Ludwig-Maximilians-Universität München
Technische Universität München

Master Thesis in
Theoretical and Mathematical Physics

The necessity of indefinite metric Hilbert spaces in covariant gauge formulations of QED

Advisor
Dr. habil. Wojciech Dybalski

Author
Jan Mandrysch

Second Referee
Prof. Dr. Dirk-André Deckert

Submitted on 21st of May 2019.

Abstract

Axiomatic formulations of quantum electrodynamics (QED) depart from standard QFTs in various regards. One of them is the presence of Krein spaces (i.e., indefinite metric 'Hilbert spaces') in covariant gauges. While the necessity of such spaces is often claimed, it is difficult to find a satisfactory justification in the literature, especially beyond the Gupta-Bleuler gauge or in the presence of interaction. The aim of this thesis is to provide a systematic treatment of these matters in terms of two no-go theorems. Firstly, we will show that a free hermitian covariant vector field A on a state space with a non-negative metric and a (not necessarily unique) vacuum state will give rise to a vanishing two-point function of its exterior derivative $F_{\mu\nu} = \partial_{[\mu} A_{\nu]}$. Secondly, we will infer from this result that F will not be able to generate massless states from the vacuum even if we let off the assumption of A to be free. In addition to the proofs, the reader will find a broad preliminary section and a brief comparison of the results with existing literature.

Contents

Structure of the document

This document consists of five main sections. The first section is an introduction and is devoted to describing the general setup for the proofs such that the to-be-proven statements can be formulated. For rigorous definitions and explanations concerning notation, motivation, etc. the reader is referred to Section 2. In Section 2 the mathematical and physical preliminaries are given to make the document self-contained in terms of mathematical definitions and physical interpretations. In Section 3 the proof of the first statement which was mentioned in the abstract will be worked out in detail. At the end of the section there will be a brief comparison with existing literature. Section 4 contains the implications for the interacting case of QED. This will mainly entail the formulation and proof of the second statement which was mentioned in the abstract. Eventually, Section 5 contains a summary of the obtained results and a look beyond the scope of this document. In the Appendix there is additional side material which complements the main part of the document.

Acknowledgements

Let me take here the pleasant opportunity to be thankful for the support which I received during the work on this thesis. I would like to thank my advisor Wojciech Dybalski for providing many good ideas, by taking time for helpful discussions, and by bringing to attention details which I may have missed. In this regard, my student colleague Maximilian Heinz Ruep deserves special mentioning. In office or while having lunch he always found time to discuss general concepts or to work through the details of a proof with me. Finally, in times with little progress there was encouraging support from my family, my friends, and my beloved Hanna, and I am glad for that.

Chapter 1

Introduction

The aim of this chapter is to introduce the bare necessities in order to formulate the to-be-proven statements in a rigorous way and for a reader who is already familiar with the Wightman framework of quantum field theory (QFT). For the detailed introduction, the definitions, and the notation, the reader is referred to the next section. Nonetheless it is intended to provide a brief overview on the topic.

An important framework which includes indefinite metric Hilbert spaces is given by the Strocchi-Wightman setting of QFT. This setting extends the Wightman framework of QFT to a QFT on indefinite metric Hilbert spaces, also called Krein spaces. The probability interpretation, which needs a positive definite scalar product, is recovered by restricting to an appropriate (physical) subspace. In this thesis it will be shown that it is indeed necessary to adopt such a generalization of the Wightman framework for any formulation of QED in terms of a hermitian covariant gauge field. The proof will take the form of two no-go theorems. The procedure is to require positive semi-definiteness of scalar product and to derive that this requirement leads to a trivial theory in the case of free QED and to a non-satisfactory theory in the case of interacting QED.

The a-priori indefinite scalar product (or hermitian non-degenerate sesquilinear form) on the Krein space will be denoted by $\langle \cdot, \cdot \rangle$. To state the to-be-proven-result lets us regard QED to be

Definition 1.1 (QED).
QED is a Strocchi-Wightman QFT $(\mathcal{H}, \langle \cdot, \cdot \rangle, U, \Omega, D, \{F, J\})$[1] including the hermitian Strocchi-Wightman fields F (antisymmetric 2-tensor), J (vector) such that the physical subspace $\mathcal{H}'$ satisfies

> **(Ex. and Inv. of the domain)**
> There exists a dense domain $D' \subset \mathcal{H}'$ which is invariant under $F^{\mu\nu}(u)$, $J^{\nu}(u)$ and $U(g)$ for arb. $u \in \mathcal{S}(\mathbb{M})$ and $g \in \mathcal{P}$ (Poincare or Poincare spinor group).

[1]The notation and definition of Strocchi-Wightman QFT is introduced in the preliminaries (see in particular Definition 2.51). Nonetheless let us explain that $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ denotes a Krein space where $\langle \cdot, \cdot \rangle$ denotes the indefinite scalar product, that U is a unitary representation of the Poincaré (spinor) group such that $(\mathcal{H}, \langle \cdot, \cdot \rangle, U, \Omega)$ forms an indefinite metric relativistic quantum theory satisfying the spectrum condition (on the induced physical Hilbert space) and with a vacuum vector $\Omega \in D$ (see the paragraphs below Definition 2.48) and that D is a common domain of the Strocchi-Wightman fields which is dense in $\mathcal{H}$.

(Eq. of motion)
For arb. $\Psi_1 \in \mathcal{H}'$ and $\Psi_2 \in D'$

$$\langle \Psi_1, (\partial_\mu F^{\mu\nu} - J^\nu)(u)\Psi_2 \rangle = 0 \quad \text{and} \quad \langle \Psi_1, \epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma}(u)\Psi_2 \rangle = 0, \quad u \in \mathcal{S}(\mathbb{M}). \tag{1.1}$$

The theory is referred to be free or non-interacting if, and only if, $J^\nu = 0$. The interacting case is usually expected to have J^ν being the Dirac current, but we will not have to deal with the specific structure of J. The theory is referred to be trivial as soon as the two-point function corresponding to $F_{\mu\nu}$ vanishes, i.e., $\langle \Omega, F_{\mu\nu}(x)F_{\rho\sigma}(y)\Omega \rangle = 0$.

In the case of classical field theory, it turns out that the theory is also describable in terms of a vector field A_μ satisfying $F_{\mu\nu} = \partial_{[\mu}A_{\nu]}$, thereby automatically solving the homogeneous Maxwell equation $\partial_{[\mu}F_{\nu\rho]} = 0$, where the $[\cdot]$ embraces antisymmetrized indices. In the case of free electrodynamics ($J^\nu = \partial_\mu F^{\mu\nu} = 0$) both descriptions, in terms of F, and in terms of A, are viable and lead to equivalent results. In the interacting case the situation is more difficult. The Dirac current $J^\mu = \bar{\psi}\gamma^\mu\psi$ for a complex spin 1/2 field ψ representing electrons and positrons, leads to field equations that are hard to express in terms of only $\psi, \bar{\psi}$, and $F_{\mu\nu}$ (see [Ste00, Chapter 3, p. 21]).

In the free quantum case (free QED) the situation is like the classical case. We can write down the field equations in terms of F or A and obtain equivalent results. For the interacting quantum case, however, it becomes inevitable to make use of the vector potential A_μ. Apart from some lower dimensional models the usual 'quantization' procedure starting from field equations, or rather a Lagrangian, is limited to Lagrangians with only the fields and its first derivatives appearing and which are separated into a sum of kinetic and interaction terms (see [Ste00, Chapter 3, p. 24]). For the complicated equations obtained when written only in terms of $F_{\mu\nu}$ one may say that there is no such interaction Lagrangian. Moreover, there are arguments indicating that $F_{\mu\nu}$ cannot mediate soft photon absorption and creation (see [Str70, Section II]).

With these things said, it should be our aim now to find a viable quantum field theory with a fundamental field A_μ and observable fields $F_{\mu\nu}$ and j_μ obeying the relation $F_{\mu\nu}(A) = \partial_{[\mu}A_{\nu]}$ and fulfilling Maxwell's equations. Such a formulation is referred to as a gauge formulation of QED. From now on we will only deal with gauge formulations of QED.

In classical field theory the relation between F and A determines A only up to a gauge transformation

$$A_\mu(x) \mapsto A'_\mu(x) = A_\mu(x) + \partial_\mu G(x) \tag{1.2}$$

for an arb. twice-differentiable function G. In quantum field theory from a mathematical point of view it is hard to control such gauge transformations. It is much more promising to choose a specific representative of A out of the equivalence class of different A's and to develop the theory according to the choice of representative. There are many such choices available and it is natural to make the most convenient choice. There is a problem, however, with QED that no matter what the choice is one must leave the framework of QFTs that we are comfortable with. The possibilities are either to insist on a state space consisting of only physical states or to insist on locality or covariance of the gauge field. Note that the choice is exclusive here, that is, when one property is chosen the other property is lost. The prototypes of the two possibilities are the Coulomb and the Gupta-Bleuler gauge, respectively.

What we will show here is that when one takes the path of covariance, and in particular the path of the Gupta-Bleuler-approach (a brief explanation will be found below), one will necessarily deal with a quantum field theory on an indefinite metric Hilbert space. In the physics literature this is usually termed as ghosts appearing in the theory. It is a consequence of the fact that the state space will entail not only physical states, but also non-physical states, that is states on which the equations of motion do not hold and which will have a negative 'norm'.[2]

That one has to introduce such non-physical states is a consequence of no-go-theorems by Strocchi (see [Str67, Str70] and [Str13, Chapter 7.8]). In particular, when one requires a vector-operator-valued distribution A_μ to satisfy translation-invariance and either Lorentz-covariance or locality and temperedness and one has a vacuum state in order to define the two-point function of A_μ, then the Maxwell equations will imply the two-point function of the associated $F_{\mu\nu}(A)$ to vanish.

The usual way to construct a local or covariant gauge formulation of QED is therefore to take modified equations of motion for F and A, respectively, and to require the non-physical modification terms to vanish on expectation values between physical states. Explicitly, we demand

$$\partial^\mu F_{\mu\nu} - \mathcal{L}_\nu = \Box A_\nu - \partial_\nu \partial^\mu A_\mu - \mathcal{L}_\nu = 0 \tag{1.3}$$

to be satisfied as operator equations (i.e., on the whole of $\mathcal{H}$) and the deformation term $\mathcal{L}_\nu$ to vanish when sandwiched between physical states ($\langle \phi, \mathcal{L}_\nu \psi \rangle = 0$ for $\phi, \psi \in \mathcal{H}'$). In this way the first Maxwell equation is recovered on expectation values within physical states. The choice of $\mathcal{L}_\nu$ determines A partially or completely and can be thought of as part of the gauge choice. From the Lagrangian viewpoint $\mathcal{L}_\mu$ corresponds to the gauge-fixing term which is added via a Lagrange-multiplier.

Among the gauge formulations of QED (or gauge formulations of any other QFT) one may distinguish between local and non-local gauges. The gauge is referred to be *local* if, and only if, the corresponding gauge field, here A_μ, satisfies the locality condition, here

$$[A_\mu(u), A_\nu(v)] = 0 \tag{1.4}$$

where u and v are test functions on Minkowski space such that their supports are spacelike separated[3] and *non-local* if it is not. Similarly, one may distinguish between covariant

[2]The reason that one is willing to take such uncomfortable consequences as the appearance of ghost states is that the other possible choice is not better. In fact, the loss of locality and covariance of A_μ implies the loss of many methods developed in the framework of quantum field theory which heavily rely on locality and covariance. Renormalization methods fall into this class.

[3]i.e., $(x-y)^2 < 0 \; \forall x \in \text{supp } u, y \in \text{supp } v$

and non-covariant gauges. The gauge is referred to be *covariant* if, and only if, the corresponding gauge field, here A_μ, satisfies the covariance condition, here

$$U(\underset{\sim}{\Lambda}, a) A_\mu(u) U(\underset{\sim}{\Lambda}, a)^{-1} = \Lambda(\underset{\sim}{\Lambda})_\mu{}^\nu A_\nu(u_{\Lambda(\underset{\sim}{\Lambda}),a}), \qquad (\underset{\sim}{\Lambda}, a) \in \tilde{\mathcal{P}}_+^\uparrow \tag{1.5}$$

for test functions u with $u_{\Lambda,a}(x) \equiv u(\Lambda^{-1}(x-a))$ and where U is a unitary representation of the Poincaré spinor group $\tilde{\mathcal{P}}_+^\uparrow$.

Important to note is that this distinction is not necessary for a formulation without gauge fields as such fields are physical thus observable, and we always want observable fields to be local and covariant.[4]

The standard example for a local and covariant gauge choice would be the Gupta-Bleuler formalism with $\mathcal{L}_\mu = (\lambda - 1)\partial_\mu \partial A$ corresponding to equations of motion $\Box A_\mu - \lambda \partial_\mu \partial A = 0$ on the whole of $\mathcal{H}$. We usually exclude $\lambda = 1$.[5] The physical states are then determined to be elements of the kernel of $(\partial A)^{(-)}$. Here $(\cdot)^{(\pm)}$ denotes the positive/negative energy part of a free field.[6] The condition that $(\partial A)^{(-)}\psi$ vanishes for a physical state ψ is called Gupta-Bleuler(GB)-subsidiary condition. This choice of gauge corresponds to the classical Lorentz gauge as it implies $\partial A \equiv \partial_\mu A^\mu$ to vanish on expectation values within physical states. This clearly implies that the same is true for $\mathcal{L}_\mu$, too.

It is now time to state the results of this thesis. It should be noted that the results themselves are already present in the literature, but without explicit derivations and the implications for the interacting case are usually not drawn. For a thorough comparison with existing literature the reader is referred to the discussion of the results in the subsections 3.5 and 4.3. The first result focusses on free QED. The statement is

Theorem 1.2. *Every covariant gauge formulation of free QED on a state space with a non-negative metric*[7] *is trivial.*

The triviality which is referred to here is the vanishing of the two-point function of the Maxwell tensor F. One may connect free QED also to the interacting case of QED by looking at massless states. On these states there is a free time evolution and the space of single-photon states is expected to lie within the space of massless states. The result however is:

Theorem 1.3. *In every covariant gauge formulation of QED on a state space with a non-negative metric the Maxwell-tensor F cannot create massless states from the vacuum. With $\mathcal{H}^{(1)}$ being the space of massless states we have*

$$\langle \mathcal{H}^{(1)}, F_{\mu\nu}(\cdot)\Omega \rangle = 0. \tag{1.6}$$

[4] Otherwise we would have observable contradictions to Einstein-locality and covariance.

[5] In the case $\lambda = 1$ the equations of motion are unmodified and Maxwell's equations hold on the whole of $\mathcal{H}$. Later (see Theorem 2.52) we will see that a covariant or local hermitian vector field A which satisfies Maxwell's equations will give rise to a trivial theory. Thus the exclusion of $\lambda = 1$.

[6] For free fields there is an explicit construction available that enables the splitting of the field into positive and negative energy part. That ∂A is a free field (in free QED) is a consequence of the equations of motion $\Box \partial A = \partial^\mu \Box A_\mu = -\lambda \partial^\mu \partial_\mu \partial A = \lambda \Box \partial A$ which imply $\Box \partial A = 0$ for $\lambda \neq 1$.

[7] i.e., a linear space equipped with a hermitian sesquilinear form with non-negative scalar square.

One could summarize both statements and obtain:

With the usual interpretation of F in QED, a satisfactory covariant gauge formulation of QED must necessarily be defined on a state space with an indefinite metric. The state space must necessarily contain states with negative 'norm'.

For a discussion of the results the reader is referred to the subsections 3.5 for the case of free QED and 4.3 for the case of interacting QED. A final discussion is contained in the summary and outlook of the document in Section 5.

1.1 Outline of the proof

From the mathematical side the Theorems 1.2 and 1.3 are proven as corollaries of the following mathematical result

Lemma 1.4. *Let W be a Lorentz-covariant tempered distribution on Minkowski space $\mathbb{M}$ which is*

(a) transforming as a Lorentz-2-tensor, i.e., written in components $W_{\mu\nu}$ one has

$$W_{\mu\nu}(u) = \Lambda_\mu^\rho \Lambda_\nu^\sigma W_{\rho\sigma}(u_{\Lambda,0}), \quad \Lambda \in \mathcal{L}_+^\uparrow, u \in \mathcal{S}(\mathbb{M}). \tag{1.7}$$

(b) fulfilling the partial differential equation

$$\left(\eta_\nu{}^\rho \square - \lambda \partial_\nu \partial^\rho\right) W_{\mu\rho} = 0, \quad \lambda \neq 1 \tag{1.8}$$

Then requiring that $W_{\mu\mu}$ is non-negative, i.e.,

$$\hat{W}_{\mu\mu}(u) \geq 0, \quad \textit{for } u \in \mathcal{S}(\mathbb{M}) \textit{ with } u \geq 0 \tag{1.9}$$

implies that there exists a tempered distribution G on $\mathbb{M}$ such that

$$W_{\mu\nu} = \partial_\mu \partial_\nu G. \tag{1.10}$$

In particular one obtains

$$G(u) = \int \left(a\hat{u}(|\vec{p}|, \vec{p}) + b\hat{u}(-|\vec{p}|, \vec{p})\right) \frac{d\vec{p}}{2|\vec{p}|}, \quad u \in \mathcal{S}(\mathbb{M}) \tag{1.11}$$

for some constants a and b, and where $\hat{f}(p) = \int d^4x f(x) e^{ipx}$ denotes the Fourier transform of f (px denoting the Minkowski product).

The role of the tempered distribution W with its properties is played by the (translation-invariant) Wightman two-point function $W_{\mu\nu}(x-y) = \langle \Omega, A_\mu(x) A_\nu(y) \Omega \rangle$ of the free(!) vector potential A. Ω denotes the vacuum vector of the theory. One can then derive that the two-point function of F is linear in W and is such that derivative terms like in the resulting eq. (1.10) do not contribute. Hence the two-point function of F vanishes.

For the interacting case one introduces the orthogonal projection operator $P^{(1)}$ onto the space of massless states. With the definition of a modified two-point function

$$W^{(1)}_{\mu\nu}(x-y) \equiv \langle \Omega, A_\mu(x) P^{(1)} A_\nu(y)\Omega\rangle \tag{1.12}$$

which has its support restricted to the massless spectrum one obtains again a tensor-valued tempered distribution satisfying the same properties as required for Lemma 1.4. This will yield to a vanishing two-point function $\langle \Omega, F_{\mu\nu}(x)P^{(1)}F_{\rho\sigma}(y)\Omega\rangle$ and thus imply that Theorem 1.3 holds.

A brief outline of the key steps to the proof of Lemma 1.4 above should be given in order to clarify the general concept of it:

Step 1 (Covariant structure of W)
By the transformation behaviour as a Lorentz two-tensor there exists a decomposition

$$W_{\mu\nu} = \eta_{\mu\nu}K + \partial_\mu\partial_\nu G \tag{1.13}$$

where K and G are Lorentz-invariant distributions.

Step 2 (Precise shape of W by equation of motion)
The general solution to the equations of motion

$$\left(\eta_\nu{}^\rho \,\Box - \lambda\partial_\nu\partial^\rho\right) W_{\mu\rho} = 0, \tag{1.14}$$

including also Lorentz-covariance and the spectral condition, is

$$\begin{aligned} W_{\mu\nu} = c_1 &\left(\eta_{\mu\nu}D^{(+)} - \frac{\lambda}{1-\lambda}\partial_\mu\partial_\nu x^2 D^{(+)}\right) - c_2\partial_\mu\partial_\nu D^{(+)} \\ &- c_3\left(\eta_{\mu\nu}x^2 - \frac{1}{24}\frac{4-\lambda}{1-\lambda}\partial_\mu\partial_\nu (x^2)^2\right) + c_4\eta_{\mu\nu}. \end{aligned} \tag{1.15}$$

for constants $c_1, ..., c_4$ and where $D^{(+)}$ is formally defined by its Fourier transform $\hat{D}^{(+)}(p) = \theta(p_0)\delta(p^2)$.

Step 3 (Final argument)
Non-negativity of the scalar square of $\langle\cdot,\cdot\rangle$ requires

$$\hat{W}_{\mu\mu}(|u|^2) \geq 0 \tag{1.16}$$

for arb. $u \in \mathcal{S}(\mathbb{M})$. As the summands of $W_{\mu\nu}$ are homogeneous generalized functions with different degrees of homogeneity (scaling degrees) it is possible to infer non-negativity of the summands with different degrees, separately. As the $\eta_{\mu\nu}$ terms show different signs for $\mu = \nu = 0$ and $\mu = \nu = 1, 2, 3$ we obtain $c_1 = c_3 = c_4 = 0$ and $c_2 \geq 0$. Thus

$$W_{\mu\nu} = -c_2\partial_\mu\partial_\nu D^{(+)}. \tag{1.17}$$

Chapter 2

Preliminaries

This section is devoted to enlightening and untying the construct of definitions underlying this thesis. It will introduce some basic notation, the calculus of generalized functions, specific topics of representation theory focussing on Lorentz and Poincare group representations, the Wightman and the Strocchi-Wightman frameworks of quantum field theory as well as a working definition for quantum electrodynamics. The preliminaries aim to be self-contained although for many proofs there is a reference to either the appendix or the literature. For more detailed information on where the contents of the preliminary section generally stem from (unless noted otherwise) the reader is referred to the following footnote.[1]

2.1 Basic notation and conventions

This subsection sets up the general conventions and specialties of this document. What is written here is applied throughout the document unless stated otherwise.
In this document we will denote n-dimensional real and complex Euclidean space by $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively. Four-dimensional real and complex Minkowski space will be denoted by $\mathbb{M}$ and $\mathbb{CM}$, respectively. Their metric tensors will be denoted by δ (Kronecker-Delta) and η (Minkowski-tensor), respectively. Written out in components they are defined by

$$\delta_{11} = ... = \delta_{nn} = +1, \delta_{ij} = 0 \text{ for } i \neq j \tag{2.1}$$

and (choosing mostly minus convention)

$$\eta_{00} = -\eta_{11} = -\eta_{22} = -\eta_{33} = +1, \eta_{\mu\nu} = 0 \text{ for } \mu \neq \nu. \tag{2.2}$$

[1]Sections 2.2, 2.3, and 2.4 are mainly based on the chapters 2 and 3 of [BLOT90]. Whereas the first two sections are presenting existing contents in a new concise format, in Section 2.4 also the contents were created by the author as an application to the preceding two sections. An alternative presentation following [BLOT90, Chapter 4.2] is given in Appendix B. Some more specialized information for the named three sections is also drawn from the chapters III.3 and V.3 of [RS80], and IX (esp. IX.1, IX.2, and appendix) of [RS75]. Section 2.5 is mainly based on chapters 4, 5 and 6.1 of [Sch12] as well as sections VIII.3 and VIII.4 of [RS80]. Section 2.6 is a conglomerate of many sources. Most notably are [SW80], [Dyb18], [BLOT90], and [Str13] although the presentation is partially different from the sources. For the last section a similar situation applies although the most important sources were [Ste00] and again [Str13].

Here we also choose the convention to denote Euclidean indices by small Latin letters $i, j, k, ...$ running through $1, 2, 3$ or $1, ..., n$ and Minkowskian/Lorentzian indices by small Greek letters $\mu, \nu, \rho, ...$ running through $0, 1, 2, 3$. In order to raise and lower indices the Minkowski tensor may be applied

$$x_\mu = \eta_{\mu\nu} x^\nu \text{ and } x^\nu = x_\mu \eta^{\mu\nu}, \tag{2.3}$$

where $\eta^{\mu\nu} \equiv \eta^{-1}_{\mu\nu} = \eta_{\mu\nu}$. Here and throughout the document summation over repeated upper and lower indices is understood. It will not be summed over equal indices that are all lower or all upper. In other words, these indices are meant to be fixed.

As indices of many types will occur which may complicate readability we will sometimes write expressions like

$$w \equiv w_l^{(\kappa)} \tag{2.4}$$

meaning that the following descriptions will assume fixed κ and l and keep them implicit where not needed. In general, the symbol '$\equiv$' will represent the definition relation in this document. So $A \equiv B$ means A is defined to be B.

Many times, also multi-index notation is used. This means to summarize a set of indices $\alpha_1, ..., \alpha_m$ for $m \in \mathbb{N}$, representing it by a single symbol α (it would be written $\alpha \equiv (\alpha_1, ..., \alpha_m)$). We then define

$$|\alpha| \equiv \alpha_1 + ... + \alpha_m \quad \text{and} \quad \alpha! \equiv \alpha_1! \cdot ... \cdot \alpha_m!. \tag{2.5}$$

as well as

$$x^\alpha = x_1^{\alpha_1} ... x_m^{\alpha_m} \qquad \text{and} \qquad D^\alpha = \partial_1^{\alpha_1} ... \partial_m^{\alpha_m} \equiv \frac{\partial}{\partial(x_1)_{\alpha_1}} ... \frac{\partial}{\partial(x_m)_{\alpha_m}}. \tag{2.6}$$

Finally, constants, functions and generalized functions always take values in $\mathbb{C}$ whenever this is not further specified. Sometimes this is later specified to be only in the reals or non-negative reals.

2.2 Calculus of generalized functions aka tempered distributions

In the Wightman picture of quantum field theory the fundamental objects are the quantum fields, which are operator-valued tempered distributions with certain additional properties. Therefore, the calculus of tempered distributions or - how we will also call them - generalized functions underlies Wightman theory and is of great importance to it. This subsection is devoted to introducing the calculus of generalized functions though in a strongly confined format. For proofs and other details the reader is referred to a more thorough version of preliminaries in [BLOT90, Chapter 2].

2.2.1 Tempered distributions

Let $\mathcal{S}(\mathbb{R}^n)$ denote the space of rapidly decreasing differentiable functions on $\mathbb{R}^n$, also known as Schwartz space. This space can be equipped with a countable system of seminorms

$$||u||_{l,m} = \max_{|\alpha|\leq l} \sup_{x\in\mathbb{R}^n} (1+|x|)^m |D^\alpha u(x)| \tag{2.7}$$

where $u \in \mathcal{S}(\mathbb{R}^n)$ and $|x|$ is the Euclidean norm of x on $\mathbb{R}^n$. With respect to this countable system of seminorms $\mathcal{S}(\mathbb{R}^n)$ is complete and thus becomes a separated complete locally convex space, also referred to a Frechét space. A tempered distribution is a continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$, or in other words, a linear map $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ subject to the condition:

$$\exists l, m \in \mathbb{N}, c \geq 0 : \quad |T(u)| \leq c||u||_{l,m} \quad \forall u \in \mathcal{S}(\mathbb{R}^n). \tag{2.8}$$

The space of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$ implying it to be the topological dual space of $\mathcal{S}(\mathbb{R}^n)$. There is a well-developed calculus of tempered distributions involving addition, multiplication, tensor products, transformations, differentiation, Fourier transform, convolution, etc. which is extending the calculus of differentiable functions. Therefore they are sometimes also called generalized functions. In this respect, it is extremely practical to adopt/extend the notation of functions to generalized functions. Thus, many times we will make use of writing $T(x)$ representing T with x clarifying to be the argument variable of T's test functions. We will sometimes distinguish between the terms 'generalized function' and 'tempered distribution' depending on whether we want to stress the characteristics of functions or distributions that are associated to these objects. But note that there is in principle no difference between them. We will stick to tempered distribution notation at first as it is more precise.

Another analogy of tempered distributions with ordinary functions is given when we define tempered distributions to be a limit

$$T(u) \equiv \lim_{n\to\infty} \int T_n(x)u(x)dx \tag{2.9}$$

of a certain type of sequence $(T_n)_n$ of continuous functions approximating T.[2] Equalities like $T(x) \equiv T'(x)$ will (where not stated otherwise) be understood in the framework of generalized functions aka tempered distributions, i.e., $T(u) = T'(u)\ \forall u \in \mathcal{S}$. Note that for a regular distribution[3] the ordinary sense of function equalities and the sense described here are equivalent. Hence this usage of notation is consistent.

Another practical representation of tempered distributions which makes the close connection to functions apparent is Schwartz' representation theorem. It states that every tempered distribution T (on $\mathbb{R}^n$) can be represented as

[2] This certain type of sequence, a fundamental sequence is given by a sequence $(T_n)_n$ of continuous functions T_ν for which there exists a sequence $(F_n)_n$ of polynomially bounded and sufficiently often differentiable functions F_n such that $D^\alpha F_n = T_n$ for some multiindex α and such that over each bounded set the sequence $(F_n)_n$ is uniformly convergent to a continuous function.

[3] A regular distribution is a function understood as a distribution. See e.g. eq. (2.11) for how this works.

$$T = \sum_{|\alpha| \leq N} c_\alpha D^\alpha \tag{2.10}$$

for some natural number N and continuous functions c_α of polynomial growth[4] on $\mathbb{R}^n$ understood as tempered distributions by means of

$$c_\alpha(u) = \int c_\alpha(x) u(x) d^n x. \tag{2.11}$$

This is also called a regular distribution.

Furthermore, there is the important Hahn-Banach-Theorem[5] which states that whenever we have a continuous linear functional T_0 acting on a linear subspace of $\mathcal{S}(\mathbb{R}^n)$ there exist a linear and continuous extension of T_0 to the whole of $\mathcal{S}(\mathbb{R}^n)$.

2.2.2 A selection of basic properties

In this section basic definitions to introduce the calculus of generalized functions are given. At first there is a list of basic definitions without further explanations of the concepts. These concepts should be already clear to the reader and the list is rather mentioned to fix the notation and for the sake of completeness. To simplify notation, from now on we shorten $\mathcal{S} \equiv \mathcal{S}_n \equiv \mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}' \equiv \mathcal{S}'_n \equiv \mathcal{S}'(\mathbb{R}^n)$ whenever it is clear that the domain is $\mathbb{R}^n$.

Definition 2.1. Let $T \in \mathcal{S}'_n$. Then there are the following definitions:

Restriction: For an open subset $\mathcal{O} \subset \mathbb{R}^n$ the tempered distribution $T\restriction_{\mathcal{O}} \in \mathcal{S}'(\mathcal{O})$, the restriction of T to $\mathcal{O}$, is defined by the restriction of the linear functional to the closed subspace $\mathcal{S}(\mathcal{O}) \subset \mathcal{S}$ consisting of the elements of $\mathcal{S}$ which vanish on $\mathbb{R}^n \backslash \mathcal{O}$. [6]

Support: The support of T is the unique closed set, denoted by supp T, which is the complement of the maximal open set $\mathcal{O} \subset \mathbb{R}^n$ such that

$$T(u) = 0, \quad u \in \mathcal{S}(\mathcal{O}). \tag{2.12}$$

Concentration: T is said to be concentrated at $x \in \mathbb{R}^n$ if, and only if, supp $T = \{x\}$.

Transformation: For a diffeomorphism ϕ of $\mathbb{R}^n$ onto itself the tempered distribution $T \circ \phi^{-1}$ is defined by

$$T \circ \phi^{-1}(u) \equiv T(|J(\phi)|\ u \circ \phi), \quad u \in \mathcal{S} \tag{2.13}$$

where $|J(\phi)|$ is the Jacobian determinant of ϕ. In generalized function notation we write $T(\phi^{-1}(y))$.

[4]This condition guarantees that the integral below is well-defined.

[5]The Hahn-Banach theorem states that a continuous linear functional defined on a subspace of a locally convex space can always be extended to a continuous linear functional on the whole of the locally convex space. For a version for normed spaces see for instance [BLOT90, Theorem 1.2, p. 15]

[6]Let us note that almost all the results extend to tempered distributions restricted to open subsets of $\mathbb{R}^n$. For the sake of definiteness, however, we will stick to using tempered distributions defined on all of $\mathbb{R}^n$.

Differentiation: The tempered distribution $D^\alpha T$ is defined by

$$D^\alpha T(u) \equiv (-1)^{|\alpha|} T(D^\alpha u). \tag{2.14}$$

Multiplication: For a function ϕ which is a multiplicator of $\mathcal{S}$, i.e., $\phi\mathcal{S} \subset \mathcal{S}$ the tempered distribution ϕT is defined by

$$\phi T(u) \equiv T(\phi u), \quad u \in \mathcal{S}. \tag{2.15}$$

Tensor Product: For tempered distributions $T_1 \in \mathcal{S}'_m$ and $T_2 \in \mathcal{S}'_n$ the tensor product $T_1 \otimes T_2 \in \mathcal{S}'_{m+n}$ is defined by

$$T_1 \otimes T_2\,(u_1 \otimes u_2) \equiv T_1(u_1)T_2(u_2), \quad u_1 \in \mathcal{S}_m, u_2 \in \mathcal{S}_n \tag{2.16}$$

and its unique extension to the whole of $\mathcal{S}_{m+n}$.

Vector-Valued: A vector-valued tempered distribution is a linear continuous map from $\mathcal{S} \to V$, where V is a complex locally convex space. We will sometimes denote such distributions by $\mathcal{S}'(\mathbb{R}^n, V)$. An elementary example is given by $V = \mathbb{C}^n$. A special case is a tensor-valued tempered distribution. Another important special case is when $V = \mathcal{H}$ is a Hilbert space. In this case note that for a linear map $\mathcal{S} \to \mathcal{H}$ it is equivalent whether the map is weakly or strongly continuous (see Appendix A, Proposition A.7).

Operator-Valued: Let $\mathcal{H}$ denote a Hilbert space with a dense subspace $D \subset \mathcal{H}$ and let $\mathcal{L}(D, \mathcal{H})$ denote the space of linear operators on $\mathcal{H}$ with domain D. Then an operator-valued tempered distribution is a linear map

$$A : \mathcal{S} \to \mathcal{L}(D, \mathcal{H}) \quad u \mapsto A(u) \tag{2.17}$$

where

$$u \mapsto \langle \phi, A(u)\psi \rangle \tag{2.18}$$

is required to be a tempered distribution for all $\phi, \psi \in D$ and where $A(u)$ is closable for each $u \in \mathcal{S}$. Note that weak ($\langle \phi, A(\cdot)\psi \rangle$ continuous for each $\phi \in \mathcal{H}, \psi \in D$) and strong continuity ($A(\cdot)\phi$ continuous for each $\phi \in D$) of $A(\cdot)$ can be inferred from the definition (see Appendix A, Proposition A.8).

Tensor-Operator-Valued: Let $D \subset \mathcal{H}$ be again a dense subspace of a Hilbert space $\mathcal{H}$ and let D_{mn} be a tensor-representation-map of some Lie group G corresponding to some basis of the representation space. Then a tensor operator A is a (for simplicity finite) collection of operators $A_m \in \mathcal{L}(D, \mathcal{H})$ which transform under the tensor representation by

$$A_m \mapsto A'_m = \sum_{n=1} D_{mn}(g) A_n, \quad g \in G. \tag{2.19}$$

We then define the operator $A \equiv \bigoplus_{m=1} A_m$ by

$$A\phi \equiv \bigoplus_m A_m\phi, \quad \text{or} \quad \langle \psi, A\phi \rangle \equiv \bigoplus_m \langle \psi, A_m\phi \rangle, \quad \psi, \phi \in D. \tag{2.20}$$

Proposition 2.2. *Let $T \in \mathcal{S}'$ and $u \in \mathcal{S}$ an arbitrary test function.*

(a) $D^\alpha D^\beta T = D^\beta D^\alpha T \quad \forall \alpha, \beta.$

(b) If T *is concentrated at the origin it has the form*

$$T(u) = \sum_{|\alpha| \leq N} c_\alpha D^\alpha u(0) \tag{2.21}$$

for some natural number N *and where* c_α *is a fixed finite set of constants.*

Proof. (a) holds by the commutativity of differentials acting on Schwartz space. The commutativity on Schwartz space holds because of smoothness and Schwarz'/Clairaut's theorem.
(b) see Appendix A Proposition A.1. ■

Proposition 2.3. *(The gluing principle for tempered distributions)*
Let $\{\mathcal{O}_j\}_{j=1,\dots,m}$ *finite open covering of* $\mathbb{R}^n$ *such that*

$$Q_j = \mathbb{R}^n \backslash \bigcup_{i \neq j} \mathcal{O}_i \tag{2.22}$$

is closed and contained in $\mathcal{O}_j$ *and that for each* $x \in Q_j$ *the distance* $d(x, \partial\mathcal{O}_j)$ *to the boundary* $\partial\mathcal{O}_j$ *of* $\mathcal{O}_j$ *satisfies*

$$d(x, \mathcal{O}_j) \geq A(1 + |x|)^{-\delta} \tag{2.23}$$

for fixed[7] *numbers* $A > 0$, $\delta \geq 0$. *Then for any family* $\{T_j\}_{j=1,\dots,m}$ *of tempered distributions on* $\mathbb{R}^n$ *satisfying*

$$(T_i - T_j)\restriction_{\mathcal{O}_i \cap \mathcal{O}_j} = 0 \quad \forall i, j = 1, \dots, m \tag{2.24}$$

there exists a unique tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ *coinciding with* T_j *in* $\mathcal{O}_j$ *for all* j.

Proof. See Appendix A Proposition A.4. ■

2.2.3 Division problem

With some basic structures set up we can tackle a more advanced question. The division problem (for tempered distributions) consists of a given tempered distribution T, a given function ϕ which is a multiplicator of $\mathcal{S}$, and an unknown tempered distribution S subject to the equation

$$\phi S = T. \tag{2.25}$$

For the sake of definiteness let here the variable space be $\mathbb{R}^n$. The elementary case of this problem is given whenever $\phi(x) \neq 0 \ \forall x \in \mathbb{R}^n$ and $\phi(x)$ is approaching zero not too fast as $|x| \to \infty$. In other words

Proposition 2.4. *The division problem where not only* ϕ, *but also* $1/\phi$, *is a multiplicator of* $\mathcal{S}$ *is uniquely solved by the tempered distribution*

$$S = \frac{1}{\phi} T. \tag{2.26}$$

[7]only dependent on the given covering

Proof. The tempered distribution $S = \frac{1}{\phi}T$ is well-defined, as $\frac{1}{\phi}$ is a multiplicator of $\mathcal{S}$ and clearly solves the division problem. The solution is unique as it is fixed on the subspace $\phi\mathcal{S}$ of $\mathcal{S}$ and this is already the whole space as $\mathcal{S} \supset \phi\mathcal{S} \supset \phi\frac{1}{\phi}\mathcal{S} = \mathcal{S}$. ■

For the specific case of ϕ being a polynomial function in x we can state

Theorem 2.5. *Let ϕ be a non-zero polynomial function on $\mathbb{R}^n$. Then a solution to the division problem S in the class of tempered distributions always exists.*

Proof. We will use the fact that $\phi u \to 0$ implies $u \to 0$ without further proof. For the details on this the reader is referred to the original work [Hö58, Theorem 1].
The linear map $\phi u \mapsto u$ from $\phi\mathcal{S}$ to $\mathcal{S}$ is well-defined[8] and as $\phi u \to 0$ implies $u \to 0$ the linear map $\phi u \mapsto T(u)$ is continuous. By the Hahn-Banach theorem we can extend this linear form on $\phi\mathcal{S} \subset \mathcal{S}$ to the (continuous) map $u \mapsto S(u)$ defined on the whole of $\mathcal{S}$. Hence we end up with a tempered distribution S fulfilling $\phi S(u) = S(\phi u) = T(u)$ for each $u \in \mathcal{S}$. ■

The last part of this section focusses on the very simple case of the division problem on $\mathbb{R}$ for polynomials. By using the gluing principle we can deal with the division problem for each of the zeroes of the polynomial function separately (e.g. by using compactly supported test functions with appropriate supports). For the separate division problems we can use a change of variables (a diffeomorphism of $\mathbb{R}$ into itself) to reduce the problem to

Proposition 2.6. *Let $T, S \in \mathbb{S}'(\mathbb{R})$ such that*

$$x^k S(x) = T(x) \tag{2.27}$$

for some natural number k. Then S has the general form

$$S = \sum_{j=0}^{k-1} c_j \delta^{(j)} + S_p, \tag{2.28}$$

for some constants c_j and a particular solution to the division problem $S_p \in \mathcal{S}'(\mathbb{R})$. The j-th derivative of the delta distribution is denoted by $\delta^{(j)}$.

Proof. By Theorem 2.5 a solution to the division problem given here always exists. When taking two particular solutions to the problem, their difference, denoted by S_0, has to satisfy the homogeneous equation

$$x^k S_0 = 0. \tag{2.29}$$

Therefore S_0 has to be concentrated at the origin and by Proposition 2.2(b) the general form of S_0 is

$$S_0 = \sum_{j=0}^{N} c_j \delta^{(j)} \tag{2.30}$$

[8]If there are to elements $u, v \in \mathcal{S}$ such that $\phi u = \phi v$ or, equivalently, such that $\phi(u - v) = 0$. Then this implies that $u - v(x) = 0$ for all but finitely many isolated x (the zeroes of ϕ). Because of continuity $u - v = 0$.

for some natural number N and some constants c_j. Let us note that the $\delta^{(j)}$'s are linear independent for different j and that

$$\begin{aligned} x^k\delta^{(j)}(u) &= (-1)^j\delta((x^k u)^{(j)}) \\ &= (-1)^j\sum_{i=k}^{j} u^{(j-i)}(0), \quad u\in\mathcal{S}. \end{aligned}$$

Here $(\ )^{(j)}$ denotes the j-th derivative of the term in brackets. In general, the result of the computation above is zero for $j<k$ and not zero for $j\geq k$. Therefore $N<k$ is necessary and sufficient such that S_0 is a solution to eq. (2.29). ■

What we will later use to discuss solutions of the d'Alembert equation in four dimensions is a bit more general

Proposition 2.7. *Let $T,S\in\mathcal{S}'(\mathbb{R}\times\mathbb{R}^n)$ such that*

$$x^k S(x,y) = T(x,y) \tag{2.31}$$

for some natural number k. Then S has the general form

$$S(x,y) = \sum_{j=0}^{k-1} c_j(y)\delta^{(j)}(x) + S_p(x,y) \tag{2.32}$$

for generalized functions $c_j\in\mathcal{S}'(\mathbb{R}^n)$ and some particular solution to the division problem $S_p\in\mathcal{S}'(\mathbb{R}\times\mathbb{R}^n)$.

Proof. This is a corollary of Proposition 2.6 and the fact that each $T\in\mathcal{S}'(\mathbb{R}^m\times\mathbb{R}^n)$ with $\operatorname{supp} T\subset\{a\}\times S$ for some $a\in\mathbb{R}^m, S\subset\mathbb{R}^n$ can be represented as

$$T(x,y) = \sum_{|\alpha|\leq N} D_x^\alpha\delta(x)h_\alpha(y) \tag{2.33}$$

for a (finite) family of $h_\alpha\in\mathcal{S}'(\mathbb{R}^m)$ with $\operatorname{supp} h_\alpha\subset S$. This fact is proven in Appendix A Proposition A.3. ■

2.2.4 Fourier transform, multiplication, and convolution

One of the great properties of tempered distributions is that they behave very nicely under Fourier transform. The reason for that is that the Fourier transform is an automorphism of Schwartz space and therefore induces an automorphism on its topological dual, as well.

To begin with, let us specify the definitions of the Fourier transform and the convolution for test functions (i.e., on Schwartz space):

Definition 2.8 (Fourier transform and convolution of test functions)**.**
For each $u,v\in\mathcal{S}$ define the following Schwartz functions

$$\begin{aligned} \textit{Fourier transform:}\qquad \hat{u}(p) &\equiv \int u(x)e^{-ipx}d^n x. \\ \textit{Inverse Fourier transform:}\qquad \check{u}(x) &\equiv \int u(p)e^{+ipx}d_n p. \\ \textit{Convolution: } u*v\,(x) &\equiv \int u(x)v(y-x)d^n x. \end{aligned}$$

Here $(p,x)\mapsto px$ denotes a fixed non-degenerate bilinear form on $\mathbb{R}^n$ and $d_n p\equiv(2\pi)^{-n}d^n p$ accounts for the normalization of the Fourier transform.

Remark 2.9. Later we will only deal with the Fourier transform on Minkowski space $\mathbb{M}$ which is $\mathbb{R}^4$ endowed with the non-degenerate bilinear form $px \equiv p_\mu x^\mu \equiv p_0 x^0 - \vec{p} \cdot \vec{x}$, called Minkowski product.

Let us now extend the definition of the Fourier transform to tempered distributions

Definition 2.10 (Fourier transform of tempered distributions)**.**
Let $T \in \mathcal{S}'$. Then the Fourier transform of T, denoted by $\hat{T}$, and its inverse, denoted by $\check{T}$, are tempered distributions defined by

$$\hat{T}(u) = T(\hat{u}), \qquad u \in \mathcal{S} \tag{2.34}$$

and

$$\check{T}(u) = T(\check{u}), \qquad u \in \mathcal{S}, \tag{2.35}$$

respectively.

Understood as a map $\hat{}: \mathcal{S}' \to \mathcal{S}'$, the Fourier transform is an automorphism on $\mathcal{S}'$ and the unique weakly continuous extension of the Fourier transform on $\mathcal{S}$.[9] Many properties from the Fourier transform on $\mathcal{S}$ carry over:

Proposition 2.11. *For each $T \in \mathcal{S}'$ and $u \in \mathcal{S}$ we have*

- *Parseval's identity:*
$$\hat{T}(\bar{\hat{u}}) = T(\bar{u}), \tag{2.36}$$
where $\bar{u}$ denotes the complex conjugate of u.

- *Constant/Delta Distribution:*
$$\hat{\delta} = 1 \quad \textit{and} \quad \hat{1} = \delta. \tag{2.37}$$

- *Polynomials/Derivatives:*
$$\widehat{x^\alpha D^\beta T}(u) = (-iD)^\alpha (-ip)^\beta \hat{T}(u). \tag{2.38}$$

- *Multiplication/Convolution:*
$$\hat{T}(u * v) = T(\hat{u}\hat{v}) \quad \textit{and} \quad \hat{T}(uv) = T(\hat{u} * \hat{v}). \tag{2.39}$$

- *Real Linear Transformation:*
Let A be a real linear transformation of $\mathbb{R}^n$ into itself. Then
$$\hat{T} \circ A^{-1}(u) = \det\left(A^\dagger\right) \cdot \widehat{T \circ A^\dagger}(u), \tag{2.40}$$
where $A^\dagger$ is the adjoint of A with respect to the fixed bilinear form px of the Fourier transform.

[9]cf. Theorem IX.2 of [RS75], p.5.

Proof.
Whenever there are two equations, we will only prove one of them:

Parseval's identity:

$$\hat{T}(\bar{\hat{u}}) = \hat{T}(\check{\bar{u}}) = T(\hat{\check{\bar{u}}}) = T(\bar{u}).$$

Constant/Delta Distribution:

$$\hat{\delta}(u) = \delta(\hat{u}) = \int u(x) d^n x = 1(u),$$

where 1 denotes the 1-function viewed as a distribution.

Polynomials/Derivatives:

$$\begin{aligned}
\widehat{x^\alpha D^\beta T}(u) &= (-1)^{|\beta|} T(D^\beta x^\alpha \hat{u}) \\
&= (-1)^{|\beta|} \hat{T}(\widecheck{D^\beta x^\alpha \hat{u}}) \\
&= (-1)^{|\beta|} \hat{T}((ip)^\beta (iD)^\alpha u) \\
&= (-1)^{|\beta|+|\alpha|} (iD)^\alpha (ip)^\beta \hat{T}(u) \\
&= (-iD)^\alpha (-ip)^\beta \hat{T}(u).
\end{aligned}$$

Multiplication/Convolution:

$$\hat{T}(u * v) = T(\widehat{u * v}) = T(\hat{u}\hat{v}).$$

Real Linear transformation:

$$\hat{T} \circ A^{-1}(u) = \hat{T}(|J(A)| \cdot u \circ A) = \det A \cdot T(\widehat{u \circ A}).$$

The Fourier transform of $u \circ A$ gives $\frac{1}{\det A} \cdot \hat{u} \circ (A^{-1})^\dagger$, where $(\ )^\dagger$ denotes the adjoint of A with respect to the bilinear form px for which the Fourier transform was defined. Finally, we obtain

$$\hat{T} \circ A^{-1}(u) = T(\hat{u} \circ (A^{-1})^\dagger) = \det A^\dagger \cdot \widehat{T \circ A^\dagger}(u).$$

■

Corollary 2.12. *(Lorentz transformation) Let Λ be a real linear transformation on $\mathbb{M}$ keeping the Minkowski product px invariant.*[10] *Then $\Lambda^\dagger = \Lambda^{-1}$ and $|det\ \Lambda| = 1$ such that for each tempered distribution*

$$\hat{T} \circ \Lambda = \widehat{T \circ \Lambda} \tag{2.41}$$

or in generalized function notation

$$\hat{T}(\Lambda p) = \widehat{T(\Lambda x)}(p). \tag{2.42}$$

[10] We will introduce Minkowski space and Lorentz transformations properly in the next subsection.

2.3 Representation theory pertaining to Lorentz- and Poincaré group

In this subsection we will pay attention to the Lorentz and Poincaré group and their representations. The main part of this subsection will focus on the Lorentz group including a classification of its irreducible and simply reducible representations as well as an analysis of the structure of Lorentz-invariant and Lorentz-covariant functions and generalized functions.

2.3.1 The Lorentz- and Poincaré group

Quantum field theory is a relativistic theory and, to be precise, it obeys the principles of special relativity. To start with, the four-dimensional Minkowski (vector) space is defined as the vector space $\mathbb{R}^4$ endowed with the Pseudo-Euclidean scalar product

$$(p,q) \mapsto pq = p^\mu q_\mu = \eta_{\mu\nu} p^\mu q^\nu = p_0 q_0 - \vec{p}\cdot\vec{q} \tag{2.43}$$

for arbitrary four-vectors $p, q \in \mathbb{M}$. The spacetime coordinates then live in the affine space of $\mathbb{M}$, also called Minkowski (coordinate) space, that inherits the Pseudo-Euclidean metric $d : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ defined by

$$d^2(x,y) = (x-y)^2 = (x_0 - y_0)^2 - (\vec{x}-\vec{y})\cdot(\vec{x}-\vec{y}), \quad x, y \in \mathbb{M}. \tag{2.44}$$

From the affine space we again get the Minkowski vector space by taking the tangent space or its dual. In particular in the latter case it will also be called Minkowski momentum space. The basic symmetry groups of special relativity are then the invariance group of the Pseudo-Euclidean scalar product and the invariance group of the Pseudo-Euclidean metric, named (real) Lorentz- and Poincaré group, respectively. The Lorentz group $\mathcal{L}$ may be written as matrix Lie group

$$\mathcal{L} = \{\Lambda \in M(4\times 4, \mathbb{R}), \Lambda g \Lambda^T = g\} \tag{2.45}$$

such that transformations $x \mapsto x' = \Lambda x$ keep the Minkowski product invariant. The Poincaré group $\mathcal{P}$ may be written as the set

$$\mathcal{P} = \{(a,\Lambda) \in \mathbb{M} \times \mathcal{L}\} \text{ together with the product } (a,\Lambda)(b,\Lambda') = (a + \Lambda a', \Lambda\Lambda') \tag{2.46}$$

for each $(a,\Lambda),(a',\Lambda') \in \mathcal{P}$ where (a,Λ) represents a transformation $x \mapsto x' = (a,\Lambda)x = \Lambda x + a$.[11]. The (real) Lorentz and Poincaré group each have four connected components from which we will select the components that contain the identity. These components are called special orthochronous - or also proper - and are selected by the conditions

$$\det \Lambda = +1 \text{ (special)} \qquad \text{and} \qquad \Lambda_{00} \geq 1 \text{ (orthochronous)}. \tag{2.47}$$

We will denote them by $\mathcal{L}_+^\uparrow$ and $\mathcal{P}_+^\uparrow$, respectively.

[11]With these definitions we have established that the Poincaré group is a semidirect product $\mathcal{P} = \mathbb{M} \rtimes \mathcal{L}$ of the translation group and the Lorentz group.

2.3.2 Representation theory of the Lorentz group

In view of physics symmetry groups and symmetry transformations are of great importance. In the section above we have learned how relativistic symmetry transformations affect the coordinates of Minkowski spacetime. Usually, we do not only want to study the behaviour of coordinates under symmetry transformations, but also of the physical entities or observables of the theory. The mathematical structure that corresponds to the transformation behaviour of a mathematical object under the action of a (continuous) symmetry group is called a (Lie group) representation.

Definition 2.13. A representation[12] of a Lie group G is given by a finite-dimensional linear space V and a Lie group homomorphism $\rho : G \to \mathrm{GL}(V)$. The homomorphism ρ is called representation map. $\mathrm{GL}(V)$ denotes the general linear group on V.

Although unspecified for the definition above, for the sake of definiteness, we should agree to use the field of complex numbers as base field for the representation spaces.
In general it turns out that the representation theory of simply connected Lie groups is particularly nice as their representations are in one-to-one correspondence with the representation theory of their Lie algebras (via the exponential map). Therefore let us study the representations of the universal covering[13] of $\mathcal{L}_+^\uparrow$, which is the complex special linear group $SL(2,\mathbb{C})$ of complex 2×2 matrices with determinant equal to one.

Proposition 2.14. *$SL(2,\mathbb{C})$ is the (unique) universal double covering of $\mathcal{L}_+^\uparrow$.*

Proof. Let us define $\underset{\sim}{x} \equiv x^\mu \sigma_\mu$, where $\sigma_0 = \mathbb{1}$ and $\sigma_i,\ i = 1,2,3$ are the Pauli-matrices[14]. The mapping $x \mapsto \underset{\sim}{x}$ defines an isomorphism from $\mathbb{M}$ to the space of hermitian 2×2-matrices $\mathcal{M}_{2\times2,Herm}$[15].
A homomorphism from $SL(2,\mathbb{C}) \to \mathcal{L}_+^\uparrow, \quad \underset{\sim}{\Lambda} \mapsto \Lambda \equiv \Lambda(\underset{\sim}{\Lambda})$ is then given by

$$\underset{\sim}{\Lambda x} \equiv \underset{\sim}{\Lambda}\underset{\sim}{x}\underset{\sim}{\Lambda}^\dagger, \tag{2.48}$$

where $\underset{\sim}{\Lambda}^\dagger \equiv \bar{\underset{\sim}{\Lambda}}^T$ denotes the adjoint of $\underset{\sim}{\Lambda}$. In order to see that this map defines a double covering of $\mathcal{L}_+^\uparrow$ and that $SL(2,\mathbb{C})$ is simply-connected the reader is referred to [BLOT90, Ex. 3.4 and above]. ∎

Matrix Lie groups have a natural action on column vectors and therefore a canonical representation called self representation consisting of the identity map and the column vector space. In the case of $SL(2,\mathbb{C})$ the self representation is given by its natural action on $\mathbb{C}^2$:

$$(\underset{\sim}{\Lambda} z)_a = \underset{\sim}{\Lambda}_a{}^b z_b, \quad a = 1,2, z \in \mathbb{C}^2. \tag{2.49}$$

The irreducible representations of $SL(2,\mathbb{C})$ are known to be of the following form:

[12]We will confine the discussion here to finite-dimensional representations (corresponding to representations with a finite dimensional linear space V) unless stated otherwise.

[13]A covering of a Lie group G is given by a Lie group H and a Lie-homomorphism $H \to G$ which is onto and a local isomorphism. The universal covering is uniquely defined by requiring H to be simply connected.

[14]For definiteness they are specified to be $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

[15]It is not difficult to show that the Pauli matrices form a basis on $\mathcal{M}_{2\times2,Herm}$. Therefore the map is clearly onto and as $\mathbb{M}$ and $\mathcal{M}_{2\times2,Herm}$ have the same dimension, it defines an isomorphism. That $\underset{\sim}{x}$ is hermitian is also easy to see via $\underset{\sim}{x}^\dagger = x^\mu \bar{\sigma}_\mu^T = x^\mu \sigma_\mu = \underset{\sim}{x}$.

Proposition 2.15. *Each irreducible representation of $SL(2,\mathbb{C})$ is characterized by a tuple (j,k) of non-negative half-integral numbers j,k. The representation $\mathcal{D}^{(j,k)}$ corresponding to (j,k) is given by $2j$ copies of the self-representation and $2k$ copies of its complex conjugate acting on the space of spinors. Explicitly, we can write*

$$\mathcal{D}^{(j,k)}(\utilde{\Lambda}) = \utilde{\Lambda}^{\otimes 2j} \otimes \bar{\utilde{\Lambda}}^{\otimes 2k} \tag{2.50}$$

acting on $S^{2j}(\mathbb{C}^2) \otimes S^{2k}(\mathbb{C}^2)$[16]*, where $S^r(V)$ is the space of symmetric tensors of rank r on a vector space V. Thus $\mathcal{D}^{(j,k)}$ is of dimension $(2j+1)(2k+1)$.*

Remark 2.16. The irreducible representations of $SL(2,\mathbb{C})$ stand in close connection to the irreducible representations of the compact Lie group $SU(2)$ of special unitary $2{\times}2$-matrices. The irreducible representations of $SU(2)$ are well-known to be characterized by a single non-negative half-integral number s called spin. For each spin s there is a unique holomorphic and a unique antiholomorphic extension to a representation of $SL(2,\mathbb{C})$ corresponding to $\mathcal{D}^{(s,0)}$ and $\mathcal{D}^{(0,s)}$ and for each irreducible rep. of $SL(2,\mathbb{C})$ we then have $\mathcal{D}^{(j,k)} \cong \mathcal{D}^{(j,0)} \otimes \mathcal{D}^{(0,k)}$. For a more thorough discussion of the relations between representations of $SU(2)$ and $SL(2,\mathbb{C})$ the reader is referred to [Zhe73, §39, §42 and §43].

It is also known that

Proposition 2.17. *For each non-negative half-integral numbers s,t,s',t' the representation $\mathcal{D}^{(s,t)} \otimes \mathcal{D}^{(s',t')}$ of $SL(2,\mathbb{C})$ decomposes into $\bigoplus_{|s-s'|\leq j\leq s+s'} \bigoplus_{|t-t'|\leq k\leq t+t'} \mathcal{D}^{(j,k)}$ where the sums run over j and k in integral steps.*

Remark 2.18. In view of the remark above we can relate this decomposition to the corresponding decomposition for $SU(2)$-representations. The latter is well-known as the decomposition of angular momentum. In order to make that statement more precise using the shorthand notation $(j,k) \equiv \mathcal{D}^{(j,k)}$ note that

$$\begin{aligned} (s,t) \otimes (s',t') &\cong (s,0) \otimes (0,t) \otimes (s',0) \otimes (0,t') && (2.51)\\ &\cong (s,0) \otimes (s',0) \otimes (0,t) \otimes (0,t'). && (2.52) \end{aligned}$$

Angular momentum decomposition then gives us $(s,0) \otimes (s',0) \cong \bigoplus_{|s-s'|\leq j\leq s+s'} (j,0)$ and similarly for the complex conjugate. Hence we arrive at the expression in the proposition above.

Finally, we want to relate the results of representation theory of $SL(2,\mathbb{C})$ to the representation theory of the Lorentz group. For this note that the representation $(1/2,1/2)$ may be realized in the space of complex hermitian $2{\times}2$-matrices and is then given by

$$\utilde{x} \mapsto \utilde{\Lambda}\utilde{x}\utilde{\Lambda}^\dagger, \quad \utilde{x} \in \mathcal{M}^{Herm}_{2\times 2}. \tag{2.53}$$

This map closely resembles the covering homomorphism defined in eq. (2.48). In fact, for each $x \in \mathbb{M}$ we have

$$\mathcal{D}^{(1/2,1/2)}(\utilde{\Lambda}) \circ \utilde{\ }\,(x) = \mathcal{D}^{(1/2,1/2)}(\utilde{\Lambda})(\utilde{x}) = \utilde{\Lambda}\utilde{x}\utilde{\Lambda}^\dagger = \utilde{\Lambda(\utilde{\Lambda})x} = \utilde{\ } \circ \Lambda(\utilde{\Lambda})\,(x). \tag{2.54}$$

[16]This is the subspace of spinors of degree $(2j,2k)$. The general spinor space is given by $S(\mathbb{C}^2) \otimes S(\mathbb{C}^2)$ with $S(V) \equiv \bigoplus_{r=0}^{\infty} S^r(V)$

Remember here that $\underset{\sim}{}$ represents the isomorphism from $\mathbb{M}$ to $\mathcal{M}_{2\times 2}^{Herm}$ mapping x to $\underset{\sim}{x} = x^\mu \sigma_\mu$ which was introduced in the proof of Proposition 2.14.
Thus $\mathcal{D}^{(1/2,1/2)}(\underset{\sim}{\Lambda}) \circ \underset{\sim}{} = \underset{\sim}{} \circ \Lambda(\underset{\sim}{\Lambda})$ and hence the two representations $(\mathcal{D}^{(1/2,1/2)}, \mathcal{M}_{2\times 2, Herm})$ and the self-representation $(\Lambda, \mathbb{M})$ are equivalent. We should introduce:

Remark 2.19. A representation ρ of $SL(2,\mathbb{C})$ is called a single-/double-valued representation of $\mathcal{L}_+^\uparrow$ depending on whether $\rho(-1) = +1$ or $\rho(-1) = -1$, where 1 stands for the identity on the corresponding space. A single-valued representation is equivalent to an ordinary representation. This can be seen as

$$\rho_1(\Lambda) = \rho(\underset{\sim}{\Lambda}) \quad \text{for } \Lambda = \Lambda(\underset{\sim}{\Lambda}) \tag{2.55}$$

is a well-defined definition of a representation of $\mathcal{L}_+^\uparrow$ precisely if $\{\pm 1\} = \ker \Lambda(\cdot) \subset \ker \rho$, i.e., if, and only if, $\rho(-1) = 1$.

The full result then is

Proposition 2.20. *Each irreducible representation $\mathcal{D}^{(j,k)}$ of $SL(2,\mathbb{C})$ with non-negative half-integral numbers j,k gives rise to an irreducible representation of $\mathcal{L}_+^\uparrow$, which is single-valued if, and only if, $j+k$ is integral and double-valued if, and only if, $j+k$ is half-integral (and not integral).*

Proof. In only the next equation we will distinguish for the sake of clarity between the identity number 1, matrix $\mathbb{1} \in SL(2,\mathbb{C})$ and tensor $\mathbb{1}_{2j\otimes 2k}$ in the space of spinors of degree $(2j, 2k)$:

$$\mathcal{D}^{(j,k)}(-\mathbb{1}) = (-\mathbb{1})^{\otimes(2j+2k)} = (-1)^{2(j+k)} \mathbb{1}_{2j\otimes 2k}. \tag{2.56}$$

Hence, disposing of the distinction again,

$$\mathcal{D}^{(j,k)}(-1) = \begin{cases} +1 & \text{for } j+k \text{ integral,} \\ -1 & \text{for } j+k \text{ half-integral and not integral.} \end{cases} \tag{2.57}$$

∎

For an irreducible representation of $SL(2,\mathbb{C})$ to be real it is required that $j = k$. As this automatically implies the single-valuedness of the representation when understood as a Lorentz-representation, we have established that the (real) irreducible representations of the proper Lorentz group are precisely the irreducible representation of $SL(2,\mathbb{C})$ with $j = k$.

As a prerequisite for the next chapter let us introduce a useful implementation of irreducible and completely reducible $SL(2,\mathbb{C})$-representations. For half-integral numbers j,k define the space $\mathcal{P}^{(j,k)}$ of polynomials in two variables $\omega, \bar{\omega} \in \mathbb{C}^2$ that are homogeneous of degree $2j$ in ω and homogeneous of degree $2k$ in $\bar{\omega}$. On this polynomial space $\mathcal{D}^{(j,k)}$ may be conveniently implemented by

$$\mathcal{D}^{(j,k)}(\underset{\sim}{\Lambda})\Psi[\omega,\bar{\omega}] \equiv T_{\underset{\sim}{\Lambda}^{-1}}\Psi[\omega,\bar{\omega}] \equiv \Psi[\underset{\sim}{\Lambda}^{-1}\omega, \bar{\underset{\sim}{\Lambda}}^{-1}\bar{\omega}], \qquad \Psi \in \mathcal{P}^{(j,k)}. \tag{2.58}$$

Here we used the insertion homomorphism $T_{\underset{\sim}{\Lambda}}\Psi[\omega,\bar{\omega}] \equiv \Psi[\underset{\sim}{\Lambda}\omega, \bar{\underset{\sim}{\Lambda}}\bar{\omega}]$, which is defined for any polynomial with a natural action of $SL(2,\mathbb{C})$ on the variable spaces.

For a completely reducible representation $\bigoplus_{(j,k)\subset S} \mathcal{D}^{(j,k)}$ where S is a finite set of tuples of half-integral numbers, we may just take $\bigoplus_{(j,k)\subset S} \mathcal{P}^{(j,k)}$ as representation space where the action of $SL(2,\mathbb{C})$ remains as defined in equation (2.58).

2.3.3 Lorentz-invariant and Lorentz-covariant generalized functions

Among the mathematical objects on which a (Lorentz group) representation may act are the functions and generalized functions on Minkowski space. In particular they may satisfy a specific transformation behaviour under the group actions, named covariance.

Definition 2.21. A Lorentz-covariant tempered distribution is a vector-valued generalized function $F \in \mathcal{S}'(\mathbb{M}, V)$ which for each $u \in \mathcal{S}(\mathbb{M})$ satisfies

$$F(u_\Lambda) = \rho(\Lambda)F(u) \tag{2.59}$$

for some representation (ρ, V) of the Lorentz group and $u_\Lambda(p) \equiv u(\Lambda^{-1}p)$. It is referred to be Lorentz-invariant if, and only if, ρ is the trivial representation

$$F(u_\Lambda) = F(u). \tag{2.60}$$

The conditions of Lorentz-co- and invariance in generalized function notation amounts to $F(\Lambda p) = \rho(\Lambda)F(p)$ and $F(\Lambda p) = F(p)$.[17]

The simpler case of Lorentz-invariant generalized functions gives rise to the following representation

Proposition 2.22. *For each Lorentz-invariant $F \in \mathcal{S}'(\mathbb{M})$ there are Lorentz-invariant $f_\pm \in \mathcal{S}'(\mathbb{R})$ coinciding on $\mathbb{R}_-$ such that*

$$F(p) = \begin{cases} f_+(p^2) & \text{for } p \notin \bar{V}^-, \\ f_-(p^2) & \text{for } p \notin \bar{V}^+, \end{cases} \tag{2.61}$$

where $\bar{V}^\pm \equiv \{p \in \mathbb{M} : p^2 \geq 0, p_0 \geq 0\} \subset \mathbb{M}$ denotes the closed upper/lower light cone.

Remark 2.23. The expressions $f_\pm(p^2)$ define generalized functions on $\mathbb{M}\backslash\bar{V}^\mp$. They have the following precise meaning: The map

$$(\tau, \vec{p}) \mapsto j_\pm(\tau, \vec{p}) \equiv \left(\pm\sqrt{\tau + |\vec{p}|^2}, \vec{p}\right), \quad |\vec{p}|^2 > -\tau \tag{2.62}$$

defines a diffeomorphism from $\{(\tau, \vec{p}) \in \mathbb{R}^4 : |\vec{p}|^2 > -\tau\}$ to $\mathbb{M}\backslash\bar{V}^\mp$ with inverse $j_\pm^{-1}(p) = (p^2, \vec{p})$. Hence for $p \notin \bar{V}^\mp$ we can write

$$f_\pm(p^2) \equiv (f_\pm \otimes 1) \circ j_\pm^{-1}(p), \tag{2.63}$$

where 1 is the constant 1 viewed as an element of $\mathcal{S}'(\mathbb{R}^3)$. In distributional notation we have

$$f_\pm(p^2)(u) = f_\pm(v_\pm), \quad v_\pm(\tau) \equiv \pm\int_{|\vec{p}|^2 > -\tau} \frac{u(\pm\sqrt{\tau + |\vec{p}|^2}, \vec{p})}{2\sqrt{\tau + |\vec{p}|^2}} d\vec{p}, \quad u \in \mathcal{S}(\mathbb{M}\backslash\bar{V}^\mp). \tag{2.64}$$

Proof. See [BLOT90, Chapter 3.2 pp. 131]. ■

[17] Throughout this section we will use variables $p, q \in \mathbb{M}$ which can be thought of either as momentum or as coordinate variables.

Corollary 2.24. *For each Lorentz-invariant function $F \in \mathcal{S}(\mathbb{M})$ there exists two functions $f_\pm \in \mathcal{S}(\mathbb{R})$ such that*

$$F(p) = f_{\epsilon(p_0)}(p^2), \tag{2.65}$$

where $\epsilon(p_0)$ is the sign of p_0. In particular, a Lorentz-invariant function that is constant with respect to p_0 is a constant.

Let us fix here F to be an arbitrary Lorentz-covariant generalized function. What we want to achieve is a decomposition of F in Lorentz-invariant generalized functions f_ρ with respect to a fixed family of Lorentz-covariant polynomials Q_ρ, the so-called standard covariants. The decomposition should be of the form

$$F(p) = \sum_\rho f_\rho(p) Q_\rho(p). \tag{2.66}$$

In order to be useful we require standard covariants to form a polynomial basis. This means that the Q_ρ span the space of Lorentz-covariant polynomials with respect to the ring of Lorentz-invariant polynomials and that $F(p) \equiv 0$ implies $f_\rho(p) \equiv 0\ \forall \rho$. Such a decomposition exists in general and the Lorentz-invariant distributions $f_\rho(p)$ will be defined up to a finite number of constants. For the details the reader is referred to [BLOT90, Chapter 3.3]. Here we just want to cite [BLOT90, Proposition 3.6] in a shortened and adapted form: In order to state the result let us take the representation $\mathcal{D}^{(j,k)}$ to be realized in the space $\mathcal{P}^{(j,k)}$ introduced in the former subsection. In the definition of Lorentz-covariant generalized functions (see Definition 2.21) this corresponds to $(\rho, V) = (\mathcal{D}^{(j,k)}, \mathcal{P}^{(j,k)})$ or, in other words, that the values of F (when smeared with a test function) are homogeneous polynomials in ω and $\bar{\omega} \in \mathbb{C}^2$. The notation is defined as $F(p;\omega,\bar{\omega}) \equiv F(p)[\omega,\bar{\omega}]$.[18] Represented in this form the covariance condition becomes an invariance condition

$$F(\Lambda(\underset{\sim}{\Lambda})p;\omega,\bar{\omega}) = F(p;\underset{\sim}{\Lambda}^{-1}\omega,\bar{\underset{\sim}{\Lambda}}^{-1}\bar{\omega}) \quad \Leftrightarrow \quad F(\Lambda(\underset{\sim}{\Lambda})p;\underset{\sim}{\Lambda}\omega,\bar{\underset{\sim}{\Lambda}}\bar{\omega}) = F(p;\omega,\bar{\omega}). \tag{2.67}$$

This explains why it is so useful to write Lorentz-covariant generalized functions in this form. A distinguished role in this covariant decomposition will be played by the invariant combination of p, ω and $\bar{\omega}$ given by $\bar{\omega}\tilde{p}\omega$, where $\tilde{p} \equiv p^\mu(\epsilon\sigma_\mu^T\epsilon^{-1}) = p^\mu(\sigma^\mu)_\mu$ is a $2{\times}2$-matrix and ω and $\bar{\omega}$ are understood as column- and row-vectors, respectively.

Proposition 2.25. *An arbitrary Lorentz-covariant generalized function F on $\mathbb{M}$ transforming according to $\mathcal{D}^{(j,k)}$ is non-zero only for $j = k = n/2$ and in this case may be represented as a tempered distribution*

$$F(p;\omega,\bar{\omega}) = (\bar{\omega}\tilde{p}\omega)^n f(p), \tag{2.68}$$

where $f(p)$ is a Lorentz-invariant generalized function defined within n arbitrary constants. More precisely, if $f_0(p)$ is a fixed solution of eq. (2.68) (regarded as an equation in $f(p)$), then the general solution has the form

$$f(p) = f_0(p) + \sum_{l=0}^{n-1} a_l \Box^l \delta(p) \tag{2.69}$$

for some constants a_l, $l = 0, ..., n-1$.

[18] In other words, F can be understood as a generalized function in not only p, but also ω and $\bar{\omega}$ as polynomials can define generalized functions.

Proof. See [BLOT90, Proposition 3.6]. ■

The result may be easily extended to arbitrary Lorentz-covariant generalized function on $\mathbb{M}$ transforming according to $T = \bigoplus_{(j,k)\subset S} \mathcal{D}^{(j,k)}$ for some finite set S consisting of tuples of half-integral numbers.

2.4 Lorentz-invariant solutions of the d'Alembert equation

The goal of this section is to determine the solutions of the linear partial differential equation

$$\Box F(x) = \eta_{\mu\nu}\partial^\mu\partial^\nu F(x) = 0 \tag{2.70}$$

in the space of Lorentz-invariant tempered distributions on $\mathbb{M}$. This problem is equivalent (by applying the Fourier transform) to the division problem

$$p^2 F(p) = 0, \tag{2.71}$$

where $F(p) \equiv \tilde{F}(p)$ denotes the Fourier transform of $F(x)$ which is also Lorentz-invariant (see e.g. Corollary 2.12) and tempered (by definition). From Eq. (2.71) one immediately sees that the support of $F(p)$ must be contained in the set $\Gamma \equiv \Gamma_0 \equiv \{p \in \mathbb{M} : p^2 = 0\} = \{p \in \mathbb{M} : p_0 = \pm|\vec{p}|\}$. The most difficult part of this division problem is to deal with the point $p = 0$ and its neighbourhood. Let us start by solving the following (sub-)problem

Lemma 2.26. *An arbitrary Lorentz-invariant tempered distribution $F(p)$ on $\mathbb{M}_\times \equiv \mathbb{M}\backslash\{0\}$ subject to the equation $p^2F(p) = 0$ is of the form*

$$F(p) = a\theta(p_0)\delta(p^2) + b\theta(-p_0)\delta(p^2) \tag{2.72}$$

for some constants a, b.

Proof. To begin with, let us denote the restrictions of F to the complements of $\bar{V}^\mp$ in $\mathbb{M}$ by $F^\pm$. This is well-defined as $\mathbb{M}\backslash\bar{V}^\mp$ are open subsets of $\mathbb{M}_\times$. Solving the division problem for the restrictions separately, by the gluing principle of generalized functions we then find F (on $\mathbb{M}_\times$) to be uniquely specified by its restrictions $F^\pm$ (see Proposition 2.3). Let us clarify that the following equations of the proof are implicitly meant to live upon the complements of $\bar{V}^\mp$ and in particular $p = 0$ is excluded.
For F^+ the division problem reads

$$p^2F^+(p) = (p_0 - |\vec{p}|)(p_0 + |\vec{p}|)F^+(p) = 0. \tag{2.73}$$

The general solution to the division problem $xf(x,y) = 0$ for $f \in \mathcal{S}(\mathbb{R}\times\mathbb{R}^n)$ is $f(x,y) = \delta(x)g(y)$ for some $g \in \mathbb{S}(\mathbb{R}^n)$ (see Proposition 2.7). Thus[19] we arrive at

$$(p_0 + |\vec{p}|)F^+(p) = \delta(p_0 - |\vec{p}|)g(\vec{p}) \tag{2.74}$$

[19] An application of the diffeomorphism $p_0 \mapsto p_0 - |\vec{p}|$ for fixed $|\vec{p}|$ relates the two problems.

or equivalently

$$F^+(p) = \frac{1}{2|\vec{p}|}\delta(p_0 - |\vec{p}|)g(\vec{p}) \tag{2.75}$$
$$= \theta(p_0)\delta(p^2)g(\vec{p}). \tag{2.76}$$

In the first line we have used that $p_0 + |\vec{p}|$ is invertible because of the support properties of F^+ and that we can apply $p_0 - |\vec{p}| = 0$ to the factors in front of the delta distribution. As $\bar{V}^\pm$ and also their complements are Lorentz-invariant, the restrictions of F to the latter must also be Lorentz-invariant. As $\theta(p_0)\delta(p^2)$ is Lorentz-invariant, also $g(\vec{p})$ must be Lorentz-invariant. By Corollary 2.24 this implies $g(\vec{p}) \equiv$ *const.*. Applying the same procedure to F^- and gluing together $F^\pm$ we finally obtain

$$F(p) = a\theta(p_0)\delta(p^2) + b\theta(-p_0)\delta(p^2). \tag{2.77}$$

■

Lemma 2.27. *An arbitrary Lorentz-invariant tempered distribution $F(p)$ on $\mathbb{M}$ subject to the equation $p^2F(p) = 0$ has the form*

$$F(p) = a\theta(p_0)\delta(p^2) + b\theta(-p_0)\delta(p^2) + c\delta(p) \tag{2.78}$$

for some constants a, b, c where we have defined[20]

$$\left(\theta(\pm p_0)\delta(p^2)\right)(u) \equiv \left(\frac{1}{2p_0}\delta(p_0 \mp |\vec{p}|)\right)(u) \equiv \int \pm u(\pm|\vec{p}|, \vec{p})\frac{d\vec{p}}{2|\vec{p}|}. \tag{2.79}$$

Proof. The restriction of F to the open subset $\{p : p \neq 0\}$ of $\mathbb{M}$ fulfils the conditions of Lemma 2.26. Hence for $p \neq 0$ we can write

$$F(p) = a\theta(p_0)\delta(p^2) + b\theta(-p_0)\delta(p^2) \tag{2.80}$$

for some constants a, b. As this expression becomes ambiguous at $p = 0$ we have to fix a proper definition for the extension of this expression to the whole of $\mathbb{M}$. This is implemented by defining $\theta(\pm p_0)\delta(p^2)$ as in eq. (2.79) above. This extension clearly solves the division problem.
Let us now take two solutions to the division problem with the same restriction (2.80) to $\mathbb{M}_\times$. Then their difference is concentrated at $p = 0$ and necessarily solves the division problem, too. That it is concentrated at $p = 0$ implies that the difference has the form $P(\Box)\delta(p)$ for a polynomial P. That it is a solution of the division problem then implies that the degree of P is zero. Hence, we can only add terms proportional to $\delta(p)$ to the above defined extension of (2.80). ■

Let us define $\hat{D}(p) \equiv \epsilon(p_0)\delta(p^2)$ and let $\hat{D}^{(\pm)}(p) \equiv \pm\theta(\pm p_0)\delta(p^2)$ denote its positive and negative energy (positive and negative p_0) parts where we understand $\theta(p_0)\delta(p^2)$ defined as in eq. (2.79). These distributions $D = D^{(+)} + D^{(-)}$ and $D^{(\pm)}$ are of great importance to quantum field theory and the former is named Pauli-Jordan commutator function.[21] What we finally obtain is

[20] See also Remark 2.23 for the general case of a p^2-dependent generalized function.

[21] Note that there are many different conventions for the normalization of the commutator function. For instance [BLOT90] uses $\hat{D}(p) = -2\pi i\epsilon(p_0)\delta(p^2)$.

Corollary 2.28. *An arbitrary Lorentz-invariant tempered distribution $F(x)$ on $\mathbb{M}$ subject to the equation $\Box F = 0$ is of the form*

$$F(x) = aD^{(+)}(x) + bD^{(-)}(x) + c \tag{2.81}$$

for some constants a, b, c.

Proof. The Fourier transformed problem is solved by Lemma 2.27. ∎

In this thesis we will not only encounter the homogeneous equation $\Box F = 0$, but also:

Theorem 2.29. *An arbitrary Lorentz-invariant tempered distribution $F(x)$ on $\mathbb{M}$ subject to the equation $\Box F = T$ for a fixed Lorentz-invariant tempered distribution $T(x)$ on $\mathbb{M}$ is of the form*

$$F = F_{hom} + F_{part}, \tag{2.82}$$

where $F_{hom} = aD^{(+)} + bD^{(-)} + c$ is an arbitrary solution to the homogeneous KG equation and F_{part} is a fixed particular solution to the inhomogeneous KG equation ($\Box F_{part} = T$).

Proof. A solution to $\Box F = T$ always exists.[22] In addition, let there be two solutions to the inhomogeneous equation, then their difference has to satisfy the homogeneous equation. ∎

2.5 Unitary Hilbert space representations and their generators

In this section we briefly discuss unitary representations on (possibly infinite-dimensional) Hilbert spaces. In particular we want to carry over a part of Lie group and Lie algebra theory to the representing Hilbert space. In order to do that we give a short account on the basic results of the spectral theory of bounded and unbounded self-adjoint operators, like the spectral theorem and Stone's theorem, which will give us the opportunity to define the representations of the generators of (a connected component of) a Lie group as unbounded self-adjoint operators on the Hilbert space.

Definition 2.30. A strongly continuous unitary (possibly infinite-dimensional) representation of a Lie group is given by a 3-tuple $(\mathcal{G}, \mathcal{H}, U)$ consisting of a Lie group $\mathcal{G}$, a Hilbert space $\mathcal{H}$ and a continuous group homomorphism $U : \mathcal{G} \to \mathcal{U}(\mathcal{H})$ where $\mathcal{U}(\mathcal{H})$ denotes the group of unitary operators on $\mathcal{H}$.

Remark 2.31. The reader may think about using other topologies than the strong operator topology for $\mathcal{U}(\mathcal{H})$. It turns out that the norm topology is usually to fine for applications in physics. The other two topologies on $\mathcal{U}(\mathcal{H})$, weak and strong, respectively, are inherited from the space $\mathcal{L}(\mathcal{H})$ of linear operators on $\mathcal{H}$. But, indeed, on $\mathcal{U}(\mathcal{H})$ these topologies are

[22] A solution to $p^2\tilde{F}(p) = \tilde{T}(p)$ on $p^2\mathcal{S}(\mathbb{M})$ is given by defining $\tilde{F}(p^2u) \equiv \tilde{T}(u)\ \forall u \in \mathcal{S}(\mathbb{M})$ and an extension to $\mathcal{S}(\mathbb{M})$ exists by the Hahn-Banach theorem. Crucial is here that $p^2u \mapsto u$ is continuous which is not trivial. For the details the reader is referred to the original work by Hörmander[Hö58].

equivalent, as for $(U_n)_n \subset \mathcal{U}(\mathcal{H})$ with $U_n \to U$ weakly[23] as $n \to \infty$ and for each $\phi \in \mathcal{H}$ we have

$$\begin{aligned} ||U_n\phi - U\phi||^2 &= \langle U_n\phi, U_n\phi\rangle + \langle U\phi, U\phi\rangle - \langle U_n\phi, U\phi\rangle - \langle U\phi, U_n\phi\rangle \\ &= 2||\phi||^2 - \langle U_n\phi, U\phi\rangle - \langle U\phi, U_n\phi\rangle \\ &\to 0. \end{aligned}$$

In the second line we used the unitarity of U_n and U. In the third line we use the weak convergence of $(U_n)_n$. Hence both notions of convergence are equivalent on $\mathcal{U}(\mathcal{H})$. It is conventional to refer to this continuity as strong continuity in order to distinguish it from norm-continuity.

The unitary representations of the Poincaré group $\mathcal{P}_+^\uparrow$ are of big importance for quantum field theory. We have already mentioned Wigner's classification theorem. What will be important here is the relation between infinitesimal generators of the Lie group and their representations as unbounded operators. Let us take for simplicity (and by choice of relevance) the translation subgroup $\mathcal{T} \cong \mathbb{R}^4$ of the Poincaré group. By the theory of Lie groups and Lie algebras, it is well known that there are generators of $\mathcal{T}$ in the Lie algebra associated to $\mathcal{T}$, i.e., there are Lie algebra elements p_μ such that for each $T \in \mathcal{T}$ we obtain the representation $T = e^{ia_\mu p^\mu}$ for a vector $a \in \mathbb{M}$ where $e^{(\cdot)}$ denotes the exponential map between the translation groups Lie algebra and the group itself.

This relation is very practical and helpful. Especially, we would like to extend this relation to hold also when mapped to the Hilbert space by the representation. The reason for this is that this gives a natural way to extrapolate the concept of energy and momentum as infinitesimal generators of time and space translation to the Hilbert space theory. On the Hilbert space this will lead us to the generators being possibly unbounded operators.

2.5.1 Spectral theory of self-adjoint operators

In this part we shall briefly state two classical results on the spectral theory of operators, namely the spectral theorem and Stone's theorem. For a more thorough account on this topic and the proofs of the results the reader is referred to [Sch12, Chapter 5 and 6] and [RS80, Chapter VIII, Sections 3 and 4]. The following content is based on these two references.

In this section let $\mathcal{H}$ be a fixed separable Hilbert space.

The spectral theorem is an enormously powerful tool in the analysis of bounded and unbounded operators. Among other things it provides the possibility to define functions of self-adjoint operators[24]. This is usually referred to as the functional calculus of (self-adjoint) operators.

In order to state the spectral theorem, we will need the concepts of spectral measures and integrals. As this is standard literature and there is good literature available as stated

[23] That is $\langle\phi, U_n\psi\rangle \to \langle\phi, U\psi\rangle$ for each $\phi, \psi \in \mathcal{H}$.
[24] It is even possible to extend the functional calculus to normal operators

above, these concepts will only be briefly sketched.

For a σ-algebra $\mathfrak{A}$ of a set Ω a spectral measure is a countably additive mapping E from $\mathfrak{A}$ to a space of orthogonal projections on $\mathcal{H}$ such that $E(\Omega)$ is the identity projection. A spectral measure gives rise to an ordinary measure for each $x \in \mathcal{H}$ via the mapping $x \mapsto \langle x, E(\cdot)x\rangle$. This measure will be denoted by $E_{x,x}$.

The spectral integral may then be constructed in a Lebesgue-type procedure as a mapping from the space of $\mathfrak{A}$-measurable functions to operators on $\mathcal{H}$. This procedure runs in three steps. On simple functions $f = \sum_{n=1}^{N} a_n 1_{M_n}$ for $N \in \mathbb{N}$, constants a_n and measurable subsets M_n of Ω, and where 1_M denotes the characteristic function of a set M, the integration with respect to a spectral measure E on Ω amounts to a mapping to $f \mapsto I(f) = \sum_n a_n E(M_n)$. In a second step we can extend this definition to bounded measurable functions as the simple functions approximate them (pointwise and uniformly) and the definition of the integral is continuous (with respect to the strong topology in $\mathcal{L}(\mathcal{H})$). Lastly, we may define the domain

$$D(I(f)) \equiv \{x \in \mathcal{H} : f \in L^2(\Omega, E_{x,x})\}. \tag{2.83}$$

On this domain we can extend the definition further by noting that for each f there is a sequence of measurable sets M_n such that $f\restriction_{M_n} \equiv f\,1_{M_n}$ is bounded for each n and that $I(f\restriction_{M_n})$ is a strongly converging sequence. The limiting operator of the sequence will be the definition of the spectral integral. We will formally write

$$I(f) = \int f(\lambda) dE(\lambda). \tag{2.84}$$

Proposition 2.32. *For measurable functions f, g and $\alpha, \beta \in \mathbb{C}$, we have the following properties*

(a) $I(\alpha f + \beta g) = \overline{\alpha I(f) + \beta I(g)}$

(b) $I(\bar{f}) = I(f)^\dagger$

(c) $I(fg) = \overline{I(f)I(g)}$

(d) $D(I(f)I(g)) = D(I(fg)) \cap D(I(g))$.

*(e) Restricting I to only bounded measurable functions, I actually becomes a norm continuous *-algebra homomorphism, i.e., in (a) and (c) there are no closures to be taken.*

(f) For any polynomial $p \in \mathbb{C}[t]$ we have $I(p(f)) = p(I(f))$.

The main structural result is then the spectral theorem. It is stating that actually any self-adjoint operator is given by the spectral integral with respect to a uniquely defined spectral measure:

Theorem 2.33. *For each self-adjoint operator A on $\mathcal{H}$ there is a unique spectral measure E_A on the Borel σ-algebra $\mathfrak{B}_\mathbb{R}$ of $\mathbb{R}$ such that*

$$A = \int \lambda \, dE_A(\lambda). \tag{2.85}$$

The combination of these results justifies the definition of a 'function' of a self-adjoint operator $f(A) \equiv \int f(\lambda)\, dE_A(\lambda)$. As a spectral integral this 'function' behaves according to the calculus defined in Proposition 2.32. This is referred to as the functional calculus of self-adjoint operators.

2.5.2 Strong commutativity and functional calculus

In the case of the translation group and the energy-momentum operator it will be necessary to generalize this result to finitely many parameters. We will see that the components of the energy-momentum operator will necessarily be strongly commuting:

Definition 2.34. Two (possibly unbounded) self-adjoint operators are strongly commuting if, and only if, their associated spectral measures commute or, equivalently, if their associated unitary groups (in the sense of Stone's theorem, see below in Subsection 2.5.3) commute.

Remark 2.35. The forward implication follows immediately from the functional calculus of (bounded) functions of self-adjoint operators. The backward implication needs some work and can be found e.g. in [RS80, Theorem VIII.13].

The spectral theorem can then be generalized to the following form:

Theorem 2.36. *Let $A \equiv (A_1, ..., A_n)$ be a finite tuple of strongly commuting self-adjoint operators A_j on $\mathcal{H}$. Then there is a unique spectral measure E_A on the Borel σ-algebra $\mathfrak{B}_{\mathbb{R}^n}$ such that*

$$A_j = \int_{\mathbb{R}^n} \lambda_j \, dE_A(\lambda_1, ..., \lambda_n), \quad j = 1, ..., n. \tag{2.86}$$

sketched. For each $j = 1, ..., n$ there exists a unique spectral measure E_j such that $A_j = \int_{\mathbb{R}} \lambda \, dE_j(\lambda)$.

Existence of E_A: For two spectral measures E, F on $\mathfrak{B}_{\mathbb{R}}$ the spectral measure $E \wedge F$ on $\mathfrak{B}_{\mathbb{R}^2}$ is defined by

$$E \wedge F\,(M \times N) = E(M)E(N), \quad M, N \subset \mathfrak{B}_{\mathbb{R}} \tag{2.87}$$

and its algebraic (according to the σ-algebra structure) extension to $\mathfrak{B}_{\mathbb{R}^2}$.
Defining $E = \bigwedge_{j=1}^n E_j$ then gives

$$\int_{\mathbb{R}^n} \lambda_j dE(\lambda_1, ..., \lambda_n) = E_1(\mathbb{R})...E_{j-1}(\mathbb{R}) \int_{\mathbb{R}} \lambda_j dE_j(\lambda_j) E_{j+1}(\mathbb{R})...E_n(\mathbb{R}) = \int_{\mathbb{R}} \lambda_j dE_j(\lambda_j) = A_j. \tag{2.88}$$

Uniqueness of E_A: Let there be a spectral measure E on $\mathfrak{B}_{\mathbb{R}^n}$ such that $\int \lambda_j dE(\lambda) = A_j$ for each $j = 1, ..., n$, where we shortened $\lambda \equiv (\lambda_1, ..., \lambda_n)$. Then we can define for each $M \in \mathfrak{B}_{\mathbb{R}}$ that $E_j(M) \equiv E(\mathbb{R}^{j-1} \times M \times \mathbb{R}^{n-j})$ and observe that $\int_{\mathbb{R}} \lambda \, dE_j(\lambda) = A_j$ as well as $E = \bigwedge_{j=1}^n E_j$. The only non-uniqueness could come from the order of the E_j's in the wedge-product. But this is inessential due to the strong commutativity of the A_j's. We will not go into the proof of this inessentiality. ∎

2.5.3 Unitary groups and Stone's theorem

A one-parameter unitary group is a strongly continuous group homomorphism from $\mathbb{R}$ to $\mathcal{U}(\mathcal{H})$, where $\mathcal{U}(\mathcal{H})$ denotes the space of unitary operators on $\mathcal{H}$. It should be noted again that for unitary operators the weak and strong topology are equivalent, but strong continuity is written to distinguish it from norm continuity (see Remark 2.31).

The main result is then Stone's theorem which establishes that for any one-parameter unitary group there is a self-adjoint operator generating the group.

Theorem 2.37. *Let U be a map $\mathbb{R} \to \mathcal{U}(\mathcal{H})$. Then the following two statements are equivalent:*

(a) U is a continuous group homomorphism

(b) There exists a possibly unbounded self-adjoint operator A on $\mathcal{H}$ such that $U(t) = e^{itA}$ for each $t \in \mathbb{R}$.

Proof. For a proof the reader is referred to Appendix D. ■

In the case of the translation group and the energy-momentum operator it will be necessary to generalize this statement to finitely many parameters and accordingly finitely many generators. These generators will be strongly commuting. The generalized version of Stone's Theorem reads

Theorem 2.38. *Let U be a map $\mathbb{R}^n \to \mathcal{U}(\mathcal{H})$. Then the following two statements are equivalent:*

(a) U is a continuous group homomorphism.

(b) There exist self-adjoint operators A_j, which generate the one-parameter subgroups $U(0,...,0,t_j,0,...,0)$ and which mutually strongly commute.

Remark 2.39. With the equivalence of the two notions of strong commutativity of self-adjoint operators already established, this theorem is an almost direct consequence of the one-parameter case of Stone's theorem.

2.5.4 The energy-momentum operator

Suppose a strongly continuous unitary representation T of the translation subgroup (identified with $\mathbb{M}$) on a Hilbert space $\mathcal{H}$. By Stone's theorem (see section above) there exist four strongly commuting self-adjoint generators P_μ that generate this subgroup. In other words, for each $a \in \mathbb{M}$ we have

$$T(a) = e^{ia^\mu P_\mu}. \tag{2.89}$$

In physical applications the generators P_μ together are referred to as energy-momentum operator. The joint spectrum of the components P_μ, or in other words the spectrum of the energy-momentum operator P, is defined by the support of the associated spectral measure. We will denote it by $\sigma(P) \equiv \sigma(P_0, P_1, P_2, P_3)$. The spectrum condition states that the spectrum of the energy momentum operator is contained in the closed upper light cone, i.e.,

$$\sigma(P) \subset \bar{V}^+. \tag{2.90}$$

This condition amounts to a non-negative energy spectrum and a causal momentum spectrum (momenta will not be spacelike).

2.6 Wightman and Strocchi-Wightman quantum field theory

The Wightman framework of quantum field theory provides one of the standard axiomatic approaches to quantum field theory. The advantage of this framework is that its axioms provide a close connection to the perturbative approach and to the physical properties of QFT. Unfortunately, this framework does not seem suitable for describing gauge theories, in particular formulations using local and/or covariant gauges, as was already emphasized in the introduction. For these gauges it seems to be inevitable (and later we will proof that for QED in covariant Gupta-Bleuler gauge it actually is) to have recourse to indefinite metric Hilbert spaces, also called Krein spaces. The extension of the Wightman framework to Krein spaces will be referred to as Strocchi-Wightman framework.

2.6.1 Quantum fields and Wightman QFT

The foundation underlying quantum field theory is relativistic quantum theory. Relativistic quantum theory can be thought of as a combination of the principles of special relativity and quantum mechanics. It should generally fit into this picture:

Definition 2.40. A *relativistic quantum theory (rel. QT)* is given by a tuple $(\mathcal{H}, U)$ consisting of a separable Hilbert space $\mathcal{H}$ and a continuous unitary representation of the Poincaré group U acting on $\mathcal{H}$ by

$$\psi \mapsto U(g)\psi, \quad \psi \in \mathcal{H} \tag{2.91}$$

for each $g \in \mathcal{P}_+^\uparrow$. The relativistic quantum theory satisfies the *spectrum condition* if, and only if,

$$\sigma(P) \subset \bar{V}^+ \tag{2.92}$$

where P is the energy-momentum operator corresponding to U.
A vector $\Omega \in \mathcal{H}$ which is invariant under U, i.e., $U(g)\Omega = \Omega \, \forall g \in \mathcal{P}_+^\uparrow$[25], and normalized to 1 is called *vacuum vector*. The vacuum vector is called *unique*, if and only if, the invariance and normalization) condition fix the vector up to a phase factor.

Remark 2.41. As was shown in Proposition 2.20 the single- and double-valued representations of the (proper) Lorentz group stand in one-to-one correspondence to the (single-valued) representations of its universal covering group, $SL(2, \mathbb{C})$. As the Poincaré group is the semidirect product of the Lorentz group and the four-dimensional translation group (which is connected and simply-connected), the above statement applies to the Poincaré

[25]It should be noted that it is enough to demand that Ω is invariant under the translation subgroup. This already implies the invariance under the full Poincaré group. See e.g. [RS75, p. 63, below Property 3].

group, too. Its universal covering is the Poincaré spinor group, given by the semidirect product of $SL(2,\mathbb{C})$ and the translation group, and denoted by $\tilde{\mathcal{P}}_+^\uparrow$. Thus we will use the Poincaré and the Poincaré spinor group interchangeably as it suits the presentation best.

A possible choice for the fundamental objects of a quantum field theory (QFT) are the quantum fields. Understanding quantum field theory as a relativistic theory we will only talk about relativistic quantum fields here and omit the term relativistic after the following definition:

Definition 2.42. A *(relativistic) quantum field* embedded in a relativistic quantum theory $(\mathcal{H}, U)$ is a tuple (D, Φ) consisting of a dense linear subspace $D \subset \mathcal{H}$ and a finite collection of operator-valued tempered[26] distributions $\{\Phi_l\}_{l=1,\dots,m}$ on $\mathbb{M}$ such that D is a common dense and invariant domain to all the operators $\Phi_l(u)$, $\Phi_l(u)^\dagger$, and $U(g)$ where $l = 1, \dots, m$, $u \in \mathcal{S}$, and $g \in \tilde{\mathcal{P}}_+^\uparrow$ and $\Phi_l(u)^\dagger$ denotes the adjoint operator of $\Phi_l(u)$. The Φ_l are referred to as *field components* of the field Φ.

Note that the properties of D are of quite technical nature and are there to guarantee that the polynomials of fields, its adjoints, and of unitary transformations can be formed without domain issues. Moreover, it should be noted that by the embedding into a relativistic theory, quantum fields (or rather their smeared components $\Phi_l(u)$) transform like

$$\Phi_l(u) \mapsto U(g)\Phi_l(u)U(g)^{-1}, \quad g \in \tilde{\mathcal{P}}_+^\uparrow \tag{2.93}$$

That $U(g)\Phi_l(u)U(g)^{-1}$ is again a densely defined operator on D is implied by the invariance of D under $U(g)$ for each $g \in \tilde{\mathcal{P}}_+^\uparrow$. Very important additional properties of quantum fields are:

Definition 2.43. A quantum field (D, Φ) (embedded in a rel. QT $(\mathcal{H}, U)$) is referred to be

- *hermitian*, if for each $u \in \mathcal{S}$ the smeared field components $\Phi_l(u)$ and $\Phi_l(\bar{u})^\dagger\restriction_D$ coincide.

- *local*, if for each $u, u' \in \mathcal{S}$ such that the supports of u and u' are spacelike separated the commutator or anticommutator of the smeared field components vanishes on D, i.e.,

  $$[\Phi_l(u), \Phi_{l'}(u')]_\mp \psi \equiv (\Phi_l(u)\Phi_{l'}(u') \pm \Phi_{l'}(u')\Phi_l(u))\,\psi = 0, \quad l, l' = 1, \dots, m \tag{2.94}$$

  for each $\psi \in D$.

- *(Poincaré-)covariant*, if (on D)

  $$U(a, \underset{\sim}{\Lambda})\Phi(u)U(a, \underset{\sim}{\Lambda})^{-1} = V(\underset{\sim}{\Lambda}^{-1}) \cdot \Phi(u_{a,\Lambda(\underset{\sim}{\Lambda})}), \quad (a, \underset{\sim}{\Lambda}) \in \tilde{\mathcal{P}}_+^\uparrow,\, u \in \mathcal{S} \tag{2.95}$$

  where $u_{a,\Lambda}(x) = u(\Lambda^{-1}(x-a))$ and $V(\underset{\sim}{\Lambda})$ is a complex or real finite dimensional matrix representation $SL(2,\mathbb{C})$. Note that the number of field components of Φ coincides with the dimension of V. The field is referred to be a bosonic/fermionic field, if and only if, $V(-1) = \pm 1$. In this case the $SL(2,\mathbb{C})$-representation gives rise to a single/double-valued representation of the proper Lorentz group.

[26] There are other choices available than Schwartz space as space of test functions. See e.g. the original paper [WG64].

The definitions above specify properties a single quantum field may have. For a (quantum field) theory of a physical system there is usually a collection of quantum fields. The Wightman framework is one of the standard frameworks to incorporate such a collection of quantum fields in a consistent and complete theory (see below the definition for further explanation).

Definition 2.44. A *Wightman quantum field theory* is a 5-tuple $(\mathcal{H}, U, \Omega, D, \{\Phi^{(\kappa)}\}_{\kappa\in I})$ consisting of a relativistic quantum theory $(\mathcal{H}, U)$ which is subject to the spectrum condition and for which there is a unique vacuum state $\Omega \in \mathcal{H}$ together with a collection of local and covariant quantum fields $\{(D, \Phi^{(\kappa)})\}_{\kappa\in I}$ embedded in $(\mathcal{H}, U)$. Moreover, it is required that $\Omega \in D$, that the collection of quantum fields contains also their adjoints[27] and that

- *locality*: Any two fields either commute or anticommute under spacelike separation. This means that for each $\kappa, \kappa' \in I$ and for each $u, u' \in \mathcal{S}(\mathbb{M})$ such that the supports of u and u' are spacelike separated either the commutator or the anticommutator of the smeared field components $\Phi_l^{(\kappa)}(u)$ and $\Phi_{l'}^{(\kappa')}(u')$ vanishes on D. If those are commuting or anticommuting depends only on κ and κ'.
- *cyclicity*: Ω is cyclic with respect to $\{\Phi^{(\kappa)}\}_{\kappa\in I}$, i.e., the set
$$D_0 \equiv \{\Phi^{(\kappa_1)}(u_1)...\Phi^{(\kappa_m)}(u_m)\Omega \,:\, m \in \mathbb{N}, u_1, ..., u_m \in \mathcal{S}, \kappa_1, ..., \kappa_m \in I\} \subset D \subset \mathcal{H} \tag{2.96}$$
is dense in $\mathcal{H}$

are fulfilled.

Remark 2.45. Note that the requirement of locality for $\kappa = \kappa'$ actually requires each of the quantum fields to be local on its own. Thus it is duplicate information that the field collection consists of local fields only. It is nonetheless written down for ease of reading.

In the following we will usually omit the term quantum and just use the term field, as we will almost exclusively deal with quantum (in contrast to classical) fields. A field that is part of a Wightman QFT is termed Wightman field. For a Wightman field associated to a given Wightman QFT $(D, ...)$ it is actually sufficient to write Φ keeping the domain implicit, as it will always be D. Also for other fields we will often keep the domain implicit. The most important Lorentz-representations V are the scalar, vector- and 2-tensor-representation corresponding to $\mathcal{D}^{(0,0)}, \mathcal{D}^{(\frac{1}{2},\frac{1}{2})}$ and $\mathcal{D}^{(1,1)}$ as described in Subsection 2.3.2.

Note that the definitions for quantum fields and the formulation of Wightman QFT were very brief and did not leave much space for explanations. For a good introduction to the Wightman framework see [SW80]. For a thorough discussion of motivation for the Wightman framework and possible alternative choices in the definition the reader is referred to the original paper [WG64]. Other accounts on the Wightman framework can be found in the books [RS75, Section IX.8] and [BLOT90, Chapter 8] and the lecture notes [Dyb18].

[27]In this regard we define $\Phi_l^{(\bar\kappa)}(u) \equiv \Phi_l^{(\kappa)}(\bar u)^\dagger\restriction_D$ and suppose that $\bar\kappa \in I$ for each $\kappa \in I$. Note that this requires that D is also an invariant domain for the adjoint operator. But this assumption was included in the definition of a quantum field.

2.6.2 Wightman distributions

As is well known, from these (Wightman) quantum fields we can form the so-called Wightman distributions. Suppose a fixed given Wightman QFT $(D, \Phi^{(\kappa)}, \mathcal{H}, U, \Omega)$[28]. We can define the Wightman distributions as follows

Definition 2.46. To each fixed $m \in \mathbb{N}$ and appropriate indices $\kappa_1, ...\kappa_m, l_1, ..., l_m$ we can associate a multilinear and separately continuous map

$$w^{(\kappa_1...\kappa_m)}_{l_1...l_m} : \mathcal{S}(\mathbb{M}) \times ... \times \mathcal{S}(\mathbb{M}) \to \mathbb{C} \quad (u_1, ..., u_m) \mapsto \langle \Omega, \Phi^{(\kappa_1)}_{l_1}(u_1)...\Phi^{(\kappa_m)}_{l_m}(u_m)\Omega \rangle \,. \tag{2.97}$$

which we will call *Wightman distribution*.

These maps are called distributions as for each argument separately they form tempered distributions. One may define

$$w^{(\kappa_1...\kappa_m)}_{l_1...l_m}(u_1 \otimes ... \otimes u_m) \equiv w^{(\kappa_1...\kappa_m)}_{l_1...l_m}(u_1, ..., u_m) \tag{2.98}$$

and extend this definition by linearity and continuity to all of $\mathcal{S}(\mathbb{M}^m)$.[29] By the nuclear theorem[30] this extension to $\mathcal{S}(\mathbb{M}^m)$ is unique. We will call both versions of the map $w^{(\kappa_1...\kappa_m)}_{l_1...l_m}$ a Wightman distribution or m-point(-correlation) function. In this way we have the freedom to choose the version which makes the notation easiest, but do not have any problems arising because of the uniqueness of the extension.

The properties of the Wightman QFT directly imply some similar properties for the Wightman distributions.[31] For instance, we can make use of the translation invariance of the vacuum state to gain a simplified manifestly translation-invariant version of the Wightman distributions:

Theorem 2.47. *To each Wightman distribution $w \equiv w^{(\kappa_1...\kappa_m)}_{l_1...l_m}$ there is a tempered distribution W on $\mathbb{M}^{m-1}$ such that in generalized function notation, introducing $\xi_j = x_j - x_{j+1}$, $j = 1, ..., n-1$,*

$$w(x_1, ..., x_m) = W(\xi_1, ..., \xi_{m-1}). \tag{2.99}$$

Proof. By translation invariance of the fields and the invariance of the vacuum vector the distribution w may depend only on a set of $m-1$ translation-invariant variables. The relative coordinates ξ_j, $j = 1, ..., m-1$ are a natural choice. For the details the reader is referred to [SW80, Chapter 2-1, pp. 38]. ■

It is well known that a certain set of properties may serve as an equivalent formulation of a Wightman QFT with the Wightman distributions as fundamental objects instead of quantum fields.[32] As only the low-order correlation functions (basically only two-point-functions) are of great importance to this thesis, we will omit this general formulation here and give situation-specific explanations when they occur.

[28]Note that so far in 1+3 dimensions there are only constructions of free, i.e., non-interacting, Wightman QFTs available

[29]Note that the space $\mathcal{S}(\mathbb{M}) \otimes ... \otimes \mathcal{S}(\mathbb{M}$ (m factors) is a dense subspace of $\mathcal{S}(\mathbb{M}^m)$

[30]See e.g. Theorem V.12 in [RS80, Appendix to V.3].

[31]See e.g. [SW80, Chapter 3-3].

[32]This is often referred as the Wightman reconstruction theorem, for the original work see [Wig56]

2.6.3 Strocchi-Wightman QFT

The framework of Wightman QFT is restricted to the description of observable fields for which there are strong physical constraints in order to ensure the principles of special relativity and quantum theory. In many physical models, however, it proves to be beneficial using fundamental variables that are not observable. For instance in classical electrodynamics the introduction of the gauge field A_μ (vector potential) provides a description of electrodynamics in a simple form although A_μ itself is not observable. In general these fundamental variables are much less restricted by physical constraints. The vector potential for instance may be non-local, non-covariant and/or does not need to be an operator on the physical state space. In this view it seems natural to introduce a broader, so-called virtual, state space, in which these fundamental non-observable fields act and on which there are less-restricted conditions imposed. A standard choice for local gauge theories[33] is the Strocchi-Wightman framework to quantum field theories. This framework allows for a metric which is indefinite on the virtual state space. As is the aim to show in this thesis this actually is necessary in order to describe QED in covariant Gupta-Bleuler gauge. Let us start by introducing the notion of an indefinite metric Hilbert space, also referred to as Krein space.

Let us start with some very basic nomenclature. Let V be a linear space equipped with a hermitian sesquilinear form, ω on V. The space is referred to as a *space with indefinite metric* if, and only if, the scalar square $\omega(u,u)$ can take all real values when ranging through $u \in V$. In the case that the scalar square can only take non-negative values in $\mathbb{R}$ the space is referred to as a *space with non-negative metric.* If, in addition, $\omega(u,u) = 0$ implies $u = 0$ then the space is referred to as a space with *positive definite metric.* In the case that ω is non-degenerate, we refer to ω as a *scalar product.* Note that a scalar product space with a non-negative metric is automatically a space with positive definite metric.[34] In the positive definite case ω is in the literature usually also referred to as an inner product (turning V into an inner product or pre-Hilbert space). In this regard in the indefinite case ω is sometimes also referred to as an *indefinite inner product.*

Definition 2.48. A *Krein space* is a triple $(\mathcal{H}, (\cdot,\cdot), \langle\cdot,\cdot\rangle)$ consisting of a Hilbert space $(\mathcal{H}, (\cdot,\cdot))$ and a scalar product $\langle\cdot,\cdot\rangle$ on $\mathcal{H}$ which is continuous with respect to the Hilbert topology defined by $(\cdot,\cdot)$.

Remark 2.49. The continuity of $\langle\cdot,\cdot\rangle$ implies the existence of a bounded self-adjoint[35] operator η with bounded inverse such that $\langle\phi,\psi\rangle = (\phi,\eta\psi)$ for each $\phi,\psi \in \mathcal{H}$. By a redefinition of the Hilbert product one may achieve $\eta^2 = \mathbb{1}$ such that the eigenvalues of η become ± 1. The case that all eigenvalues of η are $+1$ or that all are -1 reduces to the case of an ordinary Hilbert space.

Remark 2.50. The Hilbert product plays an auxiliary role and turns $\mathcal{H}$ into a topological space, whereas the indefinite scalar product $\langle\cdot,\cdot\rangle$ is the physically relevant one. Thus we

[33]These are theories with a gauge-dependent fundamental variable/field which is nevertheless required to satisfy the locality-axiom

[34]Let ω be non-degenerate and its scalar square be non-negative. Then suppose $\omega(u,u) = 0$ for some $u \in V$. Then we have for each $\lambda \in \mathbb{C}$ and each $v \in V$ that $0 \leq \omega(\lambda u + v, \lambda u + v) = \lambda^2\omega(u,u) + 2\mathrm{Re}\,\lambda\omega(u,v) + \omega(v,v) = 2\mathrm{Re}\,\lambda\omega(u,v) + \omega(v,v)$. Taking $\lambda \to \pm\infty$ and $\lambda \to \pm i\infty$ shows that $\omega(u,v) = 0$ for all $v \in V$. Thus if ω is non-degenerate one necessarily gets $u = 0$.

[35]with respect to the Hilbert product $(\cdot,\cdot)$.

will also refer to the latter as the physical (scalar) product. Important to note is that the continuity of the indefinite scalar product with respect to the Hilbert product uniquely fixes the Hilbert topology. In other words: Suppose another Hilbert product with respect to which the indefinite scalar product is continuous. Then the induced norm of the new Hilbert product will be equivalent to the induce norm of the old one.[36] As the Hilbert product is auxiliary, it will be omitted in the notation. Thus a Hilbert space (as before) will be denoted by just the space $\mathcal{H}$ and a Krein space will be denoted as a tuple $(\mathcal{H}, \langle\cdot,\cdot\rangle)$.

When passing from Hilbert spaces to Krein spaces there will be some concepts within the Wightman framework that have to be generalized. For instance the concepts of self-adjointness and unitarity. But as these generalizations are straight-forward we will omit the definitions and refer the reader to [BLOT90, Chapter 10.1, in particular pp. 419].

In order to obtain a probability interpretation for a generalized Wightman framework on Krein spaces it is important to identify a physical Hilbert space constructed from the Krein space on which most parts of ordinary Wightman theory are recovered. This identification will follow a standard procedure for the construction of Hilbert spaces. For a Krein space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ define the physical subspace $\mathcal{H}' \equiv \{\psi \in \mathcal{H} |\, \langle\psi,\psi\rangle \geq 0\}$, $\mathcal{H}'' \equiv \{\psi \in \mathcal{H} |\, \langle\psi,\psi\rangle = 0\}$ and the physical Hilbert space $\mathcal{H}_{\mathrm{ph}} \equiv \overline{\mathcal{H}'/\mathcal{H}''}^{cpl}$ as the completion of the pre-Hilbert space $\mathcal{H}'/\mathcal{H}''$.[37] The elements of $\mathcal{H}_{\mathrm{ph}}$ are equivalence classes $[\psi] \equiv \psi + \mathcal{H}''$, $\psi \in \mathcal{H}'$.

The definition of a relativistic quantum theory (see Definition 2.40) may now be generalized by replacing the Hilbert space by a Krein space and the unitary representation by a $\langle\cdot,\cdot\rangle$-unitary representation that leaves the physical subspace $\mathcal{H}'$ invariant. By the normalization condition a vacuum vector is automatically an element of $\mathcal{H}'$ with non-zero norm. Thus Ω will be a non-zero element of the physical Hilbert space. Note that the $\langle\cdot,\cdot\rangle$-unitary representation U on $\mathcal{H}$ induces a unitary representation U_{ph} on $\mathcal{H}_{\mathrm{ph}}$ by specifying

$$U_{\mathrm{ph}}(g)[\psi] \equiv [U(g)\psi], \quad \psi \in \mathcal{H}' \tag{2.100}$$

for each $g \in \tilde{\mathcal{P}}_+^\uparrow$. That the definition of U_{ph} is not depending on the choice of representative ψ of $[\psi]$ follows from the invariance of $\mathcal{H}''$ under $\langle\cdot,\cdot\rangle$-unitary transformations (as they leave the $\langle\cdot,\cdot\rangle$-norm invariant). Thus for each indefinite metric rel. QT $(\mathcal{H}, \langle\cdot,\cdot\rangle, U)$ there is a naturally induced (physical) relativistic quantum theory $(\mathcal{H}_{\mathrm{ph}}, U_{\mathrm{ph}})$ and if there are vacuum vectors in the indefinite metric rel. QT then they are also vacuum vectors of the physical rel. QT. An indefinite metric rel. QT is said to fulfil the spectrum condition if, and only if, $(\mathcal{H}_{\mathrm{ph}}, U_{\mathrm{ph}})$ does. In the same manner, a vacuum vector of an indefinite metric rel. QT is said to be unique if, and only if, it is unique (up to a phase factor) in the induced physical rel. QT.

For the definition of an indefinite metric quantum field the relativistic quantum theory $(\mathcal{H}, U)$ is replaced by an indefinite metric relativistic quantum theory $(\mathcal{H}, \langle\cdot,\cdot\rangle, U)$ (as

[36]See also [BLOT90, Proposition 10.1].

[37]Note here that on the pre-Hilbert space $\mathcal{H}'/\mathcal{H}''$ there is the norm of the underlying auxiliary Hilbert space structure and there is the norm induced by the physical product. As both are equivalent no ambiguity in the term completion arises. For reference there is an exercise problem [AI89, §2, Exercise 2]

described above) and the domain D will be again a dense subspace of $\mathcal{H}$ equipped with an additional hermitian form $\langle \cdot, \cdot \rangle$ (turning it to an indefinite metric space).[38]
The according generalization of the Wightman framework for QFT is then

Definition 2.51.
A *Strocchi-Wightman quantum field theory* is a 6-tuple $(\mathcal{H}, \langle \cdot, \cdot \rangle, U, \Omega, D, \{\Phi^{(\kappa)}\}_{\kappa \in I})$ consisting of an indefinite metric relativistic quantum theory $(\mathcal{H}, \langle \cdot, \cdot \rangle, U)$ which is subject to the spectrum condition and for which there is a unique vacuum state $\Omega \in \mathcal{H}'$ together with a finite collection of local and covariant indefinite metric quantum fields $\{(D, \Phi^{(\kappa)})\}_{\kappa \in I}$ embedded in $(\mathcal{H}, \langle \cdot, \cdot \rangle, U)$. Moreover, it is required that $\Omega \in D$, that the collection of quantum fields contains also their adjoints and that

- *locality*: Any two fields either commute or anticommute under spacelike separation. This means that for each $\kappa, \kappa' \in I$ and for each $u, u' \in \mathcal{S}(\mathbb{M})$ such that the supports of u and u' are spacelike separated either the commutator or the anticommutator of the smeared field components $\Phi_l^{(\kappa)}(u)$ and $\Phi_{l'}^{(\kappa')}(u')$ vanishes on D. If those are commuting or anticommuting depends only on κ and κ'.

- *cyclicity*: Ω is cyclic with respect to $\{\Phi^{(\kappa)}\}_{\kappa \in I}$, i.e., the set

$$D_0 \equiv \{\Phi^{(\kappa_1)}(u_1)...\Phi^{(\kappa_m)}(u_m)\Omega \,:\, m \in \mathbb{N}, u_1, ..., u_m \in \mathcal{S}, \kappa_1, ..., \kappa_m \in I\} \subset D \subset \mathcal{H} \tag{2.101}$$

  is dense in $\mathcal{H}$

are fulfilled.

2.7 Quantum electrodynamics

We have already seen that there are non-perturbative frameworks for QFT, as the Wightman and the Strocchi-Wightman setting, which are consistent (the free case can be constructed) and are believed to be close to what should be non-perturbative QFT. On the other hand still there is no construction of an interacting QFT in 1+3 dimensions. In this regard also the fully non-perturbative construction of QED is an open question and is believed to be especially hard for QED, as it suffers severe infrared problems. Therefore in this section we will be content with shortly outline what quantum electrodynamics should look like (what are the constraints?) and what framework we will pick for this thesis.

2.7.1 Constraining results and properties

It should be noted that this section is characterized by rather rough and heuristic arguments. The reason for this is that there is a tremendous amount of results and thoughts

[38]One might wonder about adding the technical assumption that $(D \cap \mathcal{H}')/\mathcal{H}''$ is dense in $\mathcal{H}_{\text{ph}}$ in order to ensure that the indefinite metric quantum field is still densely defined on the physical Hilbert space. But this assumption is not required as it is a consequence of the continuity of $\langle \cdot, \cdot \rangle$: Let $\psi \in \mathcal{H}'$ be an arbitrary representative of a state $0 \neq [\psi] \in \mathcal{H}_{\text{ph}}$. Then there exists a sequence $(\psi_n)_n \subset D$ such that $\psi_n \to \psi$ in $\mathcal{H}$ for $n \to \infty$. Then (by the continuity of $\langle \cdot, \cdot \rangle$) $\langle \psi_n, \psi_n \rangle$ converges to $\langle \psi, \psi \rangle$ which is positive by assumption. Thus there exists $N \in \mathbb{N}$ such that $\langle \psi_n, \psi_n \rangle > 0 \, \forall n \geq N$ and the sequence $(\psi_{n+N})_n \subset D \cap \mathcal{H}'$ converges to ψ and induces a corresponding convergent sequence in $\mathcal{H}_{\text{ph}}$.

that went into the formulation of quantum electrodynamics and stating all the precise results would go much beyond the scope of this document. Also in some cases the arguments have to be heuristic when physics enters the game.

The confining properties in the search for (non-perturbative) quantum electrodynamics are the reproduction of classical electrodynamics in the classical regime of the theory and the reproduction of perturbative quantum electrodynamics up to an experimentally tested degree of precision. With only this at hand it seems natural to take them as a starting point. As a very distinguishing property there are the (classical) observables of electrodynamics, the Maxwell tensor $F_{\mu\nu}$ and the electromagnetic current density J_μ subject to the (classical) equations of motion

$$\partial_\mu F^{\mu\nu} = J^\nu \quad \text{and} \quad \partial_{[\mu} F_{\rho\sigma]} = 0. \tag{2.102}$$

Note that the current J is automatically conserved, i.e., $\partial^\mu J_\mu = 0$, by the equations of motion and the antisymmetry of F. In the classical theory the Poincaré theorem infers from the closedness of a covariant two-form F (i.e., $dF = 0$) the existence of a covariant one-form A satisfying $F = dA$ or in index notation $F_{\mu\nu} = \partial_{[\mu} A_{\nu]}$. This vector potential A, although not observable[39], may serve as a fundamental field of the theory instead of F. When doing so one achieves an equivalent formulation of electrodynamics in terms of the field A:

$$\Box A_\mu - \partial_\mu \partial A = J_\mu \quad \text{and} \quad F_{\mu\nu}(A) = \partial_{[\mu} A_{\nu]}. \tag{2.103}$$

Eventually, it is our aim here to construct an interacting theory of electrons (and positrons)[40]. Thus the current density J has to take the specific form of the Dirac current:

$$J_\mu = \bar{\Psi} \gamma_\mu \Psi \tag{2.104}$$

with the classical Dirac-spinor field[41] Ψ representing the electron, the conjugate field $\bar{\Psi} \equiv \Psi^* \gamma_0$ representing the positron and the gamma matrices γ_μ which are used to implement the Dirac representation in the space of Lorentz-tensors.[42] The equations of motion for Ψ and $\bar{\Psi}$, respectively, are Dirac equations (with the same mass, but conjugated to each other). The form of J may be motivated as the simplest term obeying the properties of charge symmetry, (vector) covariance and that it is built from the spin-1/2-fields Ψ and $\bar{\Psi}$. It is also the only term leading to a renormalizable quantum field theory. The ultimate motivation, though, is its great agreement with experiments.[43]

In the classical case the formulation of the theory in terms of only F, Ψ and $\bar{\Psi}$ is possible, but complicated and impractical. In the quantum case it becomes nearly intractable. The

[39]The vector potential A is subject to a gauge ambiguity. When transforming $A_\mu(x) \to A_\mu(x) + \partial_\mu \epsilon(x)$ by an arbitrary real-valued twice differentiable function ϵ on $\mathbb{M}$, the corresponding physical field F does not change.

[40]For the sake of simplicity we will not talk about any other electrically charged particle families like myons etc.

[41]Mathematically speaking this is a (at least) twice differentiable function on Minkowski space taking values that transform according to the $SL(2,\mathbb{C})$-representation $\mathcal{D}^{(1/2,0)} \oplus \mathcal{D}^{(0,1/2)}$

[42]See e.g. [BLOT90, Chapter 7 Appendix E] for a definition of gamma matrices and the Dirac representation.

[43]See [Ste00, Chapter 3, in particular p. 24].

reason for this is that there is only a very limited class of models for which we are more or less certain of how to construct renormalizable quantum field theories from them[44] and these complicated expressions do not fall into this class. Thus we are driven to implement an interacting theory of electron with the help of an unobservable vector field A.

That this field is unobservable can easily be seen by the well-known fact that the (classical) equations of motion leave space for an ambiguity of the vector potential. For an arbitrary twice differentiable real-valued field G the transformations

$$A_\mu \mapsto A'_\mu = A_\mu + \partial_\mu G, \quad \Psi \mapsto e^{ieG}\Psi, \quad \bar{\Psi} \mapsto e^{-ieG}\bar{\Psi} \tag{2.105}$$

leave the equations of motion $\partial_\mu F^{\mu\nu} = \bar{\Psi}\gamma^\nu\Psi$ and the Dirac equations of Ψ and $\bar{\Psi}$ invariant. It is still an open question, however, if this merely a situation specific to the classical case or if these gauge symmetries (they occur in many other models) should have a quantum analogue. As it is hard to gain control over these gauge transformations in the quantum setting, the usual way to construct the quantum model is to pick a certain representative of the gauge field from the orbit of gauge-equivalent fields via a certain gauge condition. This was explained already in the introduction.

There are now two more constraining properties of free QED that will be important in the aftermath: The structure of the commutator two-point-function of $F_{\mu\nu}$ and the triviality of a formulation of free QED including a fundamental either local or covariant vector field A which is subject to Maxwell's equations.

Concerning the first property, we know that F is an observable field and therefore has to obey the principles of locality and covariance. For free QED let us define the commutator expectation value of F as

$$B_{\mu\nu\rho\sigma}(\xi) \equiv \langle \Omega, [F_{\mu\nu}(x), F_{\rho\sigma}(y)]\,\Omega\rangle \tag{2.106}$$

where $\xi = x - y$.[45] Then by the fact that F is a free field, by Lorentz covariance, by locality (B vanishes when applied to test functions with support in spacelike regions) and by the spectral condition it necessarily has the form

$$B_{\mu\nu\rho\sigma}(\xi) = r d_{\mu\nu\rho\sigma} D(\xi) \tag{2.107}$$

for some real constant r and the tensor $d_{\mu\nu\rho\sigma} = -\eta_{\nu\sigma}\partial_\mu\partial_\rho + \eta_{\nu\rho}\partial_\mu\partial_\sigma - \eta_{\mu\rho}\partial_\nu\partial_\sigma + \eta_{\mu\sigma}\partial_\nu\partial_\rho$.[46] This form for B may also be derived by explicit construction of the free theory corresponding to F. The choice $r = 0$, that is a theory with a vanishing two-point-function for F cannot give rise to a satisfactory formulation of QED.

Secondly, there is a severe triviality result. Namely:

[44] From the Lagrangian point of view we usually require the model to stem from a Lagrangian depending polynomially on only the fields and their first derivatives, as well as being of an additive form separating kinetic and interaction terms

[45] Note that for free fields one knows that the commutator of the fields is a c-number times the identity operator.

[46] See [Str13, Theorem 8.1, Proof of case ii)]. It is also possible to take this result as a corollary of Lemma 3.8 of this document together with the fact that $B_{\mu\nu\rho\sigma}(\xi) = d_{\mu\nu\rho\sigma}(K(\xi) - K(-\xi))$ where K is the Lorentz-invariant tempered distribution from the decomposition $W_{\mu\nu} = \eta_{\mu\nu}K + \partial_\mu\partial_\nu G$.

Theorem 2.52.
A Strocchi-Wightman QFT $(\mathcal{H}, \langle\cdot,\cdot\rangle, U, \Omega, D, \{F_{\mu\nu}\})$ *satisfying free Maxwell's equations*

$$\partial^\mu F_{\mu\nu} = 0 \quad \textit{and} \quad \partial_{[\mu} F_{\nu\rho]} = 0 \tag{2.108}$$

where there is an additional vector field A *such that* $F = dA$ *and* A *satisfies either*

a) locality, i.e., $[A_\mu(x), A_\nu(y)] = 0 \quad \textit{for } (x-y)^2 < 0$

or

b) covariance, i.e., $U(a,\Lambda) A_\mu(u) U(a,\Lambda)^{-1} = A_\mu(u_{a,\Lambda}), \quad u \in \mathcal{S}$,

gives rise to a trivial two-point function of F

$$\langle \Omega, F_{\mu\nu}(x) F_{\rho\sigma}(y) \Omega \rangle = 0. \tag{2.109}$$

Proof. See [Str13, Chapter 7.8 Appendix: Quantization of the electromagnetic potential, pp. 191] for a text-book proof. See [Str67] for the original result. ■

Remark 2.53. The result may be actually formulated in a stronger version, i.e., with weaker assumptions. The result goes through without uniqueness of the vacuum, without spectral condition, with covariance of the two-point function $\langle \Omega, A_\mu(x) A_\nu(y) \Omega \rangle$ instead of covariance of A_μ and with locality of A with respect to F instead of full locality (i.e., A_μ and $F_{\rho\sigma}$ weakly commuting on D for spacelike separated arguments).

Hence this result poses a severe problem for the construction of QED. We have either the option to dispose of locality and covariance for A or to modify the equations of motion governing A. In case of the former option, e.g. Coulomb gauge, little is known concerning methods of renormalization for non-local non-covariant fields. Thus here and in many cases within the literature the second option will be chosen.[47]

2.7.2 QED, Lorentz gauge, covariant gauges, and Gupta-Bleuler-formalism

In the former section it was stressed that if we are not willing to sacrifice the properties of locality and covariance we inevitably have to modify the equations of motion governing A. Also we have learnt that it is advisable to select a specific representative for A by applying a gauge condition. We will focus here on covariant gauges. In order to ensure the covariance of A the gauge condition must be Lorentz-covariant. Well known from the classical case is the Lorentz condition $\partial A = \partial_\mu A^\mu = 0$. Thus let us rewrite (classical) Maxwell's equation as a system of equations

$$\Box A_\mu - \lambda \partial_\mu \partial A = 0, \tag{2.110}$$
$$\partial A = 0 \tag{2.111}$$

with some up to now arbitrary real parameter λ. Note that eq. (2.110) is the most general covariant second-order differential equation which is linear in A i.e., describing a

[47]Note that in the Lagrangian setting of QFT this modification of the equations of motion corresponds to the addition of a gauge-fixing term to the Lagrangian.

free field. If we accept that higher-order differential equations lead to pathological QFTs (when considered to be a fundamental QFT)[48] then eq. (2.110) is the most general covariant equation for a free vector field.

The strategy for the construction of the quantum theory is now to focus at first on eq. (2.110), taking it to be the equation of motion governing A. For $\lambda \neq 1$ this deviates from free Maxwell's equation and does not run into the triviality result. But because of this deviation A will also necessarily generate unphysical states from the vacuum. In other words, we enlarge our space of states to a broader space of virtual states. The physical interpretation will be recovered by a condition for physical states that implies the validity of the Lorentz condition (2.111) between matrix elements of physical states. We cannot just use $\partial A = 0$ as an operator equation on the whole of the state space as then we just recover the ordinary free Maxwell equations and run into the triviality result. We can also not use the condition $\partial A \psi = 0$ for all physical states ψ (and in particular for the vacuum state) because then the two-point function of A runs then into the same triviality result[49] Eventually, the Gupta-Bleuler choice $\partial A^{(-)}\psi = 0$ for all physical states ψ works out: It implies that $\langle \psi, \partial A \phi \rangle = 0$ for physical states ψ, ϕ and it does not run into the triviality result. This condition is also called Gupta-Bleuler(GB) subsidiary condition. For a detailed and explicit construction of the GB formalism the reader is referred to [Sch61, Chapter 9b, pp. 242].

The described picture is well adapted to the Strocchi-Wightman framework for QFT that we have introduced before. The GB-subsidiary condition takes here the role of selecting the space of non-negative states. That is really necessary to allow for an indefinite metric on the space of virtual states will be seen later.

[48]We do not want to go into this discussion here. It should be noted, however, that higher than second-order derivatives lead to a non-renormalizable QFT which fits into the framework of effective quantum field theory but causes severe problems when we are up to constructing a fundamental quantum field theory. The non-renormalizable interaction terms become infinitely strong at high energies and therefore violate the unitarity of the theory.

[49]On way to see this is that the condition implies that $\partial^\mu W_{\mu\nu} = 0$ where $W_{\mu\nu}$ denotes the two-point function of A. As also $\Box W_{\mu\nu} = 0$ the triviality result will be applicable to W, as well. Another way to see this is to note that one has to require $\partial A \Omega$ then for the vacuum vector to be physical. By the Reeh-Schlieder property of the vacuum we get that $\partial A = 0$ and see that in fact we have not weakened the physicality condition.

Chapter 3

The Details of the proof

This section gives the precise reasoning of the proof based on the preliminaries that were worked out in the section above. The reasoning follows the three basic steps as outlined in Section 1.1. The section is therefore subdivided, accordingly. But to begin with, we will give the setup we are using in this section in detail. Part of the motivation of this setting was already given in the introduction and the outline of the proof (see Section 1).

Suppose an indefinite metric relativistic quantum theory $(\mathcal{H}, \langle\cdot,\cdot\rangle, U)$ (not necessarily satisfying the spectrum condition) and having some (not necessarily unique) vacuum vector $\Omega \in \mathcal{H}$. Then suppose that there is a hermitian covariant vector field (D, A) embedded in the rel. QT where $\Omega \in D$. Moreover, let us define $F_{\mu\nu} \equiv F_{\mu\nu}(A) \equiv \partial_{[\mu}A_{\nu]}$ on D. Then F is clearly a hermitian covariant 2-tensor field (D, F) (embedded in the rel. QT). We may define the (manifestly translation-invariant) two-point functions (in generalized function notation and according to Theorem 2.47) by

$$W(x-y) \equiv \langle\Omega, A(x)\otimes A(y)\Omega\rangle \quad \text{and} \quad B(x-y) \equiv \langle\Omega, F(x)\otimes F(y)\Omega\rangle\,. \tag{3.1}$$

In components one may write

$$W_{\mu\nu}(x-y) \equiv \langle\Omega, A_\mu(x)A_\nu(y)\Omega\rangle \quad \text{and} \quad B_{\mu\nu\rho\sigma}(x-y) \equiv \langle\Omega, F_{\mu\nu}(x)F_{\rho\sigma}(y)\Omega\rangle\,. \tag{3.2}$$

This should give the setup of the proof of this section which aims to be fairly general. Note that the special cases of a Wightman QFT with a hermitian covariant vector field (D, A) and the case of free QED are included in the setup. The result which we will prove in this section is

Theorem 3.1. *Suppose that A fulfils the equation of motion*

$$\Box A_\mu - \lambda\partial_\mu\partial A = 0, \quad \lambda \neq 1 \tag{3.3}$$

and that $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ is a space with positive semidefinite metric. Then the two-point function of F is trivial, i.e., $B = 0$.

This theorem will basically be an implication of the following result

Lemma 3.2. *Let W be an arbitrary (Lorentz-)tensor-valued tempered distribution on $\mathbb{M}$ which is subject to the conditions* *(a) W transforms covariantly under the representation $\mathcal{D}^{(1/2,1/2)} \otimes \mathcal{D}^{(1/2,1/2)}$ of $SL(2,\mathbb{C})$*

(b) W satisfies the equation of motion

$$\Box W_{\mu\nu} + \lambda \partial_\mu \partial^\rho W_{\rho\nu} = 0, \quad \lambda \neq 1. \tag{3.4}$$

Then supposing that the components $\hat{W}_{\mu\mu}$ are non-negative, namely

$$\hat{W}_{\mu\mu}(u) \geq 0 \quad \forall \text{ non-negative } u \in \mathcal{S}. \tag{3.5}$$

implies

$$W_{\mu\nu}(x) = c\partial_\mu \partial_\nu G \tag{3.6}$$

for some constant $c \geq 0$ and a non-negative Lorentz-invariant tempered distribution G.

Let us proof that under the assumption of Lemma 3.2 also Theorem 3.1 holds:

Proof. At first we have to show that the definition $W = \langle \Omega, A(x) \otimes A(y)\Omega \rangle$ in the given setup fulfils all the properties (a) and (b) and the non-negativity condition required for W:
(a): By the invariance of Ω and the covariance of A (and also the invariance of D under actions of U and A) we have that

$$\begin{aligned} W(x-y) &= \langle \Omega, U(a, \underset{\sim}{\Lambda}) A(x) U(a, \underset{\sim}{\Lambda})^{-1} \otimes U(a, \underset{\sim}{\Lambda}) A(y) U(a, \underset{\sim}{\Lambda})^{-1} \Omega \rangle \\ &\equiv \mathcal{D}^{(1/2,1/2)}(\underset{\sim}{\Lambda}) \otimes \mathcal{D}^{(1/2,1/2)}(\underset{\sim}{\Lambda}) \cdot W(x-y). \end{aligned}$$

(b): In generalized function notation the proof is quite immediate, but there is the risk of oversimplified notation. Thus let us remember that the generalized function $W(x-y)$ is defined as a tempered distribution on $\mathbb{M}$ extending the definition $W(u * \tilde{v}) \equiv \langle \Omega, A(u)A(v)\Omega \rangle$ for $u, v \in \mathcal{S}$ and where $\tilde{v}(x) = v(-x)$. Then for some differential D^α for some multi-index α we have

$$D^\alpha W(u * \tilde{v}) = W(D^\alpha(u * \tilde{v})) = W((D^\alpha u) * \tilde{v}) = (-1)^{|\alpha|} W(u * (D^\alpha \tilde{v})). \tag{3.7}$$

The second two expressions correspond to derivatives taken with respect to x and y in generalized function notation. As in the equations of motion there are only second-order differentials, the minus-sign in the last expression is absent and there is no difference if the differentials are taken with respect to x, y or $x-y$.
Non-negativity: From the hypothesis of the theorem, namely that $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a space with non-negative definite metric, and the hermiticity of A it follows that for each $u \in \mathcal{S}$

$$\begin{aligned} \hat{W}_{\mu\mu}(|\hat{u}|^2) &= W_{\mu\mu}(\tilde{\bar{u}} * u) \\ &= \langle \Omega | A_\mu(\bar{u}) A_\mu(u) \Omega \rangle \\ &= ||A_\mu(u)\Omega||^2 \\ &\geq 0. \end{aligned}$$

The first equality is given by Parseval's theorem. It is not difficult to show that $W_{\mu\mu}$ being of positive type (what is written above) implies the non-negativity of $\hat{W}_{\mu\mu}$, i.e., $\forall$ non-negative $u \in \mathcal{S} : \hat{W}_{\mu\mu}(u) \geq 0$.

From Lemma 3.2 we then obtain that $W_{\mu\nu} = \partial_\mu \partial_\nu K$ for some Lorentz-invariant tempered distribution K.
For the two-point function of F we therefore obtain

$$\begin{aligned} B_{\mu\nu\rho\sigma} &= -\partial_\mu \partial_\rho W_{\nu\sigma} + \partial_\mu \partial_\sigma W_{\nu\rho} - \partial_\nu \partial_\sigma W_{\mu\rho} + \partial_\nu \partial_\rho W_{\mu\sigma} \\ &= 0 \end{aligned} \tag{3.8}$$

by taking the differential operators acting on A out of the expectation value. In the second line we used that $W_{\mu\nu} = \partial_\mu \partial_\nu K$ and that differential operators acting on tempered distributions commute.

■

Remark 3.3. Another way to express the triviality of the resulting theory is to note that $B_{\mu\nu\rho\sigma} = 0$ implies that $||F_{\mu\nu}(u)\Omega||^2 = B_{\mu\nu\mu\nu}(\tilde{\bar{u}} * u) = 0$ for all $u \in \mathcal{S}$. On the induced physical Hilbert space we will therefore obtain $F^{\mathrm{ph}}_{\mu\nu}(\cdot)\Omega^{\mathrm{ph}} = 0$. Here F^{ph} and Ω^{ph} are the objects on the induced physical Hilbert space associated to F and Ω. The Reeh-Schlieder property of the vacuum then implies $F^{\mathrm{ph}} = 0$ and the vacuum becomes the only physical state of the theory.[1] The equation $F_{\mu\nu} = \partial_{[\mu} A_{\nu]} = 0$ is equivalent to A being gradientlike/trivial, i.e., $A_\mu = \partial_\mu \phi$ for some hermitian scalar field ϕ.[2]

With this the proof of the theorem reduces to the proof of the lemma. The proof of the lemma will be a structure analysis of the tensor-valued tempered distribution W which follows the three steps that were outlined in the introduction.

3.1 Step 1 - Covariant structure of W

The first step will focus on Condition (a), thus suppose for this part that W is an arbitrary Lorentz-covariant distribution transforming under $\mathcal{D}^{(1/2,1/2)} \otimes \mathcal{D}^{(1/2,1/2)}$. Note that we do not suppose Condition (b) and non-negativity. The important facts from the preliminaries are

> *Fact 1:* For each non-negative half-integral numbers s, t, s', t' the representation $(s,t) \otimes (s',t')$ of $SL(2,\mathbb{C})$ decomposes into $\bigoplus_{|s-s'| \leq j \leq s+s'} \bigoplus_{|t-t'| \leq k \leq t+t'} (j,k)$, where j, k run in integral steps. (see Proposition 2.17)
>
> *Fact 2:* An arbitrary Lorentz-covariant generalized function F on $\mathbb{M}$ transforming according to (j,k) is non-zero only for $j = k = n/2$ and in this case may be represented as a tempered distribution $F(p;\omega,\bar{\omega}) = (\bar{\omega}\tilde{p}\omega)^n f(p)$, where $f(p)$ is a Lorentz-invariant generalized function defined within n arbitrary constants and $\tilde{p} = p^\mu(\sigma^\mu)_\mu$ is a hermitian 2x2 matrix. More precisely, if $f_0(p)$ is a fixed solution then the general solution has the form $f(p) = f_0(p) + \sum_{l=0}^{n-1} a_l \Box^l \delta(p)$ for some constants a_l, $l = 0, ..., n-1$. (see Proposition 2.25)

Combining Condition (a) and the two facts from the preliminaries we obtain

[1] Note that here a state is understood as a ray of vectors
[2] See also Corollary 1 of [Str67].

Lemma 3.4. *Suppose W is an arbitrary tensor-valued tempered distribution in $\mathbb{M}$ satisfying Condition (a). Then W may be represented as either*

$$W(p;\omega,\bar{\omega}) = k(p) + g(p)(\bar{\omega}\tilde{p}\omega)^2 \tag{3.9}$$

or

$$W_{\mu\nu}(p) = \eta_{\mu\nu}\hat{K}(p) - p_\mu p_\nu \hat{G}(p), \tag{3.10}$$

where k, g and $K = \frac{1}{4}(k - p^2 g)$, $G = -g$ are Lorentz-invariant generalized functions. The equation (3.9) solving for the k and g uniquely determines k and determines g up to addition of a term $(a\Box + b)\delta(p)$ for some constants a, b. Accordingly, $\hat{K}$ and $\hat{G}$ are determined up to shifts $(\hat{K}, \hat{G}) \mapsto (\hat{K} + 2a\delta, \hat{G} + (a\Box + b)\delta)$.

Remark 3.5. The minus sign in (3.10) and the ˆ-symbols for K and G are just conventional. They are there to indicate the relation between momentum and coordinate space that we will later use.

Proof. According to Condition (a) the two-point-function transforms as $(1/2, 1/2) \otimes (1/2, 1/2)$, which according to Fact 1 decomposes into $(0,0) \oplus (1,0) \oplus (0,1) \oplus (1,1)$. When restricting to one of the summands, this defines an irreducible representation and Fact 2 is applicable. From this we infer that the $(1,0) \oplus (0,1)$-valued sector of the two-point function vanishes and thus obtain eq. (3.9). Fact 2 also implies that k is uniquely defined and g only up to addition of a term $(a\Box + b)\delta(p)$ for some constants a and b.

The second representation of the two-point function is obtained by a change of value space from a space of homogeneous polynomials in $\omega, \bar{\omega}$[3] to the more common space of symmetric Lorentz-two-tensors. Thus let us introduce the polynomials $e_\mu(\omega, \bar{\omega}) \equiv \bar{\omega}\sigma_\mu\omega$, where ω and $\bar{\omega}$ are understood as column- and row-vectors, respectively. By explicit calculation one checks that $e^2(\omega, \bar{\omega}) = \eta^{\mu\nu} e_\mu(\omega, \bar{\omega}) e_\nu(\omega, \bar{\omega}) = ... = 0$.[4]

With this we can write

$$\begin{aligned}
W(p;\omega,\bar{\omega}) &= k(p) + g(p)(\bar{\omega}\tilde{p}\omega)^2 \\
&= k(p)\eta_{\mu\nu}\frac{1}{4}\eta^{\mu\nu} + g(p)p_\mu p_\nu e^\mu(\omega,\bar{\omega})e^\nu(\omega,\bar{\omega}) \\
&= k(p)\eta_{\mu\nu}\left(\frac{1}{4}\eta^{\mu\nu} + e^\mu(\omega,\bar{\omega})e^\nu(\omega,\bar{\omega})\right) + g(p)p_\mu p_\nu e^\mu(\omega,\bar{\omega})e^\nu(\omega,\bar{\omega}) \\
&= \Big[\eta_{\mu\nu}(k(p) - g(p)p^2) + p_\mu p_\nu g(p)\Big]\left[\frac{1}{4}\eta^{\mu\nu} + e^\mu(\omega,\bar{\omega})e^\nu(\omega,\bar{\omega})\right] \\
&= W_{\mu\nu}(p)e^{\mu\nu}(\omega,\bar{\omega}).
\end{aligned}$$

In the last step we have defined $e^{\mu\nu}(\omega, \bar{\omega}) \equiv 1/4\ \eta^{\mu\nu} + e^\mu(\omega, \bar{\omega})e^\nu(\omega, \bar{\omega})$. These polynomials form a basis of $\mathcal{P}^{(0,0)} \oplus \mathcal{P}^{(1,1)}$. Whereas $W(p;\omega,\bar{\omega})$ was characterized by the invariance

[3] The precise value space is $\mathcal{P}^{(0,0)} \oplus \mathcal{P}^{(1,1)} \cong \mathbb{C} + \mathcal{P}^{(1,1)}$. But this space may be included in some other space of homogeneous polynomials in $\omega, \bar{\omega}$.

[4] The explicit calculation goes as follows: $e_0(\omega,\bar{\omega}) = \bar{\omega}_1\omega_1 + \bar{\omega}_2\omega_2$, $e_1(\omega,\bar{\omega}) = \bar{\omega}_1\omega_2 + \bar{\omega}_2\omega_1$, $e_2(\omega,\bar{\omega}) = -i(\bar{\omega}_1\omega_2 - \bar{\omega}_2\omega_1)$, $e_3(\omega,\bar{\omega}) = \bar{\omega}_1\omega_1 - \bar{\omega}_2\omega_2$. Thus $(e_0)^2 - (e_3)^2(\omega,\bar{\omega}) = 4\bar{\omega}_1\bar{\omega}_2\omega_1\omega_2 = (e_1)^2 + (e_2)^2(\omega,\bar{\omega})$. Hence $e^2 = \eta^{\mu\nu}e_\mu e_\nu = 0$.

condition $W(\Lambda(\underset{\sim}{\Lambda})p;\underset{\sim}{\Lambda}\omega,\bar{\underset{\sim}{\Lambda}}\bar{\omega}) = W(p;\omega,\bar{\omega})$, we can infer from this that $W_{\mu\nu}(p)$ transforms as an ordinary two-tensor:

$$\begin{aligned}
W(\Lambda p;\underset{\sim}{\Lambda}\omega,\bar{\underset{\sim}{\Lambda}}\bar{\omega}) &= W_{\mu\nu}(\Lambda p)e^{\mu\nu}(\underset{\sim}{\Lambda}\omega,\bar{\underset{\sim}{\Lambda}}\bar{\omega}) \\
&= W_{\mu\nu}(\Lambda p)\left[1/4\ \eta^{\mu\nu} + (\Lambda^T e\Lambda)^{\mu\nu}(\omega,\bar{\omega})\right] \\
&= W_{\mu\nu}(\Lambda p)\left[1/4\ (\Lambda^T\eta\Lambda)^{\mu\nu} + (\Lambda^T e\Lambda)^{\mu\nu}(\omega,\bar{\omega})\right] \\
&= {\Lambda^{-1}}_{\mu}{}^{\rho}{\Lambda^{-1}}_{\nu}{}^{\sigma}W_{\rho\sigma}(\Lambda p)\,e^{\mu\nu}(\omega,\bar{\omega}) \\
&\stackrel{!}{=} W_{\mu\nu}(p)\,e^{\mu\nu}(\omega,\bar{\omega}),
\end{aligned}$$

where we kept implicit that $\Lambda = \Lambda(\underset{\sim}{\Lambda})$ and we used the invariance of the metric $\Lambda^T\eta\Lambda = \eta$ and the transformation behaviour $e_\mu(\underset{\sim}{\Lambda}\omega,\bar{\underset{\sim}{\Lambda}}\bar{\omega}) = (\Lambda e)_\mu(\omega,\bar{\omega})$.[5]
Concerning the ambiguity of the decomposition one can simply observe that $(k,g)\mapsto(k,g-(a\Box+b)\delta)$ becomes $(\hat{K},\hat{G})\mapsto(\hat{K}+\frac{a}{4}p^2\Box\delta,\hat{G}+(a\Box+b)\delta)$ from the definitions $\hat{K}=\frac{1}{4}(k-p^2g), \hat{G}=-g$. As $p^2\Box\delta = \Box(p^2)\delta = 8\delta$ the stated ambiguity relation holds. ∎

Corollary 3.6. *Suppose W is an arbitrary tensor-valued tempered distribution in $\mathbb{M}$ satisfying Condition (a). Then W may be represented as*

$$W_{\mu\nu}(x) = \eta_{\mu\nu}K(x) + \partial_\mu\partial_\nu G(x) \tag{3.11}$$

for Lorentz-invariant tempered distributions K,G that for fixed $W_{\mu\nu}$ are defined up to shifts

$$(K,G)\mapsto(K+2a,G-ax^2+b) \tag{3.12}$$

for arbitrary constants a,b.

Proof. Let $\hat{W}(p)$ denote the Fourier transform of $W(\xi)$. Then $\hat{W}$ transforms covariantly under the same representation as W does (see Proposition 2.11). Applying the Lemma 3.4 to $\hat{W}$ yields

$$\hat{W}_{\mu\nu}(p) = \eta_{\mu\nu}\hat{K}(p) - p_\mu p_\nu\hat{G}(p). \tag{3.13}$$

An inverse Fourier transform then gives the desired result. Note that the ambiguity specified in (3.12) is the inverse Fourier transformed ambiguity specified in the referenced lemma from above. ∎

An alternative proof of the covariant decompositions presented here is to be found in Appendix E. The proof also contains an explicit check for the ambiguity of K and G.

Note that for later use we used the variables x and p to indicate which result we will use in coordinate and which result we will use in momentum space. Generally, both results, Lemma 3.4 and Corollary 3.6, can, however, be applied to both, coordinate and momentum space.

[5] The transformation behaviour of $e_\mu(\omega,\bar{\omega})$ may be checked explicitly: We know that $\bar{\omega}\tilde{p}\omega = p_\mu\bar{\omega}\sigma^\mu\omega = p_\mu e^\mu(\omega,\bar{\omega})$. Then $p_\mu e^\mu(\underset{\sim}{\Lambda}\omega,\bar{\underset{\sim}{\Lambda}}\bar{\omega}) = \bar{\omega}\underset{\sim}{\Lambda}^*\tilde{p}\underset{\sim}{\Lambda}\omega = \bar{\omega}\Lambda^{\tilde{-1}}p\omega = (\Lambda^{-1}p)_\mu e^\mu(\omega,\bar{\omega})$. In the second last step we used the definition of the covering homomorphism. Namely, that $\Lambda^{\tilde{-1}}p \equiv \underset{\sim}{\Lambda}^*\tilde{p}\underset{\sim}{\Lambda}$ (see eq. (2.53) and adapt it to the inverse $\tilde{}$ instead of $\underset{\sim}{}$). Here $\underset{\sim}{\Lambda}^* \equiv \bar{\underset{\sim}{\Lambda}}^T$ is the adjoint of $\underset{\sim}{\Lambda}$.

3.2 Step 2 - Differential structure of W

Suppose now that W is satisfying not only Condition (a), but also Condition (b). Condition (b) states that the equations of motion

$$\Box W_{\mu\nu} - \lambda \partial_\mu \partial^\rho W_{\rho\nu} = 0, \quad \lambda \neq 1 \tag{3.14}$$

are satisfied.

For simplicity in the following we will furthermore suppose that the support of $\hat{W}$ is contained in the closed upper light cone. We refer to this as the spectral condition for W. The condition makes it easier to relate the result to the special case of the Wightman setting where the spectrum condition clearly holds. In Subsection 3.4 we will explain why the lemma also holds when the spectral condition is not assumed.

The important fact from the preliminaries is

> *Fact 3:* Let F be an arbitrary Lorentz-invariant tempered distribution on $\mathbb{M}$ satisfying $p^2 F(p) = T(p)$ for some tempered distribution T. Then $F = a\hat{D}^{(+)} + b\hat{D}^{(-)} + c\delta + F_{\text{part}}$ for some constants a, b, and c and a particular solution F_{part}. (see Theorem 2.29)

Note that when Fact 3 applies to F, then the spectrum condition, supp $F \subset \bar{V}^+$, is equivalent to $b = 0$. In this regard it is important to extend the support properties of $\hat{W}$ to $\hat{K}$ and $\hat{G}$:

Lemma 3.7. *Suppose $W_{\mu\nu} = \eta_{\mu\nu} K + \partial_\mu \partial_\nu G$ for some Lorentz-invariant tempered distributions K and G and let W satisfy the spectral condition. Then supp $\hat{K} \subset \bar{V}^+$ and supp $\hat{G} \subset \bar{V}^+$.*

Proof. By the spectral condition we have

$$\text{supp } \hat{W}_{\mu\nu} \subset \bar{V}^+. \tag{3.15}$$

From Lemma 3.4 we know that $\hat{K}$ and $\hat{G}$ are fixed by $\hat{W}_{\mu\nu}$ up to distributions supported at the origin. Thus on the complement of $\bar{V}^+$ the decomposition is unique and the restriction of $\hat{W}_{\mu\nu}$ to the complement yields

$$0 = \hat{W}_{\mu\nu}\restriction_{\mathbb{M}\setminus\bar{V}^+} = \eta_{\mu\nu}\hat{K}\restriction_{\mathbb{M}\setminus\bar{V}^+} - p_\mu p_\nu \hat{G}\restriction_{\mathbb{M}\setminus\bar{V}^+}. \tag{3.16}$$

Clearly, $\hat{K}\restriction_{\mathbb{M}\setminus\bar{V}^+} = \hat{G}\restriction_{\mathbb{M}\setminus\bar{V}^+} = 0$ is a solution and therefore unique, as well. This implies that the unrestricted $\hat{K}$ and $\hat{G}$ are supported within $\bar{V}^+$. ∎

Then the result of this section is

Lemma 3.8. *Suppose W is an arbitrary tensor-valued tempered distribution in $\mathbb{M}$ satisfying Condition (a)-(c). Then W is of the form*

$$\begin{aligned} W_{\mu\nu} = {} & c_1 \left(\eta_{\mu\nu} D^{(+)} + \frac{\lambda}{1-\lambda} \partial_\mu \partial_\nu x^2 D^{(+)} \right) - c_2 \partial_\mu \partial_\nu D^{(+)} \\ & + c_3 \left(-\eta_{\mu\nu} x^2 + \frac{4-\lambda}{24(1-\lambda)} \partial_\mu \partial_\nu (x^2)^2 \right) + c_4 \eta_{\mu\nu} \end{aligned} \tag{3.17}$$

or

$$\hat{W}_{\mu\nu} = c_1\left(\eta_{\mu\nu}\hat{D}^{(+)} - \frac{\lambda}{1-\lambda}p_\mu p_\nu \hat{D}'^{(+)}\right) + c_2 p_\mu p_\nu \hat{D}^{(+)}$$
$$+c_3\left(\eta_{\mu\nu}\Box\delta - \frac{4-\lambda}{24(1-\lambda)}p_\mu p_\nu \Box^2\delta\right) + c_4\eta_{\mu\nu}\delta \tag{3.18}$$

for some constants $c_1, ..., c_4$.

Remark 3.9. The summands in the second lines of eqs. (3.17) and (3.18) are all monomials in x or is Fourier transforms, terms concentrated at $p = 0$. These terms usually do not appear in physics textbooks. The reason is that the c_3-term can be shown to vanish when existence and uniqueness of the vacuum is supposed and that the c_4-term can be absorbed into a redefinition of the field A_μ such that the new A_μ satisfies $\langle \Omega, A_\mu(\cdot)\Omega\rangle = 0$. This will be explained in more detail in the discussion of the results, Subsection 3.5. We will, however, continue to work with these terms such that we can proceed without further assumptions. In the next step these terms will be shown to drop out anyway.

Proof. The proof is a somewhat lengthy computation that can be done in either coordinate or momentum space. The momentum space computation seems to be a bit easier. The equations of motion in momentum space read

$$p^2\hat{W}_{\mu\nu} - \lambda p_\mu p^\rho \hat{W}_{\rho\nu} = 0. \tag{3.19}$$

We will use the covariant decomposition from Step 1, i.e., $\hat{W}_{\mu\nu} = \eta_{\mu\nu}\hat{K} - p_\mu p_\nu \hat{G}$ for some Lorentz-invariant tempered distributions $\hat{K}$ and $\hat{G}$. Note that the support of $\hat{K}$ and $\hat{G}$ is contained in the closed upper light cone by Condition (c) and Lemma 3.7 from above. Applying the equations of motion to the covariant decomposition yields

$$(\eta_{\mu\nu}p^2 - \lambda p_\mu p_\nu)\hat{K} - (1-\lambda)p_\mu p_\nu p^2\hat{G} = 0. \tag{E1}$$

Contracting (E1) with p^μ and dividing by $(1-\lambda)$ (note that we have assumed $\lambda \neq 1$) one obtains

$$p_\nu p^2\hat{K} - p_\nu(p^2)^2\hat{G} = 0. \tag{E2}$$

Multiplying (E1) with p^2 and inserting (E2) one obtains

$$\eta_{\mu\nu}(p^2)^2\hat{K} - p_\mu p_\nu (p^2)^2\hat{G} = 0. \tag{E3}$$

We obtain then two new equations from (E1) by contracting with $p^\mu p^\nu$ and from (E3) by taking the trace, i.e., contracting with the metric. Combining them gives us an intermediate result:

$$\begin{aligned}
&\tfrac{1}{1-\lambda}p^\mu p^\nu (\mathrm{E1})_{\mu\nu} && (p^2)^2\hat{K} - (p^2)^3\hat{G} = 0 && \text{(E4)}\\
&\eta^{\mu\nu}(\mathrm{E3})_{\mu\nu} && 4(p^2)^2\hat{K} - (p^2)^3\hat{G} = 0 && \mathrm{Tr(E3)}\\
&(\mathrm{E4}) \wedge \mathrm{Tr(E3)} && \Rightarrow (p^2)^2\hat{K} = (p^2)^3\hat{G} = 0. && \text{(E5)}
\end{aligned}$$

Inserting (E5) into (E3) yields

$$p_\mu p_\nu (p^2)^2 \hat{G} = 0. \tag{E3*}$$

As $\hat{G}$ and thus $(p^2)^2\hat{G}$ are Lorentz-invariant (E3*) is equivalent to

$$(p^2)^2 \hat{G} = a\delta \tag{E6}$$

for some constant a. Inserting (E6) into the trace of (E1) then yields

$$p^2 \hat{K} = a\frac{1-\lambda}{4-\lambda}\delta \tag{E7}$$

in no conflict to (E5). Applying Fact 3 (and the support properties of $\hat{K}$) to (E7) yields

$$\hat{K} = a\frac{1-\lambda}{4-\lambda}\frac{1}{8}\Box\delta + b\hat{D}^{(+)} + c\delta \tag{E8}$$

for some constants b and c and where the first summand is a particular solution to (E7) as $p^2\frac{1}{8}\Box\delta(p) = \frac{1}{8}\Box(p^2)\delta(p) = \delta(p)$. As an auxillary expression let us note that

$$p_\mu p_\nu \hat{K} = -a\eta_{\mu\nu}\frac{1-\lambda}{4-\lambda}\frac{1}{4}\delta + bp_\mu p_\nu \hat{D}^{(+)} \tag{E9}$$

where we used that $p_\mu p_\nu \Box\delta(p) = \Box(p_\mu p_\nu)\delta(p) = 2g_{\mu\nu}\delta(p)$.
Similarly, we can apply Fact 3 (and the support properties of $\hat{G}$) to (E6) and obtain

$$p^2\hat{G} = a\frac{1}{8}\Box\delta + d\hat{D}^{(+)} + e\delta \tag{E10}$$

for some constants d and e. Applying Fact 3 again, we obtain

$$\hat{G} = a\frac{1}{196}\Box^2\delta - d\hat{G}_{\text{part}} + e\frac{1}{8}\Box\delta + f\hat{D}^{(+)} + g\delta \tag{E11}$$

for some constants f and g and where $\hat{G}_{\text{part}} \in \mathcal{S}'$ is a particular solution to $p^2\hat{G} = -\hat{D}^{(+)}$. We will need some space to discuss this division problem: Note at first that what we are interested in is what $p_\mu p_\nu G$ looks like and not G itself. Then note that $\hat{D}'^{(+)}(p) \equiv \theta(p_0)\delta'(p^2), \quad p \neq 0$ defines a generalized function on $\mathbb{M}_\times$ which solves the noted division problem (on $\mathbb{M}_\times$) as $p^2\theta(p_0)\delta'(p^2) = \theta(p_0)p^2\delta'(p^2) = -\theta(p_0)\delta(p^2) = -\hat{D}^{(+)}(p)$ for $p \neq 0$. The problem is now that the expression $\theta(p_0)\delta'(p^2)$ is not a well-defined generalized function on the whole of $\mathbb{M}$. The good thing is that $p_\mu p_\nu\theta(p_0)\delta'(p^2)$ is well-defined on the whole of $\mathbb{M}$ as will be shown in a moment.[6] Thus we demand that $\hat{G}_{\text{part}}$ is a tempered distribution solving $p^2\hat{G}_{\text{part}} = -\hat{D}^{(+)}$ on the whole of $\mathbb{M}$ such that $p_\mu p_\nu\hat{G}_{\text{part}} = p_\mu p_\nu\hat{D}'^{(+)}$. That $p_\mu p_\nu\hat{D}'^{(+)}$ is well-defined can be seen by using distributional notation:

$$p_\mu p_\nu \hat{D}'^{(+)}(u) \equiv -\int \left[\partial_0 \frac{p_\mu p_\nu u(p)}{2p_0}\right]_{p_0=|\vec{p}|} \frac{d\vec{p}}{2|\vec{p}|} \tag{3.20}$$

which is a well-defined (Riemann) integral for each $u \in \mathcal{S}$. We are now only missing the existence $\hat{G}_{\text{part}}$ as a particular solution to the division problem subject to $p_\mu p_\nu\hat{G}_{\text{part}} = p_\mu p_\nu\hat{D}'^{(+)}$.

[6] Technically speaking, we should say that there exists an extension of the expression $p_\mu p_\nu\theta(p_0)\delta'(p^2), p \neq 0$, which defines a well-defined tempered distribution on the whole of $\mathbb{M}$.

For this note that the Lorentz-invariant extensions of $\hat{D}'^{(+)}$ to the whole of $\mathbb{M}$ fulfil these conditions. They are furthermore determined up to the addition of a δ-term which vanishes when $p_\mu p_\nu$ is applied to it (see e.g. the proof of Lemma 2.27). As an auxillary expression let us note that

$$p_\mu p_\nu p^2 \hat{G} = a\eta_{\mu\nu}\frac{1}{4}\delta + dp_\mu p_\nu \hat{D}^{(+)}. \tag{E12}$$

Lastly, we have to apply (E1) to $\hat{K}$ and $\hat{G}$. Inserting (E7), (E9), and (E12) into (E1) yields

$$\begin{aligned}
0 &= a\eta_{\mu\nu}\frac{1-\lambda}{4-\lambda}\delta - \lambda\left(a\eta_{\mu\nu}\frac{1-\lambda}{4-\lambda}\frac{1}{4}\delta + bp_\mu p_\nu \hat{D}^{(+)}\right) - a\eta_{\mu\nu}(1-\lambda)\frac{1}{4}\delta - d(1-\lambda)p_\mu p_\nu \hat{D}^{(+)} \\
&= a\eta_{\mu\nu}\left((1-\frac{\lambda}{4})\frac{1-\lambda}{4-\lambda} - \frac{1}{4}(1-\lambda)\right)\delta - (d(1-\lambda)+b\lambda)p_\mu p_\nu \hat{D}^{(+)} \\
&= a\eta_{\mu\nu}\left(\frac{4-\lambda}{4}\frac{1-\lambda}{4-\lambda} - \frac{1}{4}(1-\lambda)\right)\delta - (d(1-\lambda)+b\lambda)p_\mu p_\nu \hat{D}^{(+)} \\
&= -(d(1-\lambda)+b\lambda)p_\mu p_\nu \hat{D}^{(+)}.
\end{aligned}$$

This implies that $d = -b\frac{\lambda}{1-\lambda}$. Putting everything together we obtain

$$\hat{W}_{\mu\nu} = \eta_{\mu\nu}\left(a\frac{1-\lambda}{4-\lambda}\frac{1}{8}\Box\delta + b\hat{D}^{(+)} + \left(c-\frac{e}{2}\right)\delta\right) - p_\mu p_\nu\left(a\frac{1}{196}\Box^2\delta + b\frac{\lambda}{1-\lambda}\hat{D}'^{(+)} + f\hat{D}^{(+)}\right). \tag{3.21}$$

where we have used that $p_\mu p_\nu \Box\delta(p) = 2g_{\mu\nu}\delta(p)$ to put the c- and e-term together. Note that the g-term does not contribute as $p_\mu p_\nu \delta = 0$. A redefinition of the constants and an inverse Fourier transform yields the desired expressions. ■

3.3 Step 3 - The final argument

With the completion of Step 2 we have now obtained a precise expression for W which is determined up to four constants. Let us continue with assuming Conditions (a) and (b) as well as the spectral condition. Moreover, let us add the additional assumption of non-negativity for W. As it will turn out this will imply that all but one of the summands will drop out (i.e., the proportionality constants have to be zero for these terms). Let us restate the condition of non-negativity for W. W is called non-negative if

$$\hat{W}_{\mu\mu}(u) \geq 0 \quad \forall \text{ non-negative } u \in \mathcal{S}. \tag{3.22}$$

The idea for the proof is that many of the summands are indefinite on its own and that one can scale their contributions to W to be dominant. A non-negativity assumption will then imply that the corresponding term/constant has to vanish. An important observation for the argument is that all the summands of $\hat{W}_{\mu\nu}$ are homogeneous. This will simplify the proof as we do not have to look for specific test function families that realize the scaling idea.

Definition 3.10. A generalized function f on $\mathbb{R}^n$ is called λ-homogeneous if, and only if, for some $\lambda \in \mathbb{C}$

$$f(\rho x) = \rho^\lambda f(x) \; \forall \rho > 0. \tag{3.23}$$

There are some obvious relations:

Proposition 3.11. *Let f be an arbitrary λ-homogeneous generalized function on $\mathbb{R}^n$. Then:*

(a) The degree of homogeneity of f is uniquely defined if, and only if, $f \neq 0$ (i.e., $f(u) \neq 0 \; \forall u$).

(b) A linear combination of λ-homogeneous gf.'s is λ-homogeneous.

(c) A product of λ_i-homogeneous gf.'s, whenever it is defined, is $\prod_i \lambda_i$-homogeneous.

(d) The delta distribution is $-n$-homogeneous and a constant distribution is 0-homogeneous.

(e) The generalized function $x^\alpha f$ for an arb. multi-index alpha is $\lambda + |\alpha|$-homogeneous.

(f) The generalized function $\partial_\alpha f$ is $\lambda - |\alpha|$-homogeneous.

(g) The Fourier transform $\hat{f}$ of f is $-\lambda - n$-homogeneous.

(h) The generalized function $D^{(\pm)}$ and its Fourier transform $\hat{D}^{(+)} = \theta(p_0)\delta(p^2)$ on $\mathbb{M}$ are (-2)-homogeneous.

Proof. All the proofs are straight forward. Property (g) is also implied by the transformation formula of the Fourier transform for real linear transformations. (See Proposition 2.11). We will only prove (h) here:

$$\begin{aligned}
\hat{D}^{(\pm)} \circ \rho\ (u) &= \rho^{-4} \hat{D}^{(\pm)}(u \circ \rho^{-1}) \\
&= \rho^{-4}(-2\pi i) \int \frac{u \circ \rho^{-1}(|\vec{p}|, \vec{p})}{2|\vec{p}|} d\vec{p} \\
&= \rho^{-4}(-2\pi i) \int \frac{u \circ \rho^{-1}(|\vec{p}|, \vec{p})}{2\rho\rho^{-1}|\vec{p}|} \rho^3 d(\rho^{-1}\vec{p}) \\
&= \rho^{-2} \hat{D}^{(+)}(u).
\end{aligned}$$

For its inverse Fourier transform we obtain the same degree of homogeneity:
$-4 - (-2) = -2$ by the formula of Property (g). ■

As the summands of $\hat{W}_{\mu\nu}$ show different degrees of homogeneity, this can be used to infer that the summands with highest and lowest degree of homogeneity have to be of positive type, too. In order to prove this we start with a basic lemma about scaling

Lemma 3.12. *Let $z_1, ..., z_n \in \mathbb{C}$ and let*

$$f(\rho) \equiv \rho^{\lambda_1} z_1 + ... + \rho^{\lambda_n} z_n \geq 0 \tag{3.24}$$

hold for arbitrary $\rho > 0$ and fixed integer $\lambda_1 < ... < \lambda_n$. Then $z_1, ..., z_n$ are real and $z_1 \geq 0$ as well as $z_n \geq 0$.

Proof. The function $f(\rho)$ as a polynomial in ρ is clearly smooth with real derivatives. Taking the λ_i-th derivative with respect to ρ and $\rho \to 0$ converges to z_i. Therefore all the $z_1, ..., z_n$ have to be real.
Furthermore the expressions

$$\rho^{-\lambda_1} f(\rho) \geq 0 \text{ and } \rho^{-\lambda_n} f(\rho) \geq 0 \tag{3.25}$$

converge for $\rho \to 0$ against z_1 and for $\rho \to \infty$ against z_n, respectively. Therefore $z_1 \geq 0$ and $z_n \geq 0$. ■

Corollary 3.13. *Let $f_1, ..., f_n$ be homogeneous generalized functions on $\mathbb{R}^m$ with integer degrees of homogeneity $\lambda_1 < ... < \lambda_n$ such that*

$$\sum_{i=1}^{n} f_i \textit{ non-negative.} \tag{3.26}$$

Then $f_1, ..., f_n$ are real and f_1 and f_n are non-negative.

Proof. For each non-negative $u \in \mathcal{S}_m$ by homogeneity we have

$$\rho^{\lambda_1} f_1(u) + ... + \rho^{\lambda_n} f_n(u) \geq 0. \tag{3.27}$$

Lemma 3.12 implies that $f_1(u), ..., f_n(u)$ are real and that $f_1(u) \geq 0$ and $f_n(u) \geq 0$. ■

We are now ready to look at $\hat{W}$:

Lemma 3.14. *Let W be an arbitrary (Lorentz-)tensor-valued tempered distribution on $\mathbb{M}$ which is subject to the conditions (a) and (b) as well as the spectral condition. Then supposing that the components $\hat{W}_{\mu\mu}$ are non-negative, namely that*

$$\hat{W}_{\mu\mu}(u) \geq 0 \quad \forall \textit{ non-negative } u \in \mathcal{S}. \tag{3.28}$$

implies

$$W_{\mu\nu}(x) = c\partial_\mu \partial_\nu D^{(+)} \tag{3.29}$$

for some constant c.

Proof. By Lemma 3.8 and with degrees of homogeneity written above the summands we have

$$\begin{aligned}
&\text{degr. of hom.} && \qquad -2 \qquad\qquad\qquad -2 \qquad\qquad\qquad 0 \\
&\hat{W}_{\mu\nu} = c_1 \left(\eta_{\mu\nu} \hat{D}^{(+)} - \frac{\lambda}{1-\lambda} p_\mu p_\nu \hat{D}'^{(+)} \right) + c_2 p_\mu p_\nu \hat{D}^{(+)} \\
&\text{degr. of hom.} && \qquad\qquad -6 \qquad\qquad\qquad -6 \qquad\qquad -4 \\
&\qquad + c_3 \left(\eta_{\mu\nu} \Box \delta - \frac{1}{24} \frac{4-\lambda}{1-\lambda} p_\mu p_\nu \Box^2 \delta \right) + c_4 \eta_{\mu\nu} \delta.
\end{aligned} \tag{3.30}$$

Note that $\hat{W}_{\mu\nu}$ is a sum of constants times real generalized functions. By Corollary 3.13 and the non-negativity of $\hat{W}_{\mu\mu}$ we see that $c_1, ..., c_4$ have to be real. Moreover, as the smallest and largest scaling degrees are 0 and -6, $c_2 p_\mu p_\mu \hat{D}^{(+)}$ and the c_3-term have to be non-negative. This implies that $c_2 \geq 0$, as $p_\mu p_\mu \hat{D}^{(+)}$ is non-negative, and $c_3 = 0$, as the term in brackets after c_3 can be shown to be indefinite.[7]

Comparing the degrees of homogeneity again we obtain $\eta_{\mu\mu} c_4 \geq 0$ which implies $c_4 = 0$ by taking $\mu = 0$ and $\mu = i = 1, 2, 3$. Finally, the c_1-term has the lowest scaling degree and therefore has to be non-negative. This implies $c_1 = 0$ because of the indefiniteness of the term in brackets behind c_1. Let us show this here explicitly:
The c_1-term applied to a non-negative test function $u \in \mathcal{S}$ reads

$$\int \left[\eta_{\mu\nu} u(p) - \frac{\lambda}{2(1-\lambda)} \partial_0 \frac{p_\mu p_\nu u(p)}{p_0} \right]_{p_0 = |\vec{p}|} \frac{d\vec{p}}{2|\vec{p}|}. \tag{3.31}$$

Let us look at the 00-component and the trace over the i-components of the term in edgy brackets. Introducing $a = \frac{\lambda}{2(1-\lambda)}$ we obtain

$$u - a\partial_0 p_0 u = (1-a)u - a p_0 \partial_0 u \overset{\text{at } p_0 = |\vec{p}|}{=} (1-a)u - a|\vec{p}|\partial_0 u \tag{3.32}$$

and

$$-u - a\partial_0 \frac{|\vec{p}|^2}{p_0} u = -(1 - a\frac{|\vec{p}|^2}{p_0^2})u - a\frac{|\vec{p}|^2}{p_0} u \overset{\text{at } p_0 = |\vec{p}|}{=} -(1-a)u - a|\vec{p}|\partial_0 u. \tag{3.33}$$

Thus we obtain the same terms apart from a sign-change in the first summand. If there exists a choice of u such that the first summand becomes dominant, the indefiniteness of the whole term is shown. A possible choice of u is $u(p) = w(|\vec{p}|)e^{-p_0^2}$ for a compactly supported smooth function $w : [0, \infty) \to [0, 1]$ with support in $[0, \sqrt{c}]$ for some sufficiently small positive constant c.[8]
What we end up with is

$$W_{\mu\nu} = c_2 \partial_\mu \partial_\nu D^{(+)} \tag{3.34}$$

for a constant $c_2 \geq 0$. Thus the non-negative case is proven. ■

3.4 Proof without spectral condition

For simplicity throughout the proof we assumed the spectral condition to hold true. Although this is the case in the Wightman setting it is not necessarily the case in the general setup that we introduced in the beginning of this section. This is, however, not a big problem, as the proof goes through without spectral condition with small adjustments. The covariant decomposition of W did not assume the spectral condition and is therefore not affected. For the precise form of W by the equations of motion, i.e., in the proof of Lemma 3.8, the spectrum condition is only relevant in the equations (E8), (E10), (E11). There the condition is only used to infer that in the general solution to the homogeneous

[7]The argument here goes as follows: Take test functions $u \in \mathcal{S}$ of the form $u(x) = e^{-|x|^4 + a_{\mu\nu} x^\mu x^\nu}$ with a diagonal matrix $a = diag(b_0, b_1, b_2, b_3)$ and $b_0, ..., b_3$ real constants. Then $\partial_\mu^2 u_{\restriction x=0} = 2b_\mu$ and hence the family of test functions of this form may realize partial derivatives of second order of arbitrary sign. Therefore any tempered distribution $(\sum_\mu c_\mu (\partial_\mu)^2)\delta$ is indefinite. The c_3-term (for $\mu = \nu$) falls into this class as $p_\mu p_\nu \Box^2 \delta = (4\eta_{\mu\nu}\Box + 8\partial_\mu\partial_\nu)\delta$.

[8]Sufficiently small here means that $c < |\frac{1-a}{a}|$ as $|\vec{p}||\partial_0 u|(|\vec{p}|, \vec{p}) = |\vec{p}|w(|\vec{p}|)|\vec{p}|e^{-\frac{1}{2}|\vec{p}|^2} \leq cw(|\vec{p}|)e^{-\frac{1}{2}|\vec{p}|^2} = cu(|\vec{p}|, \vec{p}) < |\frac{1-a}{a}|u(|\vec{p}|, \vec{p})$.

equation $p^2T = 0$ for a Lorentz-invariant tempered distribution T the term proportional to $\hat{D}^{(-)}$ is not appearing. When taking it to account the precise result will be that $W_{\mu\nu} = ... + c_5 \left(\eta_{\mu\nu}D^{(-)} + \frac{\lambda}{1-\lambda}\partial_\mu\partial_\nu x^2 D^{(-)}\right) - c_6\partial_\mu\partial_\nu D^{(-)}$ for some additional constants c_5 and c_6 where the ... represents the terms that were there with assumed spectrum condition. Note that these additional terms are precisely the c_1- and c_2 terms with the only difference that there is a $D^{(-)}$ instead of a $D^{(+)}$. Thus the additional terms will also have the same scaling degrees as the c_1- and the c_2-term. This is not a problem because we can distinguish the scaling behaviour of $\hat{D}^{(+)}$ and $\hat{D}^{(-)}$ by choosing test functions with support in $\bar{V}^+$ and $\bar{V}^-$, respectively. The final result would thus be

$$W_{\mu\nu} = \partial_\mu\partial_\nu \left(aD^{(+)} + bD^{(-)}\right) \tag{3.35}$$

for some non-negative constants a, b. This completes the proof of Lemma 3.8.

3.5 Discussion of the result and comparison with the literature

What we have obtained in the preceding subsections is that the two-point function W of a hermitian covariant vector field A (in an indefinite metric QFT) subject to the equation

$$\Box A_\mu - \lambda\partial_\mu\partial A = 0, \quad \lambda \neq 1 \tag{3.36}$$

is of the form

$$\begin{aligned}\hat{W}_{\mu\nu} = c_1 &\left(\eta_{\mu\nu}\hat{D}^{(+)} - \frac{\lambda}{1-\lambda}p_\mu p_\nu \hat{D}'^{(+)}\right) + c_2 p_\mu p_\nu \hat{D}^{(+)} \\ &+ c_3\left(\eta_{\mu\nu}\Box\delta - \frac{1}{24}\frac{4-\lambda}{1-\lambda}p_\mu p_\nu \Box^2\delta\right) + c_4\eta_{\mu\nu}\delta\end{aligned} \tag{3.37}$$

for constants $c_1, ..., c_4$ and that the assumption of non-negativity of the scalar product of the underlying space implies that the two-point function of the associated field $F_{\mu\nu}(A) = \partial_{[\mu}A_{\nu]}$ vanishes.

Concerning the first result, the form of W, apart from the terms concentrated at $p = 0$, may be found in [Ste00, Chapter 5.3, Eq. (5.83)] and a bit less explicit in [Str13, Chapter 8.2, Eq. (7.8.29)]. The specific case of $\lambda = 0$ may also be found in [WG64, Eq. (2.54)]. Note that the $D^{(+)}$ in the latter two publications is differing from the $D^{(+)}$ here by a factor of i and $\frac{i}{(2\pi)^3}$, respectively. It should be explained why the additional terms here do not show up in the cited publications:

The reason that these additional terms do not appear there is that they are there ruled out by the cluster decomposition principle and the vanishing vacuum expectation value of A. Here we have not made use of it as we did not assume the uniqueness of the vacuum vector which is necessary to infer the cluster decomposition principle.[9]

[9] On a Hilbert state space the cluster decomposition principle is equivalent to the existence and uniqueness of the vacuum vector (see [BLOT90, Proposition 7.1, pp. 276]). In the general indefinite metric case

In order to give the relation to the other publications let us now assume the cluster decomposition principle and apply it to the two-point function of A. The result is

$$\langle \Omega, A_\mu(x) A_\nu(y+ka)\Omega\rangle \underset{k\to\infty}{\to} \langle \Omega, A_\mu\Omega\rangle \langle \Omega, A_\nu\Omega\rangle = 0 \quad \text{for each spacelike } a \in \mathbb{M}. \tag{3.38}$$

Important to note is that $\langle \Omega, A_\mu\Omega\rangle \equiv \langle \Omega, A_\mu(x)\Omega\rangle$ does not depend on x by translation invariance, we therefore obtain a constant which, however, must transform as a vector. This is only possible if the vacuum expectation value $\langle \Omega, A_\mu(x)\Omega\rangle$ vanishes identically.

Now let us remind ourselves that in the final argument to the proof we have established the scaling degrees for the different summands of $W_{\mu\nu}$ in momentum space. The corresponding scaling degrees in coordinate space are given by the negative momentum space scaling degrees -4 (see Proposition 3.11(g)). Thus for the c_1-, c_2-, c_3-, and c_4-terms we have scaling degrees of -2, -4, 2, and 0, respectively. Thus the c_1- and c_2-term drop to zero in the limit $k \to \infty$ and the c_3-term must vanish identically because otherwise the limit $k \to \infty$ would not be finite. The c_4-term must then coincide with $\langle \Omega, A_\mu\Omega\rangle \langle \Omega, A_\nu\Omega\rangle$ or, in other words, must vanish. Thus we end up with

$$W_{\mu\nu} = c_1\left(\eta_{\mu\nu} D^{(+)} + \frac{\lambda}{1-\lambda}\partial_\mu\partial_\nu x^2 D^{(+)}\right) - c_2 \partial_\mu\partial_\nu D^{(+)}. \tag{3.39}$$

This result coincides with the mentioned results in the literature.

In this document, however, we did not make these simplifying assumptions for $W_{\mu\nu}$. In this regard it is nice that the result of a vanishing two-point function of F when non-negativity is assumed is not affected. Thus the result here is a bit more general than in [Ste00], [Str13] and [WG64].
There are also publications that prove the necessity of indefinite metric Hilbert spaces for the covariant quantization of a vector field and especially for the gauge field of QED which, however, do not make recourse to the axiomatic setting. Many of them consider what can be understood as considerations pertaining to the construction of 1-photon states. Notable are the following two papers by Bertrand and Bracci:

The first paper [Ber71] obtains that Poincaré-invariant separately continuous hermitian sesquilinear forms on a certain space of vectorial functions $u_\mu,\ _{\mu=0,1,2,3}$ are either vanishing or indefinite.[10] The space of vectorial functions consists of smooth compactly supported functions with supports in $\partial V^+\backslash\{0\}$. The relation to this document is that the two-point function of a vector field may be understood as a Lorentz-covariant and translation-invariant continuous hermitian sesquilinear form on $\mathcal{S}(\mathbb{M})$ and that Lorentz-covariant sesquilinear forms on $\mathcal{S}(\mathbb{M})$ stand in one-to-one correspondence to Lorentz-invariant sesquilinear forms on the vectorial functions $\mathcal{S}(\mathbb{M})^{\times 4}$ (we mean here 4-tuples $u_{\mu,\mu=0,1,2,3}$ with $u_\mu \in \mathcal{S}(\mathbb{M})$). In order to see this let us denote the former by $S_{\mu\nu}$ and the latter by S. Lorentz-covariance of $S_{\mu\nu}$ means that

the cluster decomposition principle may fail even in the case of a unique vacuum vector. For this and more on the cluster decomposition principle on indefinite metric spaces the reader is referred to [Str78, Section V].

[10] In the notation of the reference the sesquilinear form is denoted by B and most of the time the μ-index for the vectorial functions is omitted.

$$S_{\mu\nu}(u_\Lambda, v_\Lambda) = \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma (u, v), \quad u, v \in \mathcal{S}(\mathbb{M}) \tag{3.40}$$

where $u_\Lambda(x) \equiv u(\Lambda^{-1}x)$. Lorentz-invariance of S means that

$$S(u_\Lambda, v_\Lambda) = S(u, v), \quad u, v \in \mathcal{S}(\mathbb{M})^{\times 4}. \tag{3.41}$$

The one-to-one correspondence is given by associating $B(u, v) \equiv B_{\mu\nu}(u^\mu, v^\nu)$ to each $B_{\mu\nu}$ and by associating $B_{\mu\nu}(u, v) \equiv B(ue_\mu, ve_\nu)$ to each B (e_μ, $\mu=0,1,2,3$, denote the standard basis vectors of $\mathbb{R}^4$).
Note that in the first case W is 2-tensor-valued and u is complex-valued where in the second case W is complex-valued and u is vector-valued.
That they restrict to test functions supported within $\partial V_\times^+ \equiv \partial V^+ \backslash \{0\}$ is the same as to consider tempered distributions $\mathcal{S}'(\partial V_\times^+)$.[11] Converted into the notation and mathematical objects which are used here the referenced paper [Ber71], in particular eq. (24), classifies Lorentz-covariant Lorentz-2-tensor-valued tempered distributions $\hat{W}$ on $\partial V_\times^+$ to be of the form

$$\hat{W}_{\mu\nu}(p) = a\eta_{\mu\nu}\hat{D}^{(+)}(p) + bp_\mu p_\nu \hat{D}^{(+)}(p) \tag{3.42}$$

where $\hat{D}^{(+)}$ is understood here as elements of $\mathcal{S}'(\partial V_\times^+)$. When extending this result to zero, one will achieve the same result as obtained here. The difference between the result of the paper and the result in this document is that the paper looks only on $\partial V_\times^+$ and not on the whole of $\mathbb{M}$, but therefore does not have to make explicit use of the equations of motion that were used in this document.

In the second paper [Bra72] again unitary Poincaré representations on vectorial functions on $\mathbb{M}$ are discussed.[12] The result is that a positive definite separately continuous hermitian sesquilinear form (i.e., a sep. cont. inner product) on a space of vectorial functions (that is spin 1 or rather helicity ± 1) transforming under a massless unitary Poincare representation requires the functions to be gradientlike, i.e., $u_\mu(p) \equiv p_\mu u(p)$ for some function u. This corresponds to the result that a Lorentz-covariant Lorentz-2-tensor-valued tempered distribution $\hat{W}$ which is positive definite is of the form

$$\hat{W}_{\mu\nu}(p) = p_\mu p_\nu K \tag{3.43}$$

for some Lorentz-invariant tempered distribution K. Thus the result agrees with what we found here. Interesting is that in the reference there is a very similar result for massless spin 2 representations which are relevant for gravitons. In the outlook of this document there will be more about this.

It would also be interesting to discuss the different escape routes from the presented no-go-theorems. We will postpone this discussion to the summary chapter, Chapter 5. At first we will continue with the implications for the interacting case of QED which can be drawn from this result.

[11] Note that so far we have only introduced tempered distributions restricted to open subsets of $\mathbb{R}^n$ (or $\mathbb{M}$). We can still apply this case by observing that $\mathbb{R}^3 \cong \partial V^+$ by the map $\vec{p} \mapsto j(\vec{p}) \equiv (|\vec{p}|, \vec{p})$. Then we characterize distributions on ∂V^+ by: $T \in \mathcal{S}'(\partial V^+) \Leftrightarrow T : \mathcal{S}(\mathbb{M}) \to \mathbb{C}$ such that $\exists S \in \mathcal{S}'(\mathbb{R}^3) : T = S \circ j^{-1}$. Then the restriction to $\partial V_\times^+$ corresponds to the restriction from $\mathbb{R}^3$ to the open subset $\mathbb{R}^3_\times$.

[12] In the reference the vectorial functions are denoted by $A_\mu(K)$.

Chapter 4

Implications for the interacting case

In the former section we have seen the completion of the proof that a formulation of free QED which is insisting on hermiticity, covariance, free evolution and non-negativity of the two-point function of A, necessarily leads to a trivial theory in the sense that the two-point function of the observable field $F_{\mu\nu}$ vanishes.

What we are finally up to is to draw conclusions for the interacting case of QED. As it is well known that there are no (non-perturbative) constructions of interacting quantum field theories available in 1+3 dimensions so far, our conclusions take the form of a no-go theorem. That is, we will show that a gauge formulation of an interacting theory of QED that relies on covariance and non-negativity of the two-point function of A is trivial in a certain sense specified below.

As we will see, the strong relation between the free and the interacting theory will make it very hard to construct an interacting theory without having a satisfactory free version of it at hand. Scattering theory provides a strong relation between the free and the interacting theory (not only in QED). In scattering theory the asymptotic in- and out-fields are expected to be free. In standard scattering theory, however, the interaction is assumed to be short-ranged. QED is a limiting case of this class of interactions and falls just out of it. Although a scattering theory for massless particles is existent [Buc75, Buc77][original work in algebraic setting] and [Str90][Wightman setting], this is much more delicate than the standard setting of scattering theory and it is in fact not really necessary to go into this construction for our purposes here.

What we will focus on is based on the idea that also the interacting theory should have sectors, which are completely free. In particular, the one-particle sectors should be free of any dynamics. In the usual interpretation of elementary quantum fields the one-particle states are generated by the action of the elementary quantum field on the vacuum.

In the setting here, the elementary quantum field which is supposed to generate the one-photon states is the Maxwell field $F_{\mu\nu}$. The space of the one-photon states should be a subset of the space of massless states. What we will show here, is that $F_{\mu\nu}$ is not able to create any massless states from the vacuum. With the usual interpretation of F such a theory is inevitably non-satisfactory.

4.1 The setup

In order to make the above statement an adequate theorem, let us clarify the setup and write down some proper definitions. For simplicity we will now start assuming the condition of a positive definite metric right away and do not aim for full generality. Let us begin by properly defining the subspace of massless states:

Definition 4.1. Suppose a relativistic quantum theory with vacuum vector $(\mathcal{H}, U, \Omega)$ satisfying the spectrum condition. Then the subspace of massless states is given by

$$\mathcal{H}^{(1)} \equiv \chi_{\partial V^+ \backslash \{0\}}(P)\mathcal{H}, \tag{4.1}$$

where P is the energy-momentum operator associated to the translation group representation induced by U and where χ denotes the characteristic function on $\mathbb{M}$.

Remark 4.2. The operator $\chi_{\partial V^+ \backslash \{0\}}(P)$ is defined by the functional calculus for finitely many strongly commuting self-adjoint operators and is basically a spectral projection such that the P_μ will have a joint spectrum lying inside $\partial V^+ \backslash \{0\}$. $\partial V^+ \backslash \{0\}$ is precisely the subset of $\mathbb{M}$ with massless dispersion relation and positive energy. In a theory with spectral condition this is equivalent to the space of massless states (without the zero-energy state/vacuum). The existence of strongly commuting self-adjoint operators P_μ is guaranteed by Stone's theorem. For the functional calculus of self-adjoint operators and Stone's theorem the reader is referred to Section 2.5 on unitary Hilbert space representations.

In order to convince us that this definition is reasonable let us prove the following result

Proposition 4.3.

$$\mathcal{H}^{(1)} \cup kerP = kerP^2 \tag{4.2}$$

where $ker\ A \equiv \{\phi \in D : A\phi = 0\}$ *for any* $A \in \mathcal{L}(D, \mathcal{H})$.

Proof. $\mathcal{H}^{(1)} \subset kerP^2$: It is clear that $p^2\chi_{\partial V}(p) = 0$ and thus

$$0 = I(p^2\chi_{\partial V^+ \backslash \{0\}}) = \overline{I(p^2)I(\chi_{\partial V^+ \backslash \{0\}})} = \overline{P^2\chi_{\partial V^+ \backslash \{0\}}(P)}.$$

$kerP^2 \subset \mathcal{H}^{(1)} \cup kerP$: Let $\phi \in kerP^2 \backslash kerP$ be arbitrary. Then by the spectrum condition we know that $(P_0 + |\vec{P}|)\phi \neq 0$[1] and by the commutativity of the P_μ's then $P^2\phi$ implies $\phi \in kerP$ or $\phi \in ker(P_0 - |\vec{P}|)$. This amounts to $\sigma(P\restriction_{kerP^2}) = \partial V^+$ and thus

$$\chi_{\partial V^+}(P)kerP^2 = \chi_{\partial V^+ \backslash \{0\}}(P)kerP^2 + \chi_{\{0\}}(P)kerP^2 \subset \mathcal{H}^{(1)} \cup kerP.$$

■

Consider now a Wightman-theory of QED $(\mathcal{H}, U, \Omega, D, \{F_{\mu\nu}, J_\nu\})$ with an additional hermitian covariant vector field (D, A) such that $F_{\mu\nu} = \partial_{[\mu}A_{\nu]}$. Let $\mathcal{H}^{(1)}$ and $P^{(1)}$ denote the subspace of massless states as specified above and the projection onto it, respectively.

[1]Note that $|\vec{P}|$ is the unique positive square root of the operator $|\vec{P}|^2 = \vec{P}^2$. For this the self-adjointness of $\vec{P}$ is required.

4.2 The properties of the 'massless' two-point function

Then define the 'massless' two-point-function $W^{(1)}_{\mu\nu}$ to be

$$W^{(1)}_{\mu\nu}(\xi) \equiv (\Omega, A_\mu(x)P^{(1)}A_\nu(y)\Omega) \tag{4.3}$$

for $\xi = x - y$. For this $W^{(1)}$ we can derive the following conditions that could appear somewhat familiar:

(a) transforms covariantly under the representation $\mathcal{D}^{(1/2,1/2)} \otimes \mathcal{D}^{(1/2,1/2)}$ of SL(2,$\mathbb{C}$)

(b) $W^{(1)}$ satisfies the free massless KG equation, namely

$$\Box W^{(1)}_{\mu\nu} = 0. \tag{4.4}$$

(c) The components $\hat{W}^{(1)}_{\mu\mu}$ are non-negative, namely

$$\hat{W}^{(1)}_{\mu\mu}(u) \geq 0 \quad \forall \text{ non-negative } u \in \mathcal{S}. \tag{4.5}$$

Concerning

(a) From the covariance of A_μ we see that

$$\begin{aligned}
W^{(1)}_{\mu\nu}(x+a, y+a) &= (\Omega, U(0,a)A_\mu(x)U(0,a)^{-1}P^{(1)}U(0,a)A_\nu(y)U(0,a)^{-1}\Omega) \\
&= (\Omega, A_\mu(x)U(0,a)^{-1}P^{(1)}U(0,a)A_\nu(y)\Omega) \\
&= (\Omega, A_\mu(x)A_\nu(y)\Omega) \\
&= W^{(1)}_{\mu\nu}(x,y)
\end{aligned}$$

and therefore $W^{(1)}$ is translation-invariant and the definition above in one variable ξ is justified. We used here the invariance of Ω and of the domain D, as well as that $P^{(1)}$ and $U(g)$ commute. The commutativity of $P^{(1)}$ and $U(g)$ is a consequence of the functional calculus of commuting self-adjoint operators. The Lorentz covariance of $W^{(1)}$ follows in the same way.

(b) We will prove that $\Box P^{(1)}A_\mu(u)\Omega = 0$ for each μ and each $u \in \mathcal{S}$:

$$\begin{aligned}
\Box P^{(1)}A_\mu(\cdot)\Omega &= P^{(1)}A_\mu(\Box\cdot)\Omega && \text{(by def)} \\
&= -iP^{(1)}[P_\nu, A_\mu(\partial^\nu\cdot)]\Omega && \text{(by Heisenberg eom)} \\
&= -iP^{(1)}P_\nu A_\mu(\partial^\nu\cdot)\Omega && \text{(by inv. of } \Omega) \\
&= -P^{(1)}P^2 A_\mu(\cdot)\Omega && \text{(same as above)} \\
&= -P^2 P^{(1)}A_\mu(\cdot)\Omega && \text{(by strong commutativity and functional calc.)} \\
&= 0 && \text{(by def of } P^{(1)}).
\end{aligned}$$

(c) The non-negativity condition is not affected by the insertion of the orthogonal projection $P^{(1)}$ and is a consequence of the hermiticity of A and the positive-definiteness of $(\cdot, \cdot)$.

With the same conditions at hand as in the case of free QED (for $\lambda = 0$) we can infer that a two-point function of F with support in the massless spectrum only (this is due to the projection $P^{(1)}$) vanishes:

$$B^{(1)}_{\mu\nu\rho\sigma}(x-y) \equiv (\Omega, F_{\mu\nu}(x)P^{(1)}F_{\rho\sigma}(y)\Omega) = 0. \tag{4.6}$$

From this we can further derive that

$$||P^{(1)}F_{\mu\nu}(u)\Omega||^2 = B^{(1)}_{\mu\nu\mu\nu}(\tilde{\bar{u}} * u) = 0 \quad \Leftrightarrow \quad P^{(1)}F_{\mu\nu}(\cdot)\Omega \sim 0. \tag{4.7}$$

The rhs denots that $P^{(1)}F_{\mu\nu}(\cdot)\Omega$ is equivalent to 0 on the induced physical Hilbert space (zero-norm states are divided out there). Thus on the induced physical Hilbert space we have that $P^{(1)}F^{\mathrm{ph}}_{\mu\nu}(\cdot)\Omega^{\mathrm{ph}} = 0$ where $^{\mathrm{ph}}$ marks the induced objects on the physical Hilbert space. The Reeh-Schlieder property of the vacuum implies that

$$P^{(1)}F^{\mathrm{ph}}_{\mu\nu} = 0 \Leftrightarrow \langle \mathcal{H}^{(1),\mathrm{ph}}, F^{\mathrm{ph}}_{\mu\nu}(\cdot)\Omega\rangle = 0. \tag{4.8}$$

Thus F is not able to create any (physical) massless states, specifically single-photon states, from the vacuum. With the usual interpretation of quantum electrodynamics this situation cannot lead to a satisfactory formulation of the theory.

4.3 Discussion of the results and comparison with existing literature

In this chapter of the document we showed that if we suppose a (possibly interacting) Wightman theory of a hermitian 2-tensor field F for which there exists a hermitian (covariant) vector field A such that $F = \partial_{[\mu}A_{\nu]}$, then F will not be able to generate massless states from the vacuum. We concluded that this makes covariant gauge formulations of QED on Hilbert spaces unsatisfactory and that it indicates the need for indefinite metric state spaces. In this section we shall briefly raise discussion points concerning this statement.

First let us note the following: That it is unsatisfactory for QED that F generates no massless states from the vacuum relies heavily on the strength of the relation between free and interacting theory. Specifically, it is dependent on how relevant the space of massless states is in an interacting theory (of QED). It could in principle be that interacting 'photons' are not particles with sharply vanishing mass and thus that F does not generate massless states would be not of great interest. However in the standard approaches to QED and also in most of the other cases where massless bosons arise we expect their mass to be sharply defined.[2]

[2]In [Buc77] for instance it is argued that there are basically two mechanisms leading to the appearance of massless bosons in the theory. Namely, spontaneous symmetry breaking which gives rise to the so-called Goldstone bosons and the appearance of gauge bosons whenever local gauge symmetries are present. In both cases there are no indications that the bosons have a not sharply vanishing mass. Also perturbative QED gives strong heuristic arguments.

If one accepts that what we are doing is indeed physically relevant, we can also argue that it is mathematically relevant. The Källen-Lehmann-representation for the non-negative(!)[3] two-point function $B_{\mu\nu\rho\sigma}(x-y) = (\Omega, F_{\mu\nu}(x)F_{\rho\sigma}(y)\Omega)$ of a hermitian covariant antisymmetric 2-tensor field F subject to the spectral condition reads

$$\hat{B}_{\mu\nu\mu\nu}(p) = \int \left(-p_\mu^2\eta_{\nu\nu} - p_\nu^2\eta_{\mu\mu}\right)\hat{D}_m^{(+)}(p)d\mu(m) \tag{4.9}$$

for some positive measure $d\mu$ on $\bar{\mathbb{R}}^+$ and the positive frequency Pauli-Jordan commutation function $\hat{D}_m^{(+)}(p) \equiv \theta(p_0)\delta(p^2-m^2)$ of mass m.[4] Let us require μ to be only supported at $m=0$, i.e., $d\mu(m) = a\delta(m)dm$ for some $a \geq 0$. Then the Källen-Lehmann representation gives

$$\hat{B}_{\mu\nu\mu\nu}(p) = a\left(-p_\mu^2\eta_{\nu\nu} - p_\nu^2\eta_{\mu\mu}\right)\hat{D}^{(+)}(p). \tag{4.10}$$

Thus the requirements on $B_{\mu\nu\rho\sigma}$ allow for a non-trivial two-point function of $P^{(1)}F$. In other words, the result of F not generating massless states from the vacuum is non-trivial.

The discussion on how to evade the necessity to introduce indefinite metric state spaces will be conducted in the next chapter.

[3] Remember that we are talking about physical fields now and thus requiring non-negativity makes sense.

[4] See [Buc86, Eq. (4)] or [Str13, Eq. (7.5.10)] for reference.

Chapter 5

Summary, final discussion and outlook

This last chapter aims to summarize the results of this thesis and put them into perspective. Moreover, we give an outlook on related topics beyond this document.

The result of this thesis is that covariant gauge formulations of QED (with standard interpretation) need to have recourse to indefinite metric Hilbert spaces. In particular we have obtained two main results. Firstly, that a free hermitian covariant vector field giving rise to a non-negative two-point function needs to be gradientlike, i.e. of the form $\partial_\mu \phi$. The consequence is that the field $F_{\mu\nu}$ vanishes (on the induced physical Hilbert space) and therefore the vector field cannot be part of a satisfactory description of free QED. Secondly, that in QED on a state space with non-negative metric the field $F_{\mu\nu} = \partial_{[\mu} A_{\nu]}$, defined in terms of a hermitian covariant vector field A, cannot generate massless states from a vacuum (within the induce physical Hilbert space).

The first result can be found in different variants in the literature. Detailed proofs like presented here are however scarce.[1] Moreover, it is new that we do not assume the cluster decomposition principle. This gave rise to two extra terms in the two-point function of A which are however still prohibited by the non-negativity condition.

The author is not aware of any explicit mention of the second result in the literature. In most of the literature only free QED is mentioned and the consequences for interacting QED are kept mostly implicit. In only some parts of the literature there are remarks to relate free QED with the asymptotic in or out fields of an interacting QED without much further explanation.[2]

[1]The references to existing literature were mostly given in the discussion section of the free QED case, Section 3.5.

[2]See e.g. [NO90, Section 2.2.2, footnote at p. 57]

What is a bit unsatisfactory about the proofs is that we still have to assume that explicit equations of motions for A, i.e., $\Box A_\mu - \lambda\partial_\mu\partial A = 0$, hold. That this is the most general equation of motion for a covariant gauge formulation was merely a heuristic argument and not a proof. It seems however unlikely to derive a triviality result, requiring only hermiticity, covariance, and non-negativity of the two-point function of A, with the same methods applied here. The problem is that one then has no information to compare the contributions of $\hat{K}$ and $\hat{G}$ to the two-point function $\hat{W} = \eta_{\mu\nu}\hat{K} - p_\mu p_\nu \hat{G}$. If $\hat{G}$ can be arbitrarily large it is unclear how to infer $\hat{K} = 0$ from the non-negativity condition.

5.1 Routes to evade the necessity of an indefinite metric state space

We should also summarize the routes to evade the no-go theorems presented here. The most obvious one is to let go of the covariance of the vector field. As A is a non-physical field it is allowed to have a more general transformation behaviour like

$$U(a,\underset{\sim}{\Lambda})A_\mu(u)U(a,\underset{\sim}{\Lambda})^{-1} = (\Lambda(\underset{\sim}{\Lambda})^{-1}A)_\mu(u_{a,\Lambda}) + \partial_\mu B(\underset{\sim}{\Lambda}, u), \quad u \in \mathcal{S} \tag{5.1}$$

for an arbitrary operator-valued B which is a function with respect to $\underset{\sim}{\Lambda}$ and a tempered distribution with respect to u. The use of such non-covariant fields is however linked with very inconvenient computations and many of the standard results for dealing with quantum fields (like extension of analyticity domains of the Wightman functions) have to be proven anew.[3]

Another way to evade the theorems is to violate the hermiticity of A. It is instructive here to look at the concrete example of the free vector potential in the Gupta-Bleuler formalism.[4] To begin with, let us sketch a part of the ordinary construction (for $\lambda = 1$). Thus we have $\Box A_\mu = 0$. In generalized function notation we obtain

$$A_\mu(x) = \sum_{\lambda=0}^{3} \int \left(\epsilon_\mu^{(\lambda)}(p)a^{(\lambda)}(p)e^{-ipx} + \epsilon_\mu^{(\lambda)*}(p)a^{(\lambda)\dagger}(p)e^{ipx} \right) d\Omega(p) \tag{5.2}$$

where $d\Omega(p) \propto \theta(p_0)\delta(p^2)dp$ is the (up to normalization) unique Lorentz-invariant measure on the massless mass shell, $\epsilon^{(\lambda)}(p)$ denotes a set of four (with respect to η) orthonormal p-dependent polarization vectors, and $a^{(\lambda)}(p)$ and $a^{(\lambda)\dagger}(p)$ are the annihilation and creation operators. $(\ )^\dagger$ denotes the adjoint with respect to the indefinite inner product of the Krein space. Note that $a^{(\lambda)}(p)$ and $a^{(\lambda)\dagger}(p)$ may be understood as generalized functions in $\vec{p}$.

Covariant canonical quantization[5] leads to

[3] See [Str67, Eq. (28) and below.].

[4] We will follow [Str13, Section 7.8.2] here.

[5] ... of the gauge-fixed Lagrangian $\mathcal{L}_{GB} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}(\partial A)^2$. This ensures non-vanishing momentum conjugate to A_0 given by $\Pi_0 = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_0)} = \partial^\mu A_\mu$. This choice of gauge-fixing term corresponds to a choice of deformation term $\mathcal{L}_\mu^{\text{deform}} = -\partial_\mu\partial A$ and thus to Feynman-Gupta-Bleuler gauge $\lambda = 1$. See also in the introductory section for this.

$$[A_\mu(x), \partial_0 A_\nu(y)]_{x_0=y_0} = -i\eta_{\mu\nu}\delta(\vec{x}-\vec{y}) \tag{5.3}$$

and by decomposition into positive and negative energy parts implies covariant canonical commutation relations for the corresponding annihilation and creation operators:

$$[a^{(\kappa)}(\vec{p}), a^{(\lambda)\dagger}(\vec{q})] = -\eta^{\kappa\lambda}|\vec{p}|\delta(\vec{p}-\vec{q}). \tag{5.4}$$

This relation clearly implies that $||a^{(0)\dagger}(\cdot)\Omega|| < 0$ due to $-\eta^{00} = -1$ where Ω is the vacuum state defined by $a^{(\lambda)}(\cdot)\Omega = 0$.

Let us choose the polarization vectors $\epsilon^\lambda(p)$ to be timelike ($\lambda = 0$), transversal ($\lambda = 1, 2$) and longitudinal ($\lambda = 3$) with respect to p. For appropriate normalization of the polarization vectors the physicality condition

$$\partial^\mu A^{(-)}_\mu(\cdot)\psi = 0 \tag{5.5}$$

is equivalent to demand

$$p^\mu a_\mu(\cdot)\psi = |\vec{p}|\left(a^{(0)} - a^{(3)}\right)(\cdot)\psi = 0, \quad p_0 = |\vec{p}|. \tag{5.6}$$

Thus the physical subspace will contain an equal amount of scalar and longitudinal photons for each momentum $\vec{p}$. The induced physical Hilbert space will be obtained by identifying zero-norm states.

We are ready now to state the proposed evasion of the consequence of an indefinite metric state space. The proposal is to replace $a^{(0)}$ with $ia^{(0)}$ and $a^{(0)\dagger}$ with $ia^{(0)\dagger}$ in A.[6] In this way a 'part' of A becomes anti-hermitian and A in total is not hermitian anymore.[7] The advantage is that the derived canonical commutation relations change to

$$[a^{(\kappa)}(p), a^{(\lambda)\dagger}(q)] = \delta^\kappa{}_\lambda|\vec{p}|\delta(\vec{p}-\vec{q}) \tag{5.7}$$

such that the minus sign of the 00-component disappears and quantization can be carried out on a state space with non-negative states. In particular we obtain now $||a^{(0)\dagger}(\cdot)\Omega|| > 0$). The non-hermiticity of A does not need to cause problems because on physical states the physical field $F_{\mu\nu}$ will remain hermitian.[8] There will be no problem with energy because the contributions from scalar and longitudinal photons cancel each other.

What we have basically done here is that we have exchanged the hermiticity of the sesquilinear form $\langle\cdot,\cdot\rangle$ with non-negativity. In this formalism we can define a new type of conjugation K with respect to which A becomes hermitian again. Representing the conjugation by a bounded invertible self-adjoint operator $\eta = \eta^\dagger, \eta^2 = 1$ we write $A^K = \eta A^\dagger \eta$. Redefining the sesquilinear form accordingly, we obtain

$$\langle\phi,\psi\rangle_{\text{new}} \equiv \langle\phi,\eta\psi\rangle \tag{5.8}$$

[6]See e.g. [Sch89, Chapter 2.11, pp. 119]. Note that in the reference a special coordinate frame is chosen such that $a_0 = a^{(0)}$.

[7]Choosing for instance $\epsilon^{(0)}_{,\mu}(p) = (1,0,0,0)$ would yield A_0 to be anti-hermitian und A_j, $j = 1,2,3$ to be hermitian.

[8]One representative space of the physical Hilbert space is the subspace with no scalar and no longitudinal photons. See the reference from above.

and thus regain the notion of an indefinite metric space. Therefore the hermitian indefinite metric case and the antihermitian (referring to the 0-component of A) non-negative metric case are two representations of the same situation and the choice of one or the other is then usually a matter of convenience.

An alternative proposal is to exchange the role of $a^{(0)}$ and $a^{(0)\dagger}$ such that $a^{(0)\dagger}$ annihilates the vacuum. In this way the sign-flip is achieved, as well, by the antisymmetry of the commutator under exchange. The problem is now that applications of $a^{(0)}$ to the vacuum successively lead to lower and lower energies opposed to the usually increasing energy. Thus the Hamiltonian is unbounded from below.[9]

Finally, there is the 'intended' way out, which is to accept that the theory has to be defined on an indefinite metric state space. This is the most common choice including the popular Gupta-Bleuler formalism. Problematic with these approaches is that the analysis on indefinite metric spaces becomes more complicated, that the physicality condition is not under good control in the interacting case and that it therefore also becomes hard to interpret. In the end the choice of route if indefinite metric state space, non-hermitian vector field, a Hamiltonian which is unbounded from below is a matter of practicability as in all cases physical consequences remain unchanged.

5.2 The necessity of an indefinite metric state space in local gauge formulations of QED

What we did not really speak about in this document (so far) is the other property which one would like to have for A, namely locality. There are also statements in the literature that a local gauge formulation of QED implies the necessity of an indefinite metric state space. The terminology in the literature is however a bit ambiguous and sometimes the covariance of A is also assumed. Here we will still follow the convention of this document: We assume only locality and hermiticity of A if we speak about a local gauge formulation of QED.
The only references which the author is aware of that formulate such a statement are [Str70, Section VII], [Str13, Proposition 3.3, p. 155], and [FPS74, Theorem 2]. In the first reference we find two relevant statements. The first one, which is however given without any motivation or sketch of proof, is that in a local gauge formulation of free QED the physical states, i.e. the states on which Maxwell's equation are satisfied in terms of expectation values, cannot form a dense subset of the state space. The second one is that a non-trivial local *and* covariant gauge formulation of free QED necessarily is defined on an indefinite metric state space. For the proof the reader is referred to the paper [WG64] which was already referenced in this document. There we will however only find a proof of the necessity of indefinite metric state spaces for the Gupta-Bleuler formalism in Feynman gauge ($\lambda = 1$, $\Box A_\mu = 0$) which was already discussed here. Although this choice ($\lambda = 1$) might be in fact the only local and covariant gauge formulation of QED, the statement remains dubious to the author. In the second reference what is referred to as 'local FGB quantization' actually assumes the covariance of A. It can be considered as a special case application of the result of the third reference. Let us briefly present the

[9]For reference the reader is referred to [Sch61, Chapter 9b, in particular p.244 below eq. (24)].

result of the third reference (Lemma 1 and Theorem 2 of the reference). Note that we will choose here a slightly different style of presentation of the result than in the reference:

The lemma runs as follows: Suppose a local field j_ν such that $j_\nu = \partial^\mu F_{\mu\nu}$ for some antisymmetric local field $F_{\mu\nu}$. Moreover, j and F are assumed to have a common dense and invariant domain D of a state space with a possibly indefinite metric. Then for any field Φ with D as an invariant domain which is local with respect to j_ν and $F_{\mu\nu}$ we obtain

$$\lim_{R\to\infty}[Q_R, \Phi(u)] = 0, \quad u \in \mathcal{S} \tag{5.9}$$

where Q_R can be considered as a charge generator and is defined by $Q_R \equiv j_0(wf_R)$ for compactly supported functions $R > 0$, $w \in \mathcal{D}(\mathbb{R})$, and $f_R \in \mathcal{D}(\mathbb{R}^3)$ such that $f_R(\vec{x}) = 1$ for $|\vec{x}| < R$ and $f_R(\vec{x}) = 0$ for $|\vec{x}| > R + \epsilon$ for some $\epsilon > 0$.

The proof of the lemma consists of not too difficult considerations on the support properties of the participating fields.
Without bigger problems one can weaken the statement by not assuming any invariance of the domain, but instead sandwiching eq. (5.9) between states of the domain.[10]

What this result establishes is that for QED one cannot expect Maxwell's equation $j = \partial F$ to hold on the whole state space - unless one is willing to have no local charged fields in the theory. Thus we have to modify the equations of motion by a deformation term $\mathcal{L}_\nu$ such that we have $j = \partial F - \mathcal{L}$. Let us take now an arbitrary subspace $\mathcal{H}'$ fulfilling only the condition that there exists a dense domain $D \subset \mathcal{H}'$ which is invariant under j and F. Then this implies that D is also a dense and invariant domain to $\mathcal{L} = \partial F - j$. The aim is now to apply the lemma again and to show that $j = \partial F$ cannot hold on $\mathcal{H}'$.

The deformation term can be interpreted as an additional fictitious current: $j = \partial F - \mathcal{L} \Leftrightarrow j' = j + \mathcal{L} = \partial F$. If $\mathcal{L}$ is local, then so is the resulting current j'. The lemma applies now to j' with charge $Q'_R \equiv Q_R^{(fict)} + Q_R$ where $Q_R^{(fict)} = \mathcal{L}_0(wf_R)$. Let Φ_q be a field of charge $q \neq 0$ with domain D. Then for states $\varphi, \psi \in D$ we have

$$-q\,\langle\varphi, \Phi_q(u)\psi\rangle = \lim_{R\to\infty}\langle\varphi, [Q_R, \Phi_q(u)]\psi\rangle = \lim_{R\to\infty}\langle\varphi, [Q_R^{(fict)}, \Phi_q(u)]\psi\rangle\,, \quad u \in \mathcal{S} \tag{5.10}$$

If $(\partial F - j)_\nu(\cdot)D = \mathcal{L}_\nu(\cdot)D = 0$ then this implies that Φ_q vanishes on expectation values within $\mathcal{H}'$. Therefore, unless all the charged fields of QED have vanishing expectation values between states in D, Maxwell's equations cannot hold as operator equations.[11]

If the existence of local charged fields (with non-trivial expectation values) in QED is assumed, then there are now two consequences of this result. The first is: Maxwell's equations can at most hold as expectation values on a subspace of the whole state space. If we demand that they are indeed satisfied as expectation values on some physical sub-

[10] That the resulting expressions are well-defined is then a consequence of the implicit assumption for quantum fields that the domain of the field is contained in the domain of the adjoint field.

[11] Note that one can also dispose of the assumption of D being an invariant domain to F, j and Φ_q and obtain the same result. This is done in the appendix of [FPS74].

space, then the second consequence is: There exist zero- and negative-norm states[12] in the state space, thus the metric is necessarily indefinite.

For the second statement the argument runs as follows: As $\mathcal{L}_\nu(u)D \neq 0$ for some $u \in \mathcal{S}$, there exists $\phi \in D$ such that $\psi \equiv \mathcal{L}_\nu(u)\phi \neq 0$. ψ is a zero-norm state as

$$\langle \psi, \psi \rangle = \langle \psi, \mathcal{L}_\nu(u)\phi \rangle = \lim_{n\to\infty} \langle \psi_n, \mathcal{L}_\nu(u)\phi \rangle = 0 \tag{5.11}$$

for some sequence $(\psi_n)_n \subset D$ converging to ψ. The existence of such a sequence is implied by the density of D and that the rhs-term vanishes by assumption. There have to be negative-norm states as well because positive semi-definiteness and non-degeneracy of the sesquilinear form imply its positive definiteness which is incompatible with the existence of zero-norm states.[13]

In this regard we obtain here a complementary result to what was obtained in this thesis. We have shown that in any covariant gauge formulation of QED one needs to have recourse to an indefinite metric state space. This result tells us that in any local gauge formulation of QED one needs to have recourse to an indefinite metric. It is an interesting, but somewhat peculiar feature of this result, that it relies on the existence of local charged fields and thus does not apply to free QED. As a conclusive statement we can formulate:

In a gauge formulation of QED the (hermitian) gauge field is either non-local and non-covariant or it has to be defined on a state space equipped with an indefinite metric.

5.3 An outlook beyond

There is a very interesting and recent proposal which tries to overcome the conflict between locality and covariance of A on the one side, and non-negativity and physicality of the state space on the other side. The proposal is a formalism of so-called string-localized fields.[14] Starting with a non-gauge formulation of QED, one may define a string-local vector potential

$$A_\mu(x, e) = \int_0^\infty ds F_{\mu\nu}(x + se)e^\nu, \tag{5.12}$$

where e is some additional Minkowski direction. In general A will be a generalized function in x and in e. The resulting field $F_{\mu\nu}(x, e) = \partial_{[\mu}A_{\nu]}(x, e)$ can in fact be shown to be independent of e and to coincide with $F_{\mu\nu}(x)$. Moreover, A satisfies so-called string-locality, i.e., (in the sense of weak commutativity on some common dense and invariant domain)

$$[A_\mu(x, e), A_\nu(y, f)] = 0 \tag{5.13}$$

[12] Meaning that the scalar square of the hermitian sesquilinear form of the state space is zero and negative, respectively

[13] See footnote [34] on p. 35.

[14] The presentation follows [JM17]. For the details the reader is referred to that reference.

whenever $x + \mathbb{R}_0^+ e$ is spacelike with respect to every point on the other string $y + \mathbb{R}_0^+ f$. This notion is clearly weaker than pointlike locality. A also satisfies a different covariance condition

$$U(a,\Lambda)A_\mu(x,e)U(a,\Lambda)^{-1} = (\Lambda A)_\mu(\Lambda x + a, \Lambda e). \tag{5.14}$$

In this way the formalism of string-localized fields may escape the conflict which is overshadowing gauge formulations of QED, in particular, and of QFT, in general.

Finally, there is another topic that deserves mention, namely linearized gravity in quantum field theory, or in other words, gravitons. Gravitons are massless spin-2 particles which arise as gauge bosons of the gravitational interaction in the weak-field approximation (linearized gravity). The methods which were used in this document should be applicable to gravitons, too. Let $h_{\mu\nu}$ be the dynamic part[15] of the metric tensor of Einstein gravity which is assumed to be small. For a quantized field we therefore supposed the corresponding linearized free (i.e., the energy-momentum tensor is assumed to be vanishing) field equations to hold, we demand that it is hermitian and it transforms covariantly as a Lorentz-2-tensor. The relevant two-point function would be

$$W_{\mu\nu\rho\sigma}(x-y) = \langle \Omega, h_{\mu\nu}(x) h_{\rho\sigma}(y)\Omega \rangle \,. \tag{5.15}$$

This two-point function can be shown to be of a certain simple form which yields a trivial two-point function of the Riemann tensor.[16] This result is in the spirit of the former triviality result of free QED with Maxwell's equations satisfied on the whole state space and is again independent of the definiteness of the metric. It is possible that similar methods as applied here (however more complicated computations) could show a similar result as obtained here.

[15]For a metric tensor $g_{\mu\nu}$ we assume a decomposition into $g = g^{(0)} + h$ where $g^{(0)}$ is a constant metric tensor and h are 'small' variations from it.

[16]See [Str68].

Appendix

The appendix collects a bunch of different material that I gathered and wrote down to learn about the foundations and background of the proof in detail. Most of the material is either basic or not specifically tailored to be relevant to the proof.

A Proof of additional properties for tempered distributions

This is a collection of basic results and properties on tempered distributions. They are mainly based on exercises within [BLOT90, Chapter 2].

Proposition A.1. *(Tempered distribution concentrated at the origin) Let $T \in \mathcal{S}'$ be concentrated at the origin and $u \in \mathcal{S}$ an arbitrary test function. Then T has the form*

$$T(u) = \sum_{|\alpha| \leq N} c_\alpha D^\alpha u(0) \tag{A.16}$$

for some natural number N and where c_α is a fixed (finite) set of constants.

Lemma A.2. *For each compactly supported $T \in \mathcal{S}'(\mathbb{R}^n)$ there exists $c \geq 0$ and a non-negative integer m such that*

$$|T(u)| \leq c||u||_{l,0} = c\, sup_{x \in \mathbb{R}^n} (1 + |x|)^m |u(x)| \quad \forall u \in \mathcal{S}'. \tag{A.17}$$

Proof. For each $w \in \mathcal{D}_n$ with $w\restriction_{\text{supp } T} \equiv 1$ we have

$$|T(u)| = |T(wu)| \leq c||wu||_{l,m} \tag{A.18}$$

for some constant $c \geq 0$ and non-negative integers l, m. Now for each multi-index α (corresponding to $\mathbb{R}^n$) with $|\alpha| \leq l$ we have

$$\begin{aligned} ||(1 + |x|)^m D^\alpha (wu)||_\infty &\leq || \sum_\beta \left[(1 + |x|)^m D^\beta w D^{\beta^c} u \right] ||_\infty \\ &\leq || \sum_\beta \left[(1 + |x|)^m D^\beta w \right] ||_\infty \sup_{|\gamma| \leq l} ||D^\gamma u||_\infty \end{aligned}$$

where the sum over β runs over decompositions of α into multi-indices β and β^c. The estimate is possible because $(1 + |x|)^m D^\beta w$ are compactly supported functions. ∎

Prop. A.1. (according to [BLOT90, Proof of Prop. 2.2, p. 52]) By Lemma A.2 we have that there exists $c \geq 0$ and a non-negative integer l such that

$$|T(u)| \leq c||u||_{l,0}, \quad \forall\, u \in \mathcal{S}. \tag{A.19}$$

Choosing $w \in \mathcal{D}_n$ with $w \equiv 1$ in an open region around the origin and a function $v \in \mathcal{S}$ with $D^\alpha v(0)$ vanishing for all α with $|\alpha| \leq l$ the norm $||w(k\cdot)v||_{l,0}$ goes to zero of $k \to \infty$. Thus $|T(v)| = |T(w(k\cdot)v)| \leq c||w(k \cdot v)||_{l,0}$ has to vanish. Choosing

$$v(x) \equiv u(x) - w(x) \sum_{|\alpha| \leq l} \frac{x^\alpha}{\alpha!} D^\alpha u(0) \tag{A.20}$$

gives the desired expression with $c_\alpha = \frac{1}{\alpha!} T(x^\alpha w)$. ∎

Proposition A.3. *(Tempered distribution concentrated at a point times a closed subset of $\mathbb{R}^n$) Let $T \in \mathcal{S}'(\mathbb{R}^m \times \mathbb{R}^n)$ with supp $T \subset \{a\} \times S$ for some $a \in \mathbb{R}^m, S \subset \mathbb{R}^n$ closed subset. Then T can be represented as*

$$T(x, y) = \sum_{|\alpha| \leq N} D_x^\alpha \delta(x) h_\alpha(y), \tag{A.21}$$

where $h_\alpha \in \mathcal{S}'(\mathbb{R}^m)$ with supp $h_\alpha \subset S$.

Proof. Let $u \in \mathcal{S}_m$ and $v \in \mathcal{S}_n$. Then supp $T(\cdot, v) \subset \{a\}$ and thus for each $v \in \mathcal{S}_n$ there exists non-negative integers N_v and constants $c_{\alpha,v}$ such that

$$T(u, v) = T(\cdot, v)(u) = \sum_{|\alpha| \leq N_v} c_{\alpha,v} D^\alpha u(a). \tag{A.22}$$

By the linearity and temperedness of $T(u, \cdot)$ for each $u \in \mathcal{S}_m$ we furthermore have that $h_\alpha(v) \equiv c_{\alpha,v}$ define tempered distributions on $\mathbb{R}^n$ and that there is a maximal non-negative integer N majorizing the numbers N_v. ∎

For simplicity we will only prove the gluing principle for finitely many patches, i.e., a finite family of tempered distributions glued together. The statement is the following:

Proposition A.4. *(The gluing principle for tempered distributions)*
Let $\{\mathcal{O}_j\}_{j=1,\ldots,m}$ finite open covering of $\mathbb{R}^n$ such that

$$Q_j = \mathbb{R}^n \setminus \bigcup_{i \neq j} \mathcal{O}_i \tag{A.23}$$

closed and contained in $\mathcal{O}_j$ and that for each $x \in Q_j$ the distance $d(x, \partial\mathcal{O}_j)$ to the boundary $\partial\mathcal{O}_j$ of $\mathcal{O}_j$ satisfies

$$d(x, \mathcal{O}_j) \geq A(1 + |x|)^{-\delta} \tag{A.24}$$

for fixed[17] *numbers $A > 0$, $\delta \geq 0$. Then for any family $\{T_j\}_{j=1,\ldots,m}$ of tempered distributions on $\mathbb{R}^n$ satisfying*

$$(T_i - T_j)\restriction_{\mathcal{O}_i \cap \mathcal{O}_j} = 0 \quad \forall i, j = 1, \ldots, m \tag{A.25}$$

there exists a unique tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ coinciding with T_j in $\mathcal{O}_j$ for all j.

[17]only dependent on the given covering

Proof.
Step 1: Construct a partition of unity subordinated to $\{\mathcal{O}_j\}_{j=1,\dots,m}$.

To begin with, assume $\delta = 0$ such that $d(x, \partial\mathcal{O}_j) \geq A > 0 \; \forall x \in Q_j$. Then for each j take $\tilde{e}_j \in \mathcal{D}(\mathbb{R}^n)$ such that

$$\tilde{e}_{j\restriction Q_j} \equiv 1 \quad \text{and} \quad \tilde{e}_{j\restriction Q_j^{\frac{1}{2}}} \equiv 0, \tag{A.26}$$

where $Q_j^{\frac{A}{2}}$ is the set of all points $x \in \mathbb{R}^n$ that have $d(x, Q_j) \leq \frac{A}{2}$. When we now define

$$e_j \equiv \frac{1}{\sum_{k=1}^m \tilde{e}_k} \tilde{e}_j \tag{A.27}$$

we obtain a finite partition of unity subordinated to the covering $\{\mathcal{O}_j\}$ as supp $e_j \subset Q_j^{\frac{A}{2}} \subset \mathcal{O}_j$ and $\sum_{j=1,\dots,m} e_j(x) \equiv 1$.

For $\delta \geq 0$ define $B_k \equiv \{x \in \mathbb{R}^n : k-1 \leq |x| \leq k\}$ for each natural number k. Then $A_{kj} \equiv B_k \cap \mathcal{O}_j$ gives rise to a locally finite countable covering of $\mathbb{R}^n$ where for each $x \in Q_{kj}$ (defined as above for $\{A_{kj}\}$)

$$d(x, \partial\mathcal{O}_j) \geq A(1+|x|)^{-\delta} \tag{A.28}$$

and therefore

$$d(x, \partial A_{kj}) \geq \inf_{x \in A_{kj}} (A(1+|x|)^{-\delta} \tag{A.29}$$

$$\geq Ak^{-\delta}. \tag{A.30}$$

This means that for each A_{kj} one can apply the ($\delta = 0$)-case (only for countably many e_{kj}) and then define

$$e_j = \sum_{k \in \mathbb{N}} e_{jk} \tag{A.31}$$

to obtain a finite partition of unity $\{e_{kj}\}$ subordinated to the covering $\{\mathcal{O}_j\}$.

Step 2: define T.

Let $u \in \mathcal{S}$ be an arbitrary test function then

$$T(u) \equiv \sum_{i=1}^m e_i T_i(u) \tag{A.32}$$

defines a generalized function[18] with

$$T_{\restriction \mathcal{O}_j} = T_{j\restriction \mathcal{O}_j} \tag{A.33}$$

as

$$\begin{aligned} e_i T_{i\restriction \mathcal{O}_j} &= e_i T_{i\restriction \mathcal{O}_i \cap \mathcal{O}_j} \\ &= e_i T_{j\restriction \mathcal{O}_i \cap \mathcal{O}_j}. \end{aligned}$$

[18]Note here that the e_j's clearly are multiplicators on $\mathcal{S}$.

and therefore

$$\sum_i e_i T_{i\upharpoonright\mathcal{O}_j} = \left(\sum_i e_{i\upharpoonright\mathcal{O}_i\cap\mathcal{O}_j}\right) T_{j\upharpoonright\mathcal{O}_j} = T_{j\upharpoonright\mathcal{O}_j}. \tag{A.34}$$

Moreover, T is uniquely specified by (A.33) as $\{\mathcal{O}_j\}$ covers $\mathbb{R}^n$. ■

Here are two lemmas that will be used for the following two propositions:

Lemma A.5. *Suppose a sequence $(T_n)_n \subset \mathcal{S}'$ such that for $\lim_{n\to\infty} T_n(u)$ exists for each $u \in \mathcal{S}$ (pointwise convergence). Then the linear map T defined by*

$$u \mapsto T(u) \equiv \lim_{n\to\infty} T_n(u) \tag{A.35}$$

is continuous (uniform convergence). In addition, there exist a constant c and natural numbers k, l such that $T_n(u) \leq c||u||_{k,l} \quad \forall n \in \mathbb{N}, u \in \mathcal{S}$.

Proof. This is a corollary from the uniform boundedness principle for Frechét spaces, i.e., for separated complete locally convex spaces (see e.g. [BLOT90, Theorem 1.7 and Corollary 1.9]). ■

Lemma A.6. *Suppose a sequence $(u_n)_n \subset \mathcal{S}$ converging to $u = \lim_{n\to\infty} u_n$ in $\mathcal{S}$ and a sequence $(T_n)_n \subset \mathcal{S}'$ converging to $T = \lim_{n\to\infty} T_n$ in $\mathcal{S}'$. Then $T_n(u_n)$ converges to $T(u)$.*

Proof. We have $||T(u) - T_n(u_n)|| \leq ||(T - T_n)(u)|| + ||T_n(u - u_n)||$. The first summand clearly converges to 0 (uniformly in u) and for the second summand we have

$$||T_n(u - u_n)|| \leq c||u - u_n||_{k,l} \quad \forall n \in \mathbb{N} \tag{A.36}$$

by Lemma A.5 from above. Important is here that c, k, and l are independent of n. Thus also the second summand converges to 0. ■

Proposition A.7. *Suppose a Hilbert space $\mathcal{H}$ and a linear map $A : \mathcal{S} \to \mathcal{H}$. Then the following properties are equivalent:*

(a) $(\phi, A(\cdot))$ is continuous for each $\phi \in D$,

(b) $(\phi, A(\cdot))$ is continuous for each $\phi \in \mathcal{H}$, i.e., A is weakly continuous

(c) A is norm continuous

Proof. (c) ⇒ (a),(b): Clear by the separate continuity of the Hilbert product.
(a) ⇒ (b): For each vector $\psi \in \mathcal{H}$ there exists a sequence $(\psi_n)_n \subset D$ converging against ψ in $\mathcal{H}$. Then $(\psi_n, A(\cdot))$ convergences pointwise, i.e., for each $u \in \mathcal{S}$, against $(\psi, A(\cdot))$ and by Lemma A.5 also in $\mathcal{S}'$.
(b) ⇒ (c): Similarly, for a sequence $(u_n)_n$ converging against u in $\mathcal{S}$ we have that $(A(u_n), A(\cdot))$ defines a sequence in $\mathcal{S}'$ converging against $(A(u), A(\cdot))$ pointwise and thus in $\mathcal{S}'$. Applying the other lemma, Lemma A.6, we obtain that $||A(u_n)||^2 = (A(u_n), A(u_n))$ converges to $||A(u)||^2$ and thus A is norm continuous. ■

Proposition A.8. *Suppose a Hilbert space $\mathcal{H}$, a dense linear subspace $D \subset \mathcal{H}$ and a linear map $A : \mathcal{S} \to \mathcal{L}(D, \mathcal{H})$. Then the following properties are equivalent:*

(a) $(\phi, A(\cdot)\psi)$ is continuous for each $\psi, \phi \in D$,

(b) $(\phi, A(\cdot)\psi)$ is continuous for each $\psi \in D$ and $\phi \in \mathcal{H}$, i.e., A is weakly continuous

(c) $A(\cdot)\psi$ is continuous for each $\psi \in D$, i.e., A is strongly continuous.

Proof. The result is an immediate corollary of Proposition A.7 by noting that $A(\cdot)\phi$ defines a vector-valued tempered distribution for each $\phi \in D$. ■

B d'Alembert equation as a Cauchy Problem

The goal of this section is to determine the solutions of the equation

$$\Box F(x) = \eta_{\mu\nu}\partial_\mu\partial_\nu F(x) = 0 \tag{B.1}$$

in the space of tempered distributions on $\mathbb{M}$. This equation is known as d'Alembert equation or may also be interpreted as massless Klein-Gordon equation. Other than in the main section we will understand this problem as a Cauchy problem instead of a division problem.

(In this section we have used slightly different notation than in the main document. For a tempered distribution T, $(T(x), u(x)) \equiv T(u)$. Moreover, the Fourier transform is denoted by $\tilde{T}$, $\underset{\sim}{T}$ or even $(\mathcal{F}_{p\to x}T)(x))$)

The first thing to note is that these solutions will be tempered distributions in $\vec{x}$ that are C^∞-dependent on x_0 (or vice versa):

Proof. Let $\tilde{F}$ denote the Fourier transformation of F in $\mathcal{S}'$. Then it satisfies

$$(p_0^2 - \vec{p}^2)\tilde{F}(p) = 0. \tag{B.2}$$

Therefore supp $\tilde{F} \subset \Gamma \subset \{|\vec{p}| \leq A(1 + |p_0|)^\delta\}$ for some positive constants A, δ. By [BLOT90, Ex. 2.32] this implies that $\tilde{F}$ is a convolute w.r.t. $\vec{p}$ and hence (see [BLOT90, Ex.2.48(b)]) F is a gen. fct. in $\vec{x}$ that is C^∞-dependent on x_0. ∎

As a result we obtain the Cauchy problem

$$\partial_t^2 F_t(\vec{x}) = \partial_{\vec{x}}^2 F_t(\vec{x}) \tag{B.3}$$

with initial values $u_0(\vec{x}) \equiv F_{t=0}(\vec{x})$, $u_1(\vec{x}) \equiv \partial_t F_{t=0}(\vec{x})$. This problem is uniquely solvable. Uniqueness can be shown in the following way:

Proof. Take IV to be 0 (i.e., $u_i(\vec{x}) \equiv 0$, $i = 1, 2$). Then by induction on m also all derivatives $\partial_t^m F_{t=0}(\vec{x})$ vanish identically (apply (B.3) and commute differentials). Now if we can proof analyticity in t we are done. Therefore take $\tilde{u} \in \mathcal{D}(\mathbb{R}^3)$ and observe that $(\tilde{F}, \tilde{u}(-\vec{p}))$ as a generalized function in p_0 has compact support (see above in the proof of C^∞-dependence) That implies that the Fourier transform

$$(F(x_0, \vec{x}), u(\vec{x})) \equiv (\mathcal{F}_{p_0\to x_0}\tilde{F}(x_0, \vec{p}), u(-\vec{p})) \tag{B.4}$$

is analytic. ∎

Existence can be shown in the following way: First show the existence of the fundamental solution $D(x)$ to the Cauchy problem (namely the solution with initial data $u_0(\vec{x}) = 0, u_1(\vec{x}) = \delta(\vec{x})$). Then show the existence of arbitrary solutions.

First:

Proof. define the tempered distribution

$$D(x) = \frac{1}{2\pi}\epsilon(x_0)\delta(x^2) \tag{B.5}$$

or equally well, its Fourier transform with respect to $\vec{p}$

$$\underset{\sim}{D}(t,\vec{p}) = \frac{\sin(|\vec{p}|t)}{|\vec{p}|}. \tag{B.6}$$

This gen. function fulfills

$$\begin{aligned} D(0,\vec{x}) &= 0 \\ \partial_t D(0,\vec{x}) &= \delta(\vec{x}). \end{aligned}$$

This can be seen from the following facts: $D(x)$ is antisymmetric in x_0 therefore D vanishes at $x_0 = 0$. For $\partial_t D = \delta$ regard

$$\partial_t \underset{\sim}{D}(0,\vec{p}) = \cos(0) = 1, \tag{B.7}$$

thus transforming it back gives δ. And lastly we see that $\Box D(x) = 0$ as

$$(\partial_t^2 + \vec{p}^2)\underset{\sim}{D}(t,\vec{p}) = (-|\vec{p}|^2 + \vec{p}^2)\underset{\sim}{D}(t,\vec{p}) = 0. \tag{B.8}$$

■

Here we will list some properties of the fundamental solution (apart from its characterising ones as a fund. solution):

$$\begin{aligned} \text{odd} \qquad & D(x) = -D(-x) \\ \text{locality} \qquad & \operatorname{supp} D \subset V \equiv \{x \in \mathbb{M} | x^2 > 0\} \\ \text{Lorentz-inv.} \qquad & D(\Lambda x) = D(x)\, \forall \Lambda \in L_+^\uparrow \end{aligned}$$

Second:

Proof. For arb. initial data $u_i(\vec{x}) \in \mathcal{S}'(\mathbb{R}^3)$

$$\begin{aligned} F(x) &= \partial_t D(t,\vec{x}) *_{\vec{x}} u_0(\vec{x}) + D(t,\vec{x}) *_{\vec{x}} u_1(\vec{x}) \\ &= D(x) * [\partial_{x_0}\delta(x_0)u_0(\vec{x}) + \delta(x_0)u_1(\vec{x})] \end{aligned}$$

is the unique solution to the Cauchy problem. $F(x)$ is well-defined as supp $D \subset V$. Therefore D and $\partial_t D$ are convolutes w.r.t. $\vec{x}$ (again by [BLOT90, Ex 2.32]) and the convolutions in the first line are well-defined. [19] Then

$$\begin{aligned} F(0,\vec{x}) &= \partial_t D\ (0,\vec{x}) *_{\vec{x}} u_0(\vec{x}) + D(0,\vec{x}) *_{\vec{x}} u_1(\vec{x}) \\ &= \delta(\vec{x}) *_{\vec{x}} u_0(\vec{x}) \\ &= u_0(\vec{x}) \end{aligned}$$

[19] For the second line note that the support of the right term is concentrated at $x_0 = 0$. Then [BLOT90, Ex. 2.41] applies such that the convolution is canonically defined

as well as

$$\partial_t F\,(0, \vec{x}) = u_1(\vec{x})$$

and most importantly

$$\begin{aligned}\Box F(x) &= \Box D\,(x) * [...]\\ &= 0.\end{aligned}$$

For the last equality we used that for arb. $u, v \in \mathcal{S}'$ such that $u*v$ is canonically defined[20] then for arbitray polynomial $P(\partial)$

$$P(\partial)(f * g) = P(\partial)f * g = f * P(\partial)g.$$

This follows from [BLOT90, Ex. 2.39(a)] and the linearity of the convolution. ■

Now we have the following theorem

Theorem B.1. *For an arbitrary $F \in \mathcal{S}'(\mathbb{M})$ which satisfies $\Box F = 0$ we have*

$$F(x) = D(x) * [\partial_{x_0}\delta(x_0)u_0(\vec{x}) + \delta(x_0)u_1(\vec{x})] \tag{B.9}$$

where $u_0 = F|_{t=0}$ and $u_1 = \partial_t F|_{t=0}$.

C Seperating distributions of advanced and retarded type

[This section is basically a detailed proof of [RS75, Problem 56, Section IX].]
In this subsection we want to address the task to decompose a given tempered distribution with support in the closed light cone $\bar{V} \subset \mathbb{M}$ into two tempered distributions $T_\pm$ with support in the closed upper/lower light cone $\bar{V}^\pm$, respectively. Such distributions are sometimes called to be of advanced/retarded type. It is clear that such a decomposition

$$T = T^{(+)} + T^{(-)} \tag{C.1}$$

can only be unique up to shifts by arbitrary tempered distributions concentrated at $x = 0$. Therefore our strategy is to begin with separating the part of T that is concentrated at the origin, naming it T_0. It will turn out that this is actually necessary for the decomposition as otherwise we cannot construct a sufficiently regular multiplicator for T in order to define $T_\pm$. This sufficient regularity will in fact not be smoothness, but only finite differentiability. To make this work let us show the following lemma

Lemma C.1. *Let T be a tempered distribution of order at most N.*[21] *Then there exists a linear and continuous extension $\tilde{T}$ of T to rapidly decreasing functions that are in C^N and for every function $\chi \in C^N(\mathbb{R}^n)$ that is of polynomial growth the map*

$$T(\chi\cdot) : \mathcal{S}(\mathbb{R}^n) \ni u \mapsto \tilde{T}(\chi u) \tag{C.2}$$

defines a well-defined tempered distribution. Moreover, supp $T(\chi\cdot) \subset$ supp χ.

[20]This is equivalent to $u(x)v(y-x)$ being of integrable type w.r.t. x. If u or v is a convolute this automatically implies that $u(x)v(y-x)$ is of integrable type.

[21]A distribution is said to be of order at most $N \in \mathbb{N}_0$ when for some constant $c \geq 0$ and $M \in \mathbb{N}_0$ it satisfies the inequality $|T(u)| \leq c||u||_{M,N} \quad \forall u \in \mathcal{S}(\mathbb{R}^n)$.

Proof. (Of the Lemma) By the Schwartz representation theorem there exist continuous functions c_α of polynomial growth such that

$$T = \sum_{|\alpha| \leq N} c_\alpha D^\alpha \tag{C.3}$$

It is clear that the above representation (rhs) extends also to rapidly decreasing functions that are in C^N and that this defines a linear and continuous extension of T, denoted above by $\tilde{T}$.
That $T(\chi\cdot)$ is tempered can be seen by the following calculation

$$\begin{aligned} T(\chi u) \quad &= \sum_{|\alpha|\leq N} \int c_\alpha(x) D^\alpha (\chi u)(x) d^n x \\ &= \sum_{|\beta|,|\gamma|\leq N} \int c_{\alpha=\beta+\gamma}(x)(-1)^{|\beta|} D^\beta \chi(x) D^\gamma u(x) d^n x \\ &= \sum_{|\gamma|\leq N} \int \tilde{c}_\gamma(x) D^\gamma u(x) d^n x \end{aligned}$$

where $\tilde{c}_\gamma = \sum_{|\beta|\leq N}(-1)^{(|\beta|)} c_{\beta\gamma}$ are continuous functions of polynomial growth, too.
That the support of $T(\chi\cdot)$ is contained in the support of χ is evident. Just take any Schwartz function u supported outside the support of χ, then $\chi u \equiv 0$ and $T(\chi u) = T(0) = 0$. ■

Proof. (Of the decomposition) One natural way to project distributions onto distributions with certain support properties is to multiply them by a function that has the desired support properties, and which acts as a multiplicator on the test function space (meaning that the product still lies in the test function space:
Thus let f be a smooth function on the unit sphere $S \subset \mathbb{R}^n$ mapping to $[0,1]$ with $S \cap V^+ \equiv 1$ and $S \cap V^- \equiv 0$. Then $\chi(x) \equiv f(x/||x||)$ is smooth except for the origin.[22] At the origin there is a singularity of order 1 meaning that $(x^2)^n\chi$ is in C^{2n-1}.[23] This is the key observation for the decomposition. We cannot construct a smooth function separating $V^\pm$, but we can make it finitely differentiable up to any order by multiplication with $(x^2)^n$.
There exist[24] natural numbers (incl. 0) M, N and $c \geq 0$ such that

$$|T(u)| \leq c||u||_{M,N} \quad \forall u \in \mathcal{S}(\mathbb{R}^n). \tag{C.4}$$

Then the next step is to observe that

$$S_\pm(u) = T((x^2)^{M+1}\chi_\pm u) \tag{C.5}$$

where $\chi_+ = \chi$ and $\chi_- = 1 - \chi$ are well-defined tempered distributions supported in the closed upper/lower light cone $\bar{V}^+$. This assertion is given by the lemma above. Together

[22] Here the norm $||x||$ denotes the (or let us say one of the) Euclidean norms on $\mathbb{R}^n$.
[23] For this statement we simply have to check that the partial derivatives (at all orders up to $2n-1$) at the origin yield continuous functions. These derivatives yield a polynomial $P(x)$ with $P(x) \to 0$ for $x \to 0$. Therefore we have $|P(x)\chi(x)| \leq |P(x)||\chi(x)| \leq |P(x)| \to 0$ for $x \to 0$. These statements are the same for any chosen norm on $\mathbb{R}^n$.
[24] valid for any tempered distribution

they give the want decomposition for the 'regularized' version of T: $(x^2)^{M+1}T = S_+ + S_-$. Hence, what we would like to have is to write down $T(u) = (x^2)^{M+1}T((x^2)^{-M-1}u)$ which is only well-defined for u that go sufficiently fast to zero towards the origin. defining the linear continuous map $H : \mathcal{S} \to \mathcal{S}$ by

$$u \mapsto Hu = \frac{1}{(x^2)^{M+1}} \left[u - h \sum_{|\beta| \leq 2M+1} \frac{1}{\beta!} D^\beta u(0)\ x^\beta \right] \tag{C.6}$$

where $h \in \mathcal{D}(\mathbb{R}^n)$ with $h \equiv 1$ near the origin, we achieve that for arbitrary $u \in \mathcal{S}(\mathbb{R}^n)$

$$T(u) = (x^2)^{M+1}T(Hu) + T_0(u) \tag{C.7}$$

where $T_0(u) = \sum_{|\beta| \leq 2M+1} T(hx^\beta) \frac{1}{\beta!} D^\beta u(0)$ such that T_0 is concentrated at the origin. The left term is then decomposable into the above defined $S_\pm$ such that we obtain

$$T(u) = S_+(Hu) + S_-(Hu) + T_0(u). \tag{C.8}$$

Splitting up T_0 in some way among S_+ and S_- will then give us a possible decomposition into $T^{(+)} + T^{(-)}$ as we wanted to achieve. ■

D Proof of Stone's theorem

Theorem D.1 (Stone's Theorem). *Let U be a map $\mathbb{R} \to \mathcal{U}(\mathcal{H})$. Then the following two statements are equivalent:*

(a) U is a continuous group homomorphism

(b) There exists a self-adjoint operator A on $\mathcal{H}$ such that $U(t) = e^{itA}$ for each $t \in \mathbb{R}$.

Lemma D.2. *Let A be a self-adjoint operator on $\mathcal{H}$. Then the unitary group $U(t) \equiv e^{itA}$ satisfies*

(a) for each $\phi \in D(A)$ the strong derivative is $\lim_{t\to 0} \frac{U(t)-1}{t}\phi = iA$.

(b) for each strongly differentiable vector $\phi \in \mathcal{H}$ it follows $\phi \in D(A)$.

Proof. (a): Let $\phi \in D(A)$ be arbitrary, then

$$\begin{aligned} \frac{e^{itA}-1}{t}\phi &= I\left(\frac{e^{it\lambda}-1}{t}\right)\phi \\ &= \lim_{n\to\infty} I\left(\frac{e^{it\lambda}-1}{t}\chi_{[-n,n]}(\lambda)\right)\phi \\ &\to \lim_{n\to\infty} I\left(i\lambda\chi_{[-n,n]}(\lambda)\right) \text{ as } t \to 0. \end{aligned}$$

In the last line we interchanged the limits $n \to \infty$ and $t \to 0$, which is possible as each of the single limits clearly exists pointwise and as one of the limits exists also uniformly. The existence of the single limits is implied by the norm continuity of I for bounded functions, see Proposition 2.32(e), although we only need strong continuity here:

$$\lim_{n\to\infty} I\left(\frac{e^{ih\lambda}-1}{h}\chi_{[-n,n]}(\lambda)\right)\phi = I\left(\frac{e^{ih\lambda}-1}{h}\right)\phi\ \forall h \in \mathbb{R}, \tag{D.1}$$

and
$$\lim_{h\to 0} I\left(\frac{e^{ih\lambda}-1}{h}\chi_{[-n,n]}(\lambda)\right)\phi = I\left(i\lambda\chi_{[-n,n]}(\lambda)\right)\phi \;\forall n\in\mathbb{N}. \tag{D.2}$$

As $([-n,n])_n$ defines a bounding sequence[25] of measurable sets for the identity function $f(\lambda)=\lambda$, the limit exists uniformly if and only if $\phi\in D(I(f))$. In other words $\int \lambda^2 \langle\phi, E_A(\lambda)\phi\rangle = ||A\phi||^2$ has to be finite, which is implied by $\phi\in D(A)$.
(b): On a strongly differentiable vector $\phi\in\mathcal{H}$ the limit $\lim\limits_{t\to 0}\frac{U(t)-1}{t}\phi$ exists. In case $\phi\in D(A)$ this limit coincides with the operator iA. Thus $B\equiv \lim\limits_{t\to 0} i\frac{U(t)-1}{t}$ (with domain of existence of the limit) fulfils $B\supset A$ and B is clearly symmetric. Hence $A\subset B\subset B^\dagger\subset A^\dagger = A \Rightarrow A=B$. ■

of Stone's Theorem. (b) ⇒ (a): Immediate as a consequence of the functional calculus for self-adjoint operators. In particular we have $U(s+t)=e^{i(s+t)A}=e^{isA}e^{itA}$ and continuity according to Proposition 2.32(f).
(a) ⇒ (b): As seen in the Lemma D.2, if (a) ⇒ (b) holds, the self-adjoint operator to be found must coincide with $-i$ times the strong derivative with respect to U. In the proof of Lemma D.2(b) this operator was defined as B. Thus we need to show that B generates the unitary group. Let us define $V(s)\equiv e^{isB}$.
For $\phi\in D(B)$ a short computation shows that
$$U'(t)\phi = iBU(t)\phi \text{ and } V'(t)\phi = iBV(t)\phi. \tag{D.3}$$

One has to note here that $U(t)\phi\in D(B)$ as $\lim\limits_{s\to 0}\frac{U(s)-1}{s}U(t)\phi = U(t)\lim\limits_{s\to 0}\frac{U(s)-1}{s}\phi = U(t)U'(0)\phi$ and $V(t)\phi\in D(B)$ likewise. Lastly let us define $w(t)\equiv U(t)\phi - V(t)\phi$ and hence $w'(t)=iBw(t)$. This implies that $\frac{d}{dt}||w(t)||^2=\langle iBw(t),w(t)\rangle+\langle w(t),iBw(t)\rangle = 0$. With $w(0)=0$ we see that actually $w(t)=0\;\forall t$ and thus $U(t)=V(t)\;\forall t$. ■

E Alternative proof of the covariant decomposition of the two-point function

Lemma E.1. *The covariance of the two-point-function $W_{\mu\nu}$ and the spectral condition together imply the form*
$$W_{\mu\nu}=\eta_{\mu\nu}K+\partial_\mu\partial_\nu G, \tag{E.1}$$
with K and G being scalar Lorentz-invariant generalized functions. For a given $W_{\mu\nu}$ the decomposition gives rise to K and G defined up to a simultaneous shift by $K\mapsto K-2a, G\mapsto G+az^2+b$.

Proof. The covariance of $W_{\mu\nu}(x)$ can be analytically continued to the covariance $W_{\mu\nu}(z)$ under the proper complex Lorentz-group $L_+(\mathbb{C})$.[26] The analyticity domain may be chosen to be the extended past tube $\mathcal{T}^-_{\text{ext}}$.[27] Using these complex Wightman distributions is

[25] A bounding sequence of a function f is a sequence of sets $(M_n)_n$ such that for each M_n the function $f\chi_{M_n}$ is bounded, $M_n\subset M_{n+1}$ and $E(\cup_n M_n)=1$, where E is the spectral measure with respect to which the spectral integral is defined.

[26] given by Bargmann-Hall-Wightman-Theorem, see e.g. [BLOT90, Chapter 9.B Theorem 9.1, p. 362].

[27] The analyticity domain can be chosen even wider to be the extended permuted tube, but this is not necessary here.

advantageous for the following reason: Analytic distributions can shown to be necessarily regular, i.e., they are analytic functions (with the natural action on test functions)[28]. This means that for the complex Wightman distribution one does not have to deal with singularites. The classification of representations of $L_+(\mathbb{C})$ in terms of such analytic functions[29] gives us the form

$$W_{\mu\nu}(z) = \eta_{\mu\nu} W_1(z) + z_\mu z_\nu W_2(z). \tag{E.2}$$

with W_i being Lorentz-invariant distributions. The Lorentz-invariance of the W_i can be directly seen by the covariance of $W_{\mu\nu}$ and the invariance of $\eta_{\mu\nu}$:

$$\begin{aligned} \eta_{\mu\nu} W_1(\Lambda z) + (\Lambda z)_\mu (\Lambda z)_\nu W_2(\Lambda z) &= W_{\mu\nu}(\Lambda z) \\ &\overset{\text{cov. of W}}{=} \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma W_{\rho\sigma}(z) \\ &= (\Lambda g \Lambda^T)_{\mu\nu} W_1(z) + (\Lambda z)_\mu (\Lambda z)_\nu W_2(z) \\ &\overset{\text{inv. of g}}{=} \eta_{\mu\nu} W_1(z) + (\Lambda z)_\mu (\Lambda z)_\nu W_2(z). \end{aligned} \tag{E.3}$$

and therefore $W_i(\Lambda z) = W_i(z)$.[30]

The analyticity of the l.h.s. of (E.2) (in the extended past tube) gives the same analyticity on the r.h.s. Taking the off-diagonal terms gives then the analyticity of W_2 and taking the trace gives the analyticity of W_1. Within their analyticity domains W_1 and W_2 are Lorentz-invariant functions (remember that analytic distributions are regular as stated above). These may be written as functions of invariants. The only invariant that can be constructed from z_μ is z^2.[31] Therefore we can write

$$W_i(z) = T_i(z^2), \quad i = 1, 2 \tag{E.4}$$

within their analyticity domains.

Now, for an arbitrary generalized functions $G(z^2)$ we have the equation

$$\partial_\mu \partial_\nu G(z^2) = 2\eta_{\mu\nu} \frac{d}{dz^2} G(z^2) + 4 z_\mu z_\nu \left(\frac{d}{dz^2} \right)^2 G(z^2). \tag{E.5}$$

defining $G(z^2)$ as the solution of the equation

$$\left(\frac{d}{dz^2} \right)^2 G(z^2) = \frac{1}{4} T_2(z^2) \tag{E.6}$$

we establish then the connection (see Eq. (E.5)) between the terms $z_\mu z_\nu T_2(z^2)$ and $\partial_\mu \partial_\nu G(z^2)$ via a shift proportional to $\eta_{\mu\nu}$ that we can put into T_1. The existence of

[28] See e.g. [BLOT90, Chapter 5.E p. 208 Prop. 5.13(a)]

[29] see Appendix C, based on original work from Hepp [Hep63] and comments of Araki within

[30] That the one can induce the separate invariance of the W_i can be seen in the following way: Taking the off-diagonal terms of eq. (E.3) only the second term survives. One can divide by $(\Lambda z)_\mu (\Lambda z)_\nu$ ($z \neq 0$ because zero is not contained in the extended past tube, thus this term is not zero) and gets the invariance of W_2, hence of W_1, too.

[31] For the more general setting of many variables see [BLOT90, Chapter 5.F pp. 209].

such a generalized function $F(z^2)$ is guaranteed by giving the explicit solution of Eq. (E.6)

$$F(z^2) = \frac{1}{4}\int_{const.}^{z^2}\int_{const.}^{s^2} T_2(t^2)dt^2ds^2, \tag{E.7}$$

where the constants are choosen inside the analyticity domain of $T_2(z^2)$.

Thus, finally, we can express Eq. (E.2) in the desired way

$$W_{\mu\nu}(z) = \eta_{\mu\nu}W_1(z) + \partial_\mu\partial_\nu W_2(z), \tag{E.8}$$

where $W_1(z) = T_1(z^2) - 2\eta_{\mu\nu}\frac{d}{dz^2}F(z^2)$ and $W_2(z) = F(z^2)$.
Taking the boundary value of this eq. then gives the proof of eq. (E.1).

To see the uniqueness of W_1 and of W_2 up to linear shifts we take two pairs K, G and K', G' satisfying eq. (E.8) (i.e., the complex one). Taking the off-diagonal terms we obtain

$$\partial_\mu\partial_\nu G = \partial_\mu\partial_\nu G' \quad \forall\mu \neq \nu. \tag{E.9}$$

Let us take fixed indices μ, ν. Then integrating this eq. two times we get that the possible shift terms relating G and G' can be only of the following form:

$$G(z) = G'(z) + \hat{\mu} + \hat{\nu}). \tag{E.10}$$

Here we introduced the short-hand notation where $\hat{\mu}$ means not dependent on the component z_μ. Now running through the μ with $\mu \neq \nu$ we get that $G(z) = G'(z) + \nu + \hat{\nu}$ meaning that the terms are either only ν-dependent or not ν-dependent. Running through all the ν we get that $G(z) = G'(z) + h_0(z_0) + h_1(z_1) + h_2(z_2) + h_3(z_3)$ where h_i are generalized functions which are only dependent on z_i. As we required Lorentz-invariance for G and G' this has to be true for $\sum_i h_i$ as well. Thus $\sum_i h_i$ has to be a function which is only z^2-dependent. As a consequence $\sum_i h_i(z) = az^2 + b$.
As a result we obtain that $G = G' + az^2 + b$ for some constants a, b and that $K = K' - \frac{1}{4}\Box(az^2 + b) = K' - 2a$. ∎

F Scalar Wightman two-point function

Let ϕ be a free massless scalar Wightman Quantum Field in a Wightman Theory with $W(\xi)$ being its two-pt. fct. in translation-invariant form

Theorem F.1.

$$W(\xi) = a + bD_0^{(+)}(\xi) \tag{F.1}$$

with $a, b \in \bar{\mathbb{R}}_+$ *and* $D_m^{(+)}(\xi) = \int \theta(p^0)\delta(p^2 - m^2)e^{ip\xi}d_4p$ *being the positive frequency Pauli-Jordan commutation function of mass* m.

Proof. By positivity we know that the Fourier transform $\hat{W}(p)$ is a non-negative gen. fct. and therefore a non-negative measure of power growth. By the spectral condition it has support in the upper light cone $\bar{V}^+$. As we are considering a scalar field it is furthermore

Lorentz-invariant. A Lorentz-invariant, non-negative measure with support in $\bar{V}^+$ has the general form

$$d\mu(p) = \left(\int_0^\infty \theta(p^0)\delta(\tau - p^2)d\rho(\tau) + a(2\pi)^4\delta(p)\right) d_4p \tag{F.2}$$

where $a \geq 0$ and $\rho(\tau)$ is a monotone decreasing function of power growth[32]. Thus we have

$$d\hat{W}(p) = \left[a(2\pi)^4\delta(p) + \int_0^\infty 2\pi\theta(p^0)\delta(p^2 - m^2)d\sigma(m^2)\right] dp \tag{F.3}$$

with $\sigma(\lambda)$ a monotone decreasing function of power growth. As ϕ is a free massless scalar field we know that $\Box\phi = 0$ and therefore (for every $p \in \mathbb{M}$) $p^2 d\hat{W}(p) = 0$ or in other words $p^2 d\hat{W}(p)$ is the zero-measure. Therefore if we take the upper light cone and an arbitrary measurable set $X \subset \mathbb{R}$ and $\bar{V}^+_X \equiv \bar{V}^+ \cap \{p^2 \in X\}$ then

$$0 = (p^2 d\hat{W}(p))(\bar{V}^+_X) \propto (\lambda d\sigma(\lambda))(X). \tag{F.4}$$

In this notation the functions $p^2 d\hat{W}(p)$ and $\lambda d\sigma(\lambda)$ are understood as measures evaluated on the measurable sets $\bar{V}^+_X$ and X. The proportionality follows from the eq. (F.3) together with the fact that $p^2\delta(p)$ vanishes everywhere and that $d\sigma(X) = \int_X d\sigma(m^2) = \int_{[0,\infty]}\int_X \delta(p^2 - m^2)dp^2 d\sigma(m^2)$.
Hence we know that $\lambda d\sigma(\lambda) = 0$ meaning that $d\sigma(\lambda) = 0$ for $\lambda > 0$ such that

$$d\sigma(\lambda) = b\delta(\lambda) \tag{F.5}$$

with $b \geq 0$. Thus we arrive at

$$\hat{W}(p) = a(2\pi)^4\delta(p) + b2\pi\theta(p^0)\delta(p^2) \tag{F.6}$$

or in coordinate space redefining the constants a and b at

$$W(\xi) = a + bD_0^{(+)}(\xi). \tag{F.7}$$

■

For the commutator we arrive at (omitting the 0-index for D)

$$\langle[\phi(x), \phi(y)]\rangle = W(\xi) - W(-\xi) = b(D^{(+)}(\xi) - D^{(+)}(-\xi)) = bD(\xi). \tag{F.8}$$

It should be noted that the constants a and b can also be absorbed into redefinitions of ϕ. In order to do that proceed as follows (the trivial generalized function $1(f) = \int f(x)dx$ is omitted):

$$\text{define:} \quad \phi'(f) \equiv \phi(f) - \langle\phi(f)\rangle \tag{F.9}$$

$$\text{define:} \quad \phi''(f) \equiv \sqrt{b}\phi'(f) + \sqrt{a} \tag{F.10}$$

Both redefinitions do not alter the Wightman conditions. The first redefinition ensures that $\langle\phi\rangle = 0$, the second redefinition absorbs the constants a and b. Thus we end up with a possible choice of ϕ to yield

$$W(\xi) = D^{(+)}(\xi). \tag{F.11}$$

It should be noted that the absorption of a as in eq. (F.9) is usually done in physics. Whereas b is usually fixed by experimental results.

[32]'of power growth' means that the associated measure gives $\int_{\tau<r} d\rho(\tau)$ being bounded from above by a power of r.

G Relation between Wightman and commutator function

Let us take a scalar field ϕ and let us denote its two-point function by W and the corresponding (vev of the) commutator function by D. If we have been given the Wightman two-point functions it is easy to obtain the commutator functions. We just have

$$D(\xi) \equiv \langle [\phi(x), \phi(y)] \rangle = W(\xi) - W(-\xi). \tag{G.1}$$

Therefore we know that the commutator function is local (vanishes for spacelike vectors), satisfies the generalized spectral condition (support of its Fourier transform lies in the light cone) and transforms covariantly as the two-point function does.

Here we want to take a look at the inverse problem. Let $D_{\mu\nu}(\xi)$ be a fixed (vacuum expectation value) of the commutator of A with the relation from eq. (G.1). What can we say about W if W is an arbitrary generalized Lorentz-invariant function which satisfies the spectral condition.

Let us decompose W into symmetric and antisymmetric parts $W_\pm$, respectively, such that $W = W_+ + W_-$. It is easy to show that this decomposition exists, is unique and that also $W_\pm$ have to be Lorentz-invariant and satisfy the spectral condition.[33] and therefore

$$D(\xi) = W(\xi) - W(-\xi) = (W_+(\xi) - W_+(-\xi)) + (W_-(\xi) - W_-(-\xi)) = 2W_-(\xi). \tag{G.2}$$

Thus we know that $W = \frac{1}{2}D + W_+$ with W_+ satisfying Lorentz-invariance and spectral condition. If we would also know that W is of positive type, this would imply that W_+ is of positive type, too.

Let us now decompose $\hat{D}$ and $\hat{X} \equiv \hat{W}_+$ into tempered distributions that have support in $\bar{V}^\pm$. For $\hat{D}$ we can just define $\hat{D}^{(\pm)} = \theta(\pm p_0)\hat{D}$ as $\hat{D}$ is supported outside the origin. For $\hat{X}$ it is more complicated, but also possible. Such a decomposition is achieved for an arbitrary tempered distribution with support in the closed light cone in the next section of the Appendix, Subsection C. This condition is satisfied for $\hat{X} = \hat{W} - \frac{1}{2}\hat{D}$. Here $\hat{X}^{(\pm)}$ shall denote a specific choice of such a decomposition which is unique up to distributions concentrated at 0. Let us therefore introduce $\backsimeq$ to denote equalities up to distributions concentrated at 0. The symmetry of $\hat{X}$ implies $\hat{X}^{(\pm)}(-p) = \hat{X}^{(\mp)}(p)$. By the spectral condition we know that the negative energy parts of W should give zero, hence

$$\frac{1}{2}\hat{D}^{(-)}(p) + \tilde{X}^{(-)}(p) \backsimeq 0 \tag{G.3}$$

and therefore

$$\tilde{X}^{(+)}(p) = \tilde{X}^{(-)}(-p) \backsimeq -\frac{1}{2}\hat{D}^{(-)}(-p) = \frac{1}{2}\hat{D}^{(+)}(p). \tag{G.4}$$

Finally, we obtain

[33] define $W_\pm(\xi) \equiv \frac{1}{2}(W(\xi) \pm W(-\xi))$ and note that the upper light cone is closed under addition.

$$\hat{W}(p) \backsimeq \tilde{D}^{(+)}(p). \tag{G.5}$$

As $\hat{W}$ and $\hat{D}^{(+)}$ are Lorentz-invariant we arrive at

$$\hat{W}(p) = \hat{D}^{(+)}(p) + P(\Box)\delta(p). \tag{G.6}$$

or

$$W(x) = D^{(+)}(x) + P(x^2). \tag{G.7}$$

Bibliography

[AI89] T. Ya. Azizov and I. S. Iokhvidov. *Linear Operators in Spaces with an Indefinite Metric*. John Wiley & Sons, 1989.

[Ber71] J. Bertrand. *Poincaré Covariance and Quantization of Zero-Mass Fields*. Il Nuovo Cimento A, 1(1):1–12, jan 1971.

[BLOT90] N. N. Bogolubov, A. A. Logunov, A. I. Oksak, and I. T. Todorov, editors. *General Principles of Quantum Field Theory*. Springer Netherlands, 1990.

[Bra72] L. Bracci. *Vector and Tensor Representations of the Poincaré Group*. Il Nuovo Cimento A, 8(1):129–146, mar 1972.

[Buc75] D. Buchholz. *Collision Theory for Massless Fermions*. Communications in Mathematical Physics, 42(3):269–279, oct 1975.

[Buc77] D. Buchholz. *Collision Theory for Massless Bosons*. Communications in Mathematical Physics, 52(2):147–173, jun 1977.

[Buc86] D. Buchholz. *Gauss' Law and the Infraparticle Problem*. Physics Letters B, 174(3):331–334, jul 1986.

[Dyb18] W. Dybalski. *Lectures on Mathematical Foundations of QFT*. Lecture Notes, `www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/WojciechDybalski/Notes-QFT41.pdf`, 2018.

[FPS74] R. Ferrari, L. E. Picasso, and F. Strocchi. *Some Remarks on Local Operators in Quantum Electrodynamics*. Communications in Mathematical Physics, 35(1):25–38, mar 1974.

[Hö58] L. Hörmander. *On the Division of Distributions by Polynomials*. Arkiv för Matematik, 3(6):555–568, dec 1958.

[Hep63] K. Hepp. *Lorentz-kovariante analytische Funktionen*. PhD thesis, ETH Zürich, 1963.

[JM17] B. Schroer J. Mund, K.-H. Rehren. *Helicity Decoupling in the Massless Limit of Massive Tensor Fields*. Nuclear Physics B, 924:699–727, nov 2017.

[NO90] N Nakanishi and I Ojima. *Covariant Operator Formalism of Gauge Theories and Quantum Gravity*. World Scientific, nov 1990.

[RS75] M. Reed and B. Simon. *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness*. Elsevier Science, 1975.

[RS80] M. Reed and B. Simon. *Methods of Modern Mathematical Physics: Functional analysis*. Academic Press, 1980.

[Sch61] Schweber. *An Introduction to Relativtistic Quantum Field Theory*. Row, Peterson, 1961.

[Sch89] G. Scharf. *Finite QED*. Springer-Verlag Berlin Heidelberg, 1989.

[Sch12] K. Schmüdgen. *Unbounded Self-adjoint Operators on Hilbert Space*. Springer Netherlands, 2012.

[Ste00] Othmar Steinmann. *Perturbative Quantum Electrodynamics and Axiomatic Field Theory*. Springer Berlin Heidelberg, 2000.

[Str67] F. Strocchi. *Gauge Problem in Quantum Field Theory*. Physical Review, 162(5):1429–1438, oct 1967.

[Str68] F. Strocchi. *Gauge Problem in Quantum Field Theory. II. Difficulties of Combining Einstein Equations and Wightman Theory*. Physical Review, 166(5):1302–1307, feb 1968.

[Str70] F. Strocchi. *Gauge Problem in Quantum Field Theory. III. Quantization of Maxwell Equations and Weak Local Commutativity*. Physical Review D, 2(10):2334–2340, nov 1970.

[Str78] F. Strocchi. *Local and Covariant Gauge Quantum Field Theories. Cluster Property, Superselection Rules, and the Infrared Problem*. Physical Review D, 17(8):2010–2021, apr 1978.

[Str90] D. Strube. *Collision Theory for Massless Particles in the Framework of a Wightman Field Theory*. Journal of Mathematical Physics, 31(9):2244–2252, sep 1990.

[Str13] F. Strocchi. *An Introduction to Non-Perturbative Foundations of Quantum Field Theory*. Oxford University Press, feb 2013.

[SW80] R. F. Streater and A. S. Wightman. *PCT, Spin and Statistics, and All That*. Princeton University Press, jan 1980.

[WG64] A. S. Wightman and L. Garding. *Fields as Operator-Valued Distributions in Relativistic Quantum Theory*. Arkiv för Fysik, 1964.

[Wig56] A. S. Wightman. *Quantum Field Theory in Terms of Vacuum Expectation Values*. Physical Review, 101(2):860–866, jan 1956.

[Zhe73] D. P. Zhelobenko. *Compact Lie Groups and Their Representations*. Translations of Mathematical Monographs. American Mathematical Society, 1973.

Statement of Authorship

I hereby declare that I am the sole author of this master thesis and that I have identified work of other people as such. Where this is more specific than generic statements or common knowledge there are references to the bibliography. I further declare that I have not submitted this thesis at any other institution in order to obtain a degree.

Münster, 21st of May, 2019 J. Mandysch