1))|_{s_1 = s_2 = 0} \\ &= \omega_2(f_2, f_1) + i \langle f_1, E f_2 \rangle \end{aligned}$$

which heuristically follows from $[\phi(f_1), \phi(f_2)] = i \langle f_1, E f_2 \rangle$.

Proposition 3.2.2. The two-point distribution of any regular state on $\mathcal{A}_M^{(\mathcal{L}_{\rm fsf})}$ has the following properties:

(1) **positive type:** $\omega_2(\bar{f}, f) \geq 0$ for all $f \in C_0^\infty(\mathcal{M}, \mathbb{C})$,

(2) **equation of motion:** $P^{(x)} \omega_2(x, y) = P^{(y)} \omega_2(x, y) = 0$,

(3) **canonical commutation relations (CCR):** $\omega_2(f, h) - \omega_2(h, f) = i \langle f, Eh \rangle$.

Conversely, given a distribution density ω_2 on $\mathcal{M}^2$ with these three properties, there exists a state on $\mathcal{W}_M^{(\mathcal{L}_{\rm fsf})}$ defined by

$$\omega(W(f)) := e^{-\frac{1}{2} \omega_2(f, f)} \tag{3.5}$$

for all $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$, which has ω_2 as its two-point distribution.

The proof of this result is omitted, cf. [38].

Definition 3.2.3. A state ω on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ of the form (3.5) is called *quasi-free* or *Gaussian.*

Prime examples of Gaussian states include e.g. vacuum and thermal states on stationary Lorentzian manifolds. In general, given any two-point distribution ω_2 there can be multiple states with that two-point distribution, but among them there is exactly one Gaussian state. For Gaussian states all other n-point distributions can be expressed in terms of ω_2 using Wick's theorem.

We now consider the GNS-quadruple $(\mathcal{H}_\omega, \Omega_\omega, \mathcal{D}_\omega, \pi_\omega)$ for any regular state ω on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$. One may show that the map

$$\begin{aligned}\Phi_1 : C_0^\infty(\mathcal{M}, \mathbb{R}) &\to \mathcal{H}_\omega \\ f &\mapsto -i\partial_s \pi_\omega(W(sf))\Omega_\omega|_{s=0}\end{aligned}$$

defines an $\mathcal{H}_\omega$-valued distribution density on $\mathcal{M}$. Note that $\Phi_1(Pf) = 0$, because ω_2 is a bi-solution to the Klein-Gordon equation. We are interested in the singularities of the distribution Φ_1, which we may investigate using microlocal techniques. We refer to [21] for details and give here only some basic definitions and a brief description of results relevant to our situation.

Definition 3.2.4. Let $\mathcal{H}$ be a Hilbert space, X a d-dimensional smooth manifold and $u : C_0^\infty(X) \to \mathcal{H}$ a Hilbert space-valued distribution density. A point $(x, \xi) \in T_x^*\mathcal{M}$ is called a *regular direction* iff there is a coordinate chart $\kappa : O \to U \subset \mathbb{R}^d$ on a contractible region $O \subset \mathcal{M}$ containing x, a test function $f \in C_0^\infty(O, \mathbb{R})$ with $f(x) \neq 0$ and an open conic subset $\Gamma \subset \mathbb{R}^d$ such that for each $N \in \mathbb{N}$ there is a $C_n > 0$ with

$$(1 + \|\xi\|^N)\|(\kappa_* u)((\kappa_* f)e^{-i\xi \cdot \cdot})\| \leq C_N \qquad \xi \in \Gamma\,.$$

The *(smooth) wave front set* of u is the set $\{(x, \xi) \in T^*\mathcal{M} \mid \xi \neq 0,\ (x, \xi)$ is not a regular direction for $u\}$.

One may show that that the wave front set does not depend on the choice of f or κ, so it is a well-defined closed subset of $T^*\mathcal{M} \setminus \mathcal{M} \times \{0\}$. We can think of $(x, \xi) \in WF(u)$ as a point x and a direction ξ in which u is singular. In particular, if $WF(u) \cap T^*O = \emptyset$ for some open region $O \subset \mathcal{M}$, then u is smooth on O.

For solutions of partial differential equations, wave front sets have especially nice properties. In our case, $Pu = 0$, and one can show that this implies $WF(u) \subset \{(x,\xi) \in T^*\mathcal{M} \mid \xi_a\xi^a = 0\}$, so only null directions can be singular. Moreover, any singularity in $(x,\xi) \in WF(u)$ determines a unique null geodesic γ through x with tangent vector ξ^a and one can chose that the singularity propagates, in the sense that $(\gamma(s), d\gamma(s)) \in WF(u)$ for all parameter values s of γ.

Definition 3.2.5. We will call a regular state ω on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ a *Hadamard* state, if the corresponding distribution Φ_1 satisfies

$$WF(\Phi_1) \subset \{(x,\xi) \mid \xi \text{ is future pointing}\}\,. \tag{3.6}$$

One may show that this definition is equivalent to a definition in terms of $WF(\omega_2)$ [35]. For us the most important consequence will be that for any two Hadamard two-point distributions ω_2 and ω_2' the difference $\omega_2 - \omega_2'$ is smooth $C^\infty(\mathcal{M}^2, \mathbb{R})$.

The set of Hadamard states on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$ can be shown to be a well-behaved state space in the sense of Section 2.1. Furthermore, ground and thermal states on stationary Lorentzian manifolds are Hadamard and any state which is Hadamard in an open neighbourhood of a Cauchy surface is Hadamard on the entire Lorentzian manifold, so the property is preserved under time evolution.

We now want to define the quantum stress tensor of the free scalar field in analogy to the classical one, which is given in equation (1.11). For this we first consider the GNS-representation $(\mathcal{H}_\omega, \Omega_\omega, \mathcal{D}_\omega, \pi_\omega)$ of a regular state ω on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$. For any $f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ we define the represented smeared quantum field operator $\phi_\omega(f)$ on the domain $\mathcal{D}_\omega = \pi_\omega(\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})})\Omega_\omega$ by

$$\phi_\omega(f)\psi = -i\partial_s \pi_\omega(W(sf))\psi|_{s=0}\,,$$

where the derivative is well-defined. This operator is symmetric, because $W(sf)^* = W(-sf)$. Similarly one can define any polynomial of smeared field operators acting on $\mathcal{D}_\omega$ by taking derivatives.

Note that the classical expression for the stress tensor in equation (1.11) involves products of the classical fields taken at every point. However, because ϕ_ω is a distribution, we cannot simply take pointwise products. Indeed, a careful examination shows that such products are not well defined[1] and we will need to regularise and renormalise the stress tensor.

[1]For a Hadamard state the distribution $\phi_\omega(x)\psi$ has its wave front set in the forward

For this purpose we first write the classical stress tensor in a point-split form. We will use the embedding $\iota : \mathcal{M} \to \mathcal{M}^2$ as the diagonal subset, $\iota(x) := (x, x)$, so we can pull-back tensors of type $(0, n)$ from $\mathcal{M}^2$ to $\mathcal{M}$. We will indicate points on $\mathcal{M}^2$ by $(\overline{x}, \underline{x})$ and we will use $T_{(\overline{x},\underline{x})}\mathcal{M}^2 \simeq T_{\overline{x}}\mathcal{M} \oplus T_{\underline{x}}\mathcal{M}$ to identify indices on $\mathcal{M}^2$ with those on $\mathcal{M}$. We then want to choose a partial differential operator $T^{\mathrm{split}}_{\overline{a}\underline{b}}$ on $\mathcal{M}^2$ such that $T_{ab}[\phi] = \iota^*(T^{\mathrm{split}}_{\overline{a}\underline{b}}\phi \otimes \phi)$. For simplicity we will only consider the minimally coupled case and choose

$$T^{\mathrm{split}}_{\overline{a}\underline{b}} := \partial_{\overline{a}}\partial_{\underline{b}} - \frac{1}{2} g_{\overline{a}\underline{b}} \left(g^{\overline{c}\underline{d}} \partial_{\overline{c}} \partial_{\underline{d}} + m^2 \right) ,$$

where $g_{\overline{a}\underline{b}}$ is a tensor field on $\mathcal{M}^2$ such that $\iota^* g_{\overline{a}\underline{b}} = g_{ab}$ and $g^{\overline{c}\underline{d}}(\overline{x}, \underline{x}) = g^{\overline{ca}}(\overline{x}) g^{\underline{db}}(\underline{x}) g_{\overline{a}\underline{b}}(\overline{x}, \underline{x})$.

Proceeding to the quantum case we want to replace the classical field ϕ by a (represented) quantum field ϕ_ω. There is no problem in applying $T^{\mathrm{split}}_{\overline{a}\underline{b}}$ in a weak sense,

$$\langle \eta, T^{\mathrm{split}}_{\overline{a}\underline{b}} \phi_\omega(\overline{x})\phi_\omega(\underline{x})\eta \rangle = T^{\mathrm{split}}_{\overline{a}\underline{b}} \langle \eta, \phi_\omega(\overline{x})\phi_\omega(\underline{x})\eta \rangle$$

for all $\eta \in \mathcal{D}_\omega$. (Expectation values between two different vectors $\xi, \eta \in \mathcal{D}_\omega$ can be obtained using the polarisation identity.) In this way we can define $T^{\mathrm{split}}_{\overline{a}\underline{b}} \phi_\omega(\overline{x})\phi_\omega(\underline{x})$ as a distribution density on $\mathcal{M}^2$ with values in the quadratic forms on $\mathcal{D}_\omega$. The difficulty is that the pull-back of this expression under ι is ill-defined.

Now we will make essential use of the properties of Hadamard states. If ω is a Hadamard state and $\eta \in \mathcal{D}_\omega$ has $\|\eta\| = 1$, then η also defines a Hadamard state, ω_η, and hence $\omega_2 - (\omega_\eta)_2$ is smooth. Moreover, one can show that the singular structure of all Hadamard two-point distributions depends in a local and covariant way on the metric g_{ab}. Indeed, there exist distributions $H_2(\overline{x}, \underline{x})$ on a neighbourhood of the diagonal in $\mathcal{M}^2$, defined in a local way terms of the metric g_{ab}, such that $\omega_2 - H_2$ is a C^2 function near the diagonal. The same is then true for all Hadamard two-point distributions. (The distribution H_2 can be defined using Hadamard series expansions for parametrices of the Klein-Gordon equation, which is the origin of the name

light cone. If the wave front set of $\phi_\omega(x)\phi_\omega(y)\psi$ were similarly restricted to two copies of the forward light cone, then one could restrict the distribution to $x = y$ and take pointwise products. Unfortunately this is not the case, however, as one can see by considering $\phi_\omega(x)\phi_\omega(y)\psi - \phi_\omega(y)\phi_\omega(x)\psi = iE(x, y)\psi$, because the wave front set of E does not satisfy the desired restriction.

Hadamard state.) We then define the *regularised stress tensor* by

$$\langle \eta, T^{\mathrm{reg}}_{ab}(x)\eta\rangle := \iota^*(T^{\mathrm{split}}_{\overline{a}\underline{b}}((\omega_\eta)_2 - H_2))(x)$$

if $\|\eta\| = 1$ (and scaling quadratically in η otherwise).

Apart from the functional analytic questions about extending this quadratic form to an operator, the regularised stress tensor still has an important drawback. The expectation values of the regularised stress tensor $\omega(T^{\mathrm{reg}}_{ab}(x))$ yield a smooth function which is typically not conserved, i.e. $\nabla^a\omega(T^{\mathrm{reg}}_{ab}(x)) \neq 0$. This is problematic if we want to consider the semi-classical Einstein equation, $R_{ab} - \frac{1}{2}g_{ab}R + \Lambda g_{ab} = 8\pi\omega(T^{\mathrm{reg}}_{ab})$, because the left-hand side is conserved and hence the equation cannot be satisfied. One can show, however, that the divergence $\nabla^a\omega(T^{\mathrm{reg}}_{ab}(x))$ depends in a local and covariant way on the metric and it is in fact independent of the state ω. Moreover, one can prove the existence of a local and covariant function Q of the metric such that the *renormalised stress tensor* $T^{\mathrm{ren}}_{ab} = T^{\mathrm{reg}}_{ab} - Qg_{ab}I$, defined through

$$\omega(T^{\mathrm{ren}}_{ab}(x)) := \iota^*(T^{\mathrm{split}}_{\overline{a}\underline{b}}(\omega_2 - H_2))(x) - Q(x)g_{ab}(x) \tag{3.7}$$

$$\tag{3.8}$$

does lead to a conserved quantity.

Remark 3.2.6. (a) (3.7) can be viewed as a distribution density with values in the quadratic forms. For each Hadamard state ω the expectation value is a *smooth* tensor field in x.

(b) The definition (3.7) of the renormalized stress tensor involves some arbitrariness in the choice of H_2. Even if we impose a number of natural conditions on H_2, including local covariance and suitable scaling behaviour, there still remains a certain amount of *renormalisation freedom*. This freedom has been determined and can be expressed in terms of a finite number of parameters and curvature tensors. In the dynamical algebra approach this renormalisation freedom is expected to appear in the GNS-representation, once a state has been chosen.

(c) For interacting fields, a renormalised stress tensor can be defined perturbatively. In that context it has also been shown, that one can choose it to be divergence free [20].

3.3 Energy conditions and QEIs in QFT

For a free scalar QFT as in Section 3.1.1 we will now consider the validity of the energy conditions, particularly the WEC. We cannot expect these pointwise conditions to hold, even in Minkowski space, due to the renormalization. Roughly speaking, if we were to take the expectation value of the 00-component (in some inertial coordinates) of the point split stress tensor and take the limit of coinciding points, $\overline{x} = \underline{x} = x$, then we get something positive, but it is $+\infty$. When we regularise, we subtract a term that also gives us $+\infty$, in such a way that the difference $\omega(T^{\mathrm{ren}}_{ab}(x))$ is finite when ω is any Hadamard state (in any suitable representation). A priori there is no guarantee, however, that this expectation value is positive.

3.3.1 Violation of the NEC in QFT

We consider a free scalar field in d-dimensional Minkowski space with $d \geq 3$ or $m > 0$, so that a vacuum state exists. For any Hadamard state the NEC at the point 0 requires that for every null vector n^a

$$n^a n^b \partial_a \partial_b (\omega_2 - H_2)(0) \geq 0 \,.$$

We will show, however, that the left-hand side can be made arbitrarily negative by choosing a suitable ω_2. If the NEC can be violated by an arbitrary amount, then so can the DEC, WEC and SEC.

Any Hadamard two-point distribution ω_2 in Minkowski space defines a 2×2-matrix of (distributional) initial data on Σ^2, where $\Sigma = \{x_0 = 0\}$ in inertial coordinates $x = (x_0, \mathbf{x})$. We can then express the initial data as

$$\begin{pmatrix} \omega_{2,00}(\mathbf{x},\mathbf{y}) & \omega_{2,01}(\mathbf{x},\mathbf{y}) \\ \omega_{2,10}(\mathbf{x},\mathbf{y}) & \omega_{2,11}(\mathbf{x},\mathbf{y}) \end{pmatrix} = \begin{pmatrix} \omega_2(x,y)|_{x_0=y_0=0} & \partial_{y_0}\omega_2(x,y)|_{x_0=y_0=0} \\ \partial_{x_0}\omega_2(x,y)|_{x_0=y_0=0} & \partial_{x_0}\partial_{y_0}\omega_2(x,y)|_{x_0=y_0=0} \end{pmatrix} .$$

One way to obtain these initial data is to consider $f, h \in C_0^\infty(\mathcal{M})$ and write $(f_0, f_1) = (Ef|_\Sigma, \partial_0 Ef|_\Sigma)$ and similarly for h. Noting that ω_2 is a solution of the Klein-Gordon equation in each variable we can use the computation (2.8) to find

$$\begin{aligned} \omega_2(f,h) &= \int_\Sigma f_0 \partial_0 \omega_2(.,h) - f_1 \omega_2(.,h) \mathrm{d}vol_\Sigma \\ &= \omega_{2,11}(f_0,h_0) - \omega_{2,10}(f_0,h_1) - \omega_{2,01}(f_1,h_0) + \omega_{2,00}(f_1,h_1) \,. \end{aligned}$$

The Minkowski vacuum state $\omega^{(0)}$ has the two-point distribution

$$\omega_2^{(0)}(x,y) = \frac{1}{2(2\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} e^{-i\nu(p)(y_0-x_0)+ip\cdot(\mathbf{y}-\mathbf{x})} \frac{1}{\nu(p)} \mathrm{d}^{\,d-1}p\,, \tag{3.9}$$

where $\nu(p) = \sqrt{|p|^2+m^2}$. Its initial data on Σ can be viewed as intergral kernels for operators on $L_2(\mathbb{R}^{d-1})$, which can be computed from (3.9). We can express them in terms of the operator $A := -\Delta + m^2$ as

$$\begin{pmatrix} \omega_{2,00}^{(0)} & \omega_{2,01}^{(0)} \\ \omega_{2,10}^{(0)} & \omega_{2,11}^{(0)} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} A^{-\frac{1}{2}} & -iI \\ iI & A^{+\frac{1}{2}} \end{pmatrix}.$$

Note that the matrix of operators on the right-hand side is of positive type, because for any $f_0, f_1 \in C_0^\infty(\Sigma)$ we have

$$\left\langle (f_1, f_0), \frac{1}{2} \begin{pmatrix} A^{-\frac{1}{2}} & -iI \\ iI & A^{\frac{1}{2}} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_0 \end{pmatrix} \right\rangle = \frac{1}{2} \|A^{-\frac{1}{4}} f_1 - iA^{\frac{1}{4}} f_0\|^2 \geq 0\,. \tag{3.10}$$

We will now prove that the violation of the NEC can be arbitrarily large. (See the appendix of [13] for a related argument.)

Proposition 3.3.1. There exists a one-parameter family of Hadamard two-point distributions $\omega_2^{(\lambda)}$ such that $\lim_{\lambda\to\infty} n^a n^b \partial_a \partial_b (\omega_2^{(\lambda)} - H_2)(0,0) = -\infty$ for any null vector n^a.

Proof: We will choose $\omega_2^{(\lambda)}$ by modifying the initial data of $\omega_2^{(0)}$. Let $u \in \mathcal{S}^\infty(\mathbb{R}^{d-1}, \mathbb{R})$ be a Schwarz function with $\nabla u(0) = 0$, $c' := \|(-\Delta)^{\frac{1}{4}} u\|^2 \neq 0$ in terms of the L_2-norm and $C := (\sqrt{-\Delta} u)(0) \neq 0$. (E.g. a Gaussian will do.)

Let $v := A^{\frac{1}{4}} u \neq 0$, which is also smooth, and set $c := \|v\|^2$ and P_v be the orthogonal projection in $L_2(\Sigma)$ onto $\frac{1}{c} v$. We then want to define ω_2 by the matrix of initial data

$$(\omega_{2,ij}(\mathbf{x},\mathbf{y})) = \frac{1}{2} \begin{pmatrix} A^{-\frac{1}{4}}(I+P_v)A^{-\frac{1}{4}} & -iI \\ iI & A^{\frac{1}{4}}(I - \frac{1}{2}P_v)A^{\frac{1}{4}} \end{pmatrix}.$$

Note that the diagonal entries are inverse to each other, so ω_2 is of positive type by a computation analogous to (3.10). The equations of motion are built in by the initial value definition of ω_2, the off diagonal terms ensure the CCR and the difference $\omega_2 - \omega_2^{(0)}$ has smooth initial data $w_{2,00}(\mathbf{x},\mathbf{y}) = \frac{1}{2c} u(\mathbf{x}) u(\mathbf{y})$, $w_{2,11}(\mathbf{x},\mathbf{y}) = \frac{-1}{4c} (A^{\frac{1}{2}} u)(\mathbf{x}) (A^{\frac{1}{2}} u)(\mathbf{y})$ and $w_{2,01} = w_{2,10} = 0$, so ω_2 is Hadamard.

We have

$$\begin{aligned}\omega_2(n^a n^b T^{\mathrm{ren}}_{ab}(0)) - \omega_2^{(0)}(n^a n^b T^{\mathrm{ren}}_{ab}(0)) &= n^a n^b \partial_{x_a}\partial_{y_b} w_2(0,0) \\ &= (n^0)^2 \partial_{x_0}\partial_{y_0} w_2(0) \\ &= -\frac{1}{4c}(n^0)^2((A^{\frac{1}{2}}u)(0)) < 0\,. \qquad (3.11)\end{aligned}$$

because the spatial derivatives vanish due to $\nabla u(0) = 0$.

For any $\lambda > 0$ we repeat the same argument for $u_\lambda(\mathbf{x}) := \lambda^{\frac{d}{2}-1}u(\lambda\mathbf{x})$ $\widehat{u_\lambda}(p) = \lambda^{-\frac{d}{2}}\hat{u}\left(\frac{p}{\lambda}\right)$. Instead of c we have

$$\begin{aligned}c_\lambda &:= \|A^{\frac{1}{4}}u_\lambda\|^2 \\ &= \frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}} \sqrt{|p|^2+m^2}|\widehat{u_\lambda}(p)|^2 \mathrm{d}^{\,d-1}p \\ &= \frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}} \sqrt{|p|^2+\lambda^{-2}m^2}|\hat{u}(p)|^2 \mathrm{d}^{\,d-1}p\,,\end{aligned}$$

so that c_λ converges to $c' \neq 0$ as $\lambda \to \infty$. Furthermore

$$\begin{aligned}(A^{\frac{1}{2}}u_\lambda)(0) &= \frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}} \sqrt{|p|^2+m^2}\widehat{u_\lambda}(p)\mathrm{d}^{\,d-1}p \\ &= \frac{\lambda^{\frac{d}{2}}}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}} \sqrt{|p|^2+\lambda^{-2}m^2}\hat{u}(p)\mathrm{d}^{\,d-1}p\,,\end{aligned}$$

so $\lambda^{-\frac{d}{2}}(A^{\frac{1}{2}}u_\lambda)(0)$ converges to $C \neq 0$ as $\lambda \to \infty$. Hence

$$\lim_{\lambda\to\infty} -\lambda^{-d}\frac{1}{4c_\lambda}(A^{\frac{1}{2}}u_\lambda)(0)^2 = -\frac{C^2}{4c'}\,.$$

It then follows from (3.11) that the violation of the NEC can be arbitrarily large as $\lambda \to \infty$. □

3.3.2 Quantum energy inequalities

Recall that the classical energy conditions were not only useful assumptions to prove interesting theorems in GR about singularities or cosmology. They are also natural attempts to try and express the stability of matter and/or the attractive nature of gravity. The fact that these conditions fail for QFTs could have significant consequences for their stability, unless there is some other property to replace them. Indeed, Ford [16] has argued that macroscopic violations of the second law of thermodynamics can be avoided if QFTs satisfy some form of quantum energy inequality. Roughly speaking, these *quantum energy inequalities* state that violations of a classical energy condition can occur, but they cannot be too negative for too long. Like the Heisenberg uncertainty principle in quantum mechanics, one expects that such bounds can be derived from the existing framework, including in particular the CCR. We will now show that this is indeed the case, at least for minimally coupled free scalar fields. In fact, we will present the following even stronger result.

Theorem 3.3.2. Let $M = (\mathcal{M}, g_{ab})$ be a globally hyperbolic Lorentzian manifold, $h, f \in C_0^\infty(\mathcal{M}, \mathbb{R})$ such that $f \equiv 1$ on $\mathrm{supp}(h)$ and t^a be a smooth timelike vector field on M. Then there exist $C, c > 0$ such that

$$\omega_2(h,h) \le C\left(\omega(T^{\mathrm{ren}}_{ab}(t^a t^b f^2)) + c\right)$$

for all Hadamard states ω on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$.

Let us first consider the significance of this result. If ω is any Hadamard state on $\mathcal{W}_M^{(\mathcal{L}_{\mathrm{fsf}})}$, then any unit vector $\psi \in \mathcal{D}_\omega$ also defines a Hadamard state. Applying the theorem we therefore have

$$\langle \psi, \phi_\omega(h)^2 \psi \rangle \le C \langle \psi, T^{\mathrm{ren}}_{ab}(t^a t^b f^2)\psi \rangle + Cc$$

in a (hopefully) obvious notation for quadratic forms. It follows in particular that the right-hand side is ≥ 0, which is a quantum energy inequality analogous to the WEC. Furthermore, with a bit of extra work we can replace this statement about quadratic forms by an inequality of operators,

$$\phi_\omega(h)^* \phi_\omega(h) \le CE + CcI$$

where E is a self-adjoint operator on $\mathcal{H}_\omega$ that extends the quadratic form $T^{\mathrm{ren}}_{ab}(t^a t^b f^2)$ and that we may think of as a measure of the total energy in the region $\mathrm{supp}(f)$. This inequality can be rewritten as

$$(E + cI)^{-\frac{1}{2}} \phi_\omega(h)^* \phi_\omega(h) (E + cI)^{-\frac{1}{2}} \le CI$$

from which it follows that

$$\|\phi_\omega(h)(E + cI)^{-\frac{1}{2}}\| \le \sqrt{C}\,.$$

This means that the unbounded operator $\phi_\omega(h)$ can be made into a bounded operator if we multiply it by the bounded operator $(E+cI)^{-\frac{1}{2}}$, which dampens the bad high energy behaviour of $\phi_\omega(h)$. Recall from the theorem that C, c are independent of the choice of ω, so this estimate holds in the representation of any Hadamard state and it actually tells us something about the properties of the abstract algebra $\mathcal{W}_M^{(\mathcal{L}_{\rm fsf})}$ combined with the state space of Hadamard states.

3.3.3 Sketch of proof of theorem 3.3.2

To avoid some technicalities we prove the result only in d-dimensional Minkowski space $M_0 = (\mathbb{R}^d, \eta_{ab})$ with timelike vector field ∂_0 and we assume $h = 0$ for simplicity. (See [34] for a full proof).

We can then write the desired result as

$$\inf_\omega \omega(T_{00}^{\rm ren}(f^2)) > -\infty\,,$$

where the infimum is taken over all Hadamard states ω. Note that the right-hand side only depends on ω_2. Let $\omega^{(0)}$ denote the two-point distribution of the Minkowski vacuum. Then this inequality is equivalent to

$$\inf_\omega (\iota^* T_{00}^{\rm split}(\omega_2 - \omega_2^{(0)}))(f^2) > -\infty\,.$$

The strategy of the proof is as follows. Suppose that we can write the expression on the left-hand side of the last inequality as the infimum over $(\omega_2 - \omega_2^{(0)})(u)$ for a compactly supported distribution u on $\mathbb{R}^{2d}$ with the following properties: (i) $\omega_2(u)$ is well-defined (in $\mathbb{R}$) for one (and hence all) Hadamard two-point distributions and (ii) $\omega_2(u) \geq 0$ for all Hadamard two-point distributions. Then we have

$$\inf_{\omega_2} (\omega_2 - \omega_2^{(0)})(u) \geq -\omega_2^{(0)}(u) > -\infty$$

and the proof is complete. It remains to show that a suitable u exists.

It may seem tempting to try $u = u_1$ for

$$u_1(x, y) = T_{00}^{\rm split} f(x) f(y) \delta(x - y)\,,$$

which gives the correct values for $(\omega_2 - \omega_2^{(0)})(u_1)$, but unfortunately $\omega_2(u_1)$ is not defined for any Hadamard two-point distribution ω_2, because ω_2 is singular on the diagonal.

As a second attempt we can use the fact that $(\omega_2 - \omega_2^{(0)})(x,y)$ is symmetric in (x,y) to replace u_1 by a distribution u with the property

$$\frac{1}{2}(u(x,y)+u(y,x)) = T_{00}^{\mathrm{split}} f(x)f(y)\delta(x-y)\,.$$

Writing $\tilde{u}(x,y) = u(y,x)$ we then have $(\omega_2 - \omega_2^{(0)})(u) = \frac{1}{2}(\omega_2 - \omega_2^{(0)})(u+\tilde{u})$, which gives the correct result. Writing $\delta(x-y) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i(x-y)\cdot p}\mathrm{d}^d p$ (in a distributional sense) we can therefore try $u = u_2$ with

$$u_2(x,y) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}_{<0}\times\mathbb{R}^{d-1}} T_{00}^{\mathrm{split}} f(x)f(y)e^{i(x-y)\cdot p}\mathrm{d}^d p\,,$$

where we restricted the integration region to $p_0 < 0$. We then have $\frac{1}{2}(u_2 + \tilde{u}_2) = u_1$ and we removed half of the singularities of $\delta(x-y)$. In particular, we removed the singularities at future pointing null vectors, which could clash with the singularities of Hadamard states. We will show that $u = u_2$ does the job.[2]

We use the fact that $\omega_2(f_1, f_2) = \langle \Phi_1(f_1), \Phi_1(f_2)\rangle_{\mathcal{H}_\omega}$ for any $f_1, f_2 \in C_0^\infty(\mathbb{R}^d, \mathbb{R})$. From the Hadamard condition we can deduce that for each $N \in \mathbb{N}$ there exists a $C_N > 0$ such that

$$\|\Phi_1(fe^{-ip\cdot\cdot})\|^2 \le C_N(1+|p|^2)^{-N} \tag{3.12}$$

when $k \in \mathbb{R}^d$ with $k_0 < 0$. (When k gets close to a future pointing null direction this estimate might fail). Now note that we can write

$$\omega_2(u_2) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}_{<0}\times\mathbb{R}^{d-1}} m^2\|\Phi_1(fe^{-ip\cdot\cdot})\|^2 + \sum_{\mu=0}^{d-1}\|\Phi_1(\partial_\mu fe^{-ip\cdot\cdot})\|^2 \mathrm{d}^d p \ge 0\,.$$

To see that the result is finite we use the estimate (3.12) with $N = d+2$, together with $\partial_\mu(fe^{-ip\cdot\cdot}) = ((\partial_\mu f) - ip_\mu f)e^{-ip\cdot\cdot}$ and the triangle inequalitiy to find a $C' > 0$ such that

$$\omega_2(u_2) \le C'\int_{\mathbb{R}_{<0}\times\mathbb{R}^{d-1}} (m^2 + 2d + 2|p|^2)(1+|p|^2)^{-d-2}\mathrm{d}^d p < \infty\,.$$

[2] If $h \neq 0$ it doesn't do the job, but we can take instead $u(x,y) = 2C\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} T_{00}^{\mathrm{split}} f(x)f(y)e^{i(x-y)\cdot p}\hat{c}(p_0)\mathrm{d}^d p - h(x)h(y)$, where $\hat{c}(p_0)$ takes values in $(0,1)$, falls off sufficiently fast when $p_0 > 0$ and satisfies $\hat{c}(-p_0) + \hat{c}(p_0) = 1$. We can then choose C large enough to make u of positive type and complete the argument.

Bibliography

[1] *Quantum Field Theory on Curved Spacetimes*,
eds. C. Bär and K. Fredenhagen, Springer, Heidelberg, Lecture Notes in Physics Vol. 786 (2009)

[2] A.N. Bernal and M. Sánchez, *On Smooth Cauchy Hypersurfaces and Geroch's Splitting Theorem*,
Commun. Math. Phys. **243**, 461–470 (2003)

[3] A.N. Bernal and M. Sánchez, *Further Results on the Smoothability of Cauchy Hypersurfaces and Cauchy Time Functions*,
Lett. Math. Phys. **77**, 183–197 (2006)

[4] O. Bratteli and D.W. Robinson, *Operator algebras and quantum statistical mechanics*,
Vol. 2, Springer, Berlin, Heidelberg, New York (2002)

[5] R. Brunetti, M. Dütsch, K. Fredenhagen and K. Rejzner, *The unitary master Ward identity: time slice axiom, Noether's theorem and anomalies*,
Ann. Henri Poincaré **24**, 469–539 (2023)

[6] 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)

[7] D. Buchholz and K. Fredenhagen, *A C^*-algebraic Approach to Interacting Quantum Field Theories*,
Commun. Math. Phys. **377**, 947–969 (2020)

[8] D. Buchholz and K. Fredenhagen, *Dynamical C^*-algebras and kinetic perturbations*,
Ann. Henri Poincaré **22**, 1001–1033 (2021)

[9] Y. Choquet-Bruhet, *General Relativity and the Einstein Equations*,
Oxford University Press, Oxford (2009)

[10] E. Curiel, *A primer on energy conditions*,
In: D. Lehmkuhl, G. Schiemann and E. Scholz (Eds.), Towards a theory of spacetime theories,
Einstein Studies Vol.13, Birkhäuser, New York NY (2017)

[11] M. Dütsch, *From classical field theory to perturbative quantum field theory*,
Progress in Mathematical Physics, Vol.74, Birkhäuser (2019)

[12] D.E. Evans and J.T. Lewis, *Dilations of irreversible evolutions in algebraic quantum theory*,
Comm. of the DIAS, Series A: Theor. Phys., **24**, Dublin (1977)

[13] C.J. Fewster, *Energy inequalities in quantum field theory*,
in: World Scientific, XIVth ICMP (proceedings), 559–568 (2006)

[14] C.J. Fewster and R. Verch, *Algebraic Quantum Field Theory in Curved Spacetimes*,
In: R. Brunetti, C: Dappiaggi, K. Fredenhagen and J. Yngvason (eds), Advances in Algebraic Quantum Field Theory,
Mathematical Physics Studies, Springer, Cham (2015)

[15] C.J. Fewster, *Quantum Energy Inequalities*,
In F. Lobo (eds): Wormholes, Warp Drives and Energy Conditions,
Fundamental Theories of Physics, Vol.189, Springer, Cham (2017)

[16] L.H. Ford, *Quantum coherence effects and the second law of thermodynamics*,
Proc. Roy. Soc. Lond., A**364**, 227–236 (1978).

[17] D. Giulini, *Remarks on the notions of general covariance and background independence*,
In: IO. Stamatescu, E. Seiler (eds), Approaches to fundamental physics,
Lecture Notes in Physics, Vol.**271**, Springer, Berlin, Heidelberg (2007)

[18] R. Haag, *Local quantum physics*,
Springer Verlag, Berlin (1996)

[19] S.W. Hawking and G.F.R. Ellis, *The large scale structure of space-time*,
Cambridge University Press, Cambridge (1973)

[20] S. Hollands and R.M. Wald, *Conservation of the stress tensor in interacting quantum field theory in curved spacetimes*,
Rev. Math. Phys. **17**, 227–312 (2005)

[21] L. Hörmander, *The analysis of linear partial differential operators*,
Vol. I, Springer, Berlin Heidelberg (1990)

[22] R.V. Kadison and J.R. Ringrose, *Fundamentals of the theory of operator algebras*,
Vol.I, AMS (1997)

[23] E.-A. Kontou and K. Sanders, *Energy conditions in general relativity and quantum field theory*,
Topical Review, Class. Quantum Grav. **37**, 193001 (2020)

[24] G. Lechner and K. Sanders, *Modular nuclearity: a generally covariant perspective*,
Axioms 5(1), 5 (2016)

[25] S. Mac Lane, *Categories for the working mathematician*,
Springer, New York (1998)

[26] O. Müller and M. Sánchez, *Lorentzian manifolds isometrically embeddable in* $\mathbb{L}^N$,
Trans. Amer. Math. Soc. **363**, 5367–5379 (2011)

[27] B. O'Neill, *Semi-Riemannien Geometry*,
Academic Press, San Diego (1983)

[28] R. Penrose, *Gravitational collapse and space-time singularities*,
Phys. Rev. Lett. **14**, 57–59 (1965)

[29] M. Reed and B. Simon, *Methods of modern mathematical physics*,
Vol.I, Academic Press, San Diego (1980)

[30] H. Ringström, *The Cauchy Problem in General Relativity*,
ESI Lectures in Mathematics and Physics, EMS (2009)

[31] K. Sanders, *A note on spacelike and timelike compactness*,
Class. Quantum Grav. **30**, 115014 (2013)

[32] K. Sanders, C. Dappiaggi and T.-P. Hack, *Electromagnetism, local covariance, the Aharonov-Bohm effect and Gauss' law*,
Commun. Math. Phys. **328**, 625–667 (2014)

[33] K. Sanders, *What can (mathematical) categories tell us about spacetime?*,
In: N. Hugget, K. Matsubara and C. Wüthrich (eds), Beyond Spacetime: The Foundations of Quantum Gravity,
pp. 338–357, Cambridge University Press, Cambridge (2020)

[34] K. Sanders, *Stress tensor bounds on quantum fields*,
Commun. Math. Phys. **405**, 132 (2024)

[35] A. Strohmaier, R. Verch and M. Wollenberg, *Microlocal analysis of quantum fields on curved space-times: Analytic wave front sets and Reeh-Schlieder theorems*,
J. Math. Phys. **43**, 5514–5530 (2002)

[36] R.M. Wald, *Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant*,
Phys. Rev. D **28**, 2118–2120 (1983)

[37] R.M. Wald, *General Relativity*,
University of Chicago Press, Chicago (1984)

[38] R.M. Wald, *Quantum Field Theory in Curved Spacetimes and Black Hole Thermodynamics*,
Chicago Lectures in Physics, University of Chicago Press, Chicago (1994)

[39] D. Yau, *Homotopical Quantum Field Theory*,
World Scientific Publishing Co., Singapore (2020)