Existence of Minimizers for Causal Variational Principles on Non-Compact Manifolds

Existenz von Minimierern kausaler Variationsprinzipien auf nichtkompakten Mannigfaltigkeiten

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Mathematik
der Universität Regensburg

vorgelegt von
Christoph Matthias Langer
aus Neumarkt i.d.OPf.
im Jahr 2021

Promotionsgesuch eingereicht am 18. Juni 2021.

Die Arbeit wurde angeleitet von Prof. Dr. Felix Finster.

Prüfungsausschuss:	Vorsitzender:	Prof. Dr. Walter Gubler
	1. Gutachter:	Prof. Dr. Felix Finster
	2. Gutachter:	Prof. Dr. Franz Gmeineder (Universität Konstanz)
	weiterer Prüfer:	Prof. Dr. Harald Garcke
	Ersatzprüfer:	Prof. Dr. Georg Dolzmann

Abstract

The theory of causal fermion systems is a new approach to describe theoretical physics by means of a variational principle, the so-called *causal action principle*. It aims to provide a unified description of quantum mechanics and general relativity. Following the principle of least action, minimizers of the causal action principle in the class of regular Borel measures are considered as relevant physical objects whose support is interpreted as physical spacetime. In order to analyze the existence theory of the causal action principle in sufficient generality, it is useful to consider more general *causal variational principles*. The variational principle is then to minimize integral expressions of an underlying non-negative Lagrangian on topological spaces in a suitable class of measures. The aim of this thesis is to analyze the existence of minimizers of causal variational principles on non-compact topological spaces in more detail. It is shown that, under suitable assumptions on the Lagrangian, the existence of minimizers in the class of regular Borel measures on non-compact second-countable locally compact Hausdorff spaces can be proven. Afterwards, this existence result is extended to causal variational principles on non-locally compact Polish subspaces of infinite-dimensional separable Banach spaces. Finally, for minimizers in the class of so-called *strictly negative definite measures in the strong sense*, the Euler-Lagrange equations in momentum space in the *homogeneous setting* are derived.

Zusammenfassung

Die Theorie kausaler Fermionsysteme ist ein neuartiger Zugang zur Beschreibung theoretischer Physik mithilfe eines Variationsprinzips, das als *kausales Wirkungsprinzip* bezeichnet wird. Die Theorie ermöglicht eine mathematisch einheitliche Formulierung relativistischer Quantentheorie und Allgemeiner Relativitätstheorie. Dem Prinzip der kleinsten Wirkung folgend werden Minimierer des kausalen Wirkungsprinzips in der Klasse der regulären Borel-Maße als physikalisch relevante Objekte aufgefasst, deren Träger als physikalische Raumzeit interpretiert wird. Zur Entwicklung und Analyse der Existenztheorie von Minimierern des kausalen Variationsprinzips in hinreichender Allgemeinheit erfolgt die Betrachtung allgemeinerer *kausaler Variationprinzipien*, bei welchen ein Wirkungsintegral über eine zugrundeliegende positive Lagrange-Funktion auf topologischen Räumen in einer geeigneten Klasse von Maßen minimiert wird. Das Ziel der vorliegenden Arbeit besteht in einer eingehenden Untersuchung der Existenz von Minimierern kausaler Variationsprinzipien auf nichtkompakten topologischen Räumen. Unter geeigneten Annahmen an die zugrundeliegende Lagrange-Funktion wird die Existenz von Minimierern kausaler Variationsprinzipien in der Klasse der regulären Borel-Maße auf nichtkompakten zweitabzählbaren lokalkompakten topologischen Hausdorff-Räumen bewiesen. Im Anschluss daran wird dieses Existenzresultat auf kausale Variationsprinzipien auf nichtlokalkompakten polnischen Unterräumen unendlich-dimensionaler separabler Banachräume übertragen. Im *homogenen Fall* werden schließlich die Euler-Lagrange-Gleichungen im Impulsraum für Minimierer in der Klasse sogenannter *strikt negativ definiter Maße im starken Sinne* hergeleitet.

Contents

CHAPTER 1

Introduction

The theory of causal fermion systems is a new approach to describe theoretical physics by means of a variational principle. It aims to provide a unified description of quantum theory and general relativity. In the physical theory of causal fermion systems, spacetime and the structures therein are described by a minimizer of the so-called causal action principle (for an introduction to the physical background and the mathematical context, we refer the interested reader to Chapter 2, the textbook [**59**], the survey articles [**62, 64**] as well as the web platform [**1**]). For clarity, let us first recall the definition of a causal fermion system according to [**59**].

DEFINITION 1.0.1. *Let $\mathcal{H}$ be a separable complex Hilbert space together with a scalar product $\langle .|. \rangle_{\mathcal{H}}$, and denote by $\mathrm{L}(\mathcal{H})$ the set of bounded linear operators on $\mathcal{H}$. Moreover, given a parameter $n \in \mathbb{N}$ (the so-called* **"spin dimension"***), we let $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ be the set of all self-adjoint operators on $\mathcal{H}$ of finite rank, which (counting multiplicities) have at most n positive and at most n negative eigenvalues. Furthermore, let $d\rho$ be a positive measure on $\mathcal{F}$ (defined on a σ-algebra of subsets of $\mathcal{F}$), the so-called* **universal measure**. *We refer to $(\mathcal{H}, \mathcal{F}, d\rho)$ as a* **causal fermion system**.

Given a causal fermion system $(\mathcal{H}, \mathcal{F}, d\rho)$ together with a non-negative function

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+ := [0, \infty)$$

(the *Lagrangian*), the causal action principle is to minimize the *causal action* $\mathcal{S}$ defined as the double integral over the Lagrangian

$$\mathcal{S}(\rho) = \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x, y)$$

under variations of the measure $d\rho$ within the class of regular Borel measures on $\mathcal{F}$ under suitable side conditions. *Causal variational principles* evolved as a mathematical generalization of the causal action principle [**56, 65**] in order to study the existence of minimizers in sufficient generality. The aim of this thesis to work out the existence theory of minimizers for causal variational principles under suitable assumptions on the Lagrangian $\mathcal{L}$ in the case that the underlying topological space $\mathcal{F}$ is *non-compact*.

In order to put the thesis into the mathematical context, let us review the results concerning the existence theory of minimizers for causal variational principles so far. To begin with, in [**53**] it was proposed to formulate physics by minimizing a new type of variational principle in discrete spacetime. The suggestion in [**53**, Section 3.5] led to the causal action principle in discrete spacetime, which was first analyzed mathematically in [**54**]. A more general and systematic inquiry of causal variational principles on measure spaces was carried out in [**56**]. In particular, variational principles in the so-called *homogeneous setting* were considered in [**56**, Section 4] in order to deal with an infinite number of particles in an infinite spacetime volume. The main advantage of the homogeneous setting is that it allows for Fourier methods, thus giving rise to a natural

correspondence between position and momentum space. As a consequence, one is led to minimize the causal action by varying in the class of so-called *negative definite measures*, and the existence of minimizers on bounded subsets of momentum space is proven in [**56**, Theorem 4.2]. The structure of minimizers was analyzed in more detail in [**12, 68**]. Following the basic concept of causal fermion systems, the resulting Euler-Lagrange (EL) equations are of particular physical interest. Apart from the physical consequences of the EL equations outlined in [**53**], a mathematical analysis of the EL equations was carried out in [**12**]. Moreover, assuming the existence of a minimizing measure, in [**65**] the corresponding EL equations are worked out in the case that the underlying topological space $\mathcal{F}$ is a compact manifold. The EL equations in the homogeneous setting, on the other hand, have not yet been worked out.

The aim of this thesis is to extend the previous existence results to the non-compact setting. More specifically, the existence theory of minimizers for causal variational principles shall be developed for a non-compact topological space $\mathcal{F}$. The first step towards this goal is to infer the existence of minimizers on compact subsets of second-countable, locally compact topological Hausdorff spaces in the class of normalized Borel measures. By exploiting σ-compactness of such topological spaces, the existence theory of minimizers is then extended to the whole second-countable, locally compact topological Hausdorff space under appropriate assumptions on the Lagrangian (the so-called "locally compact setting"). Since every finite-dimensional manifold is a second-countable, locally compact Hausdorff space, the obtained existence result applies in particular in the case of an underlying finite-dimensional non-compact manifold. Next, under suitable assumptions on the Lagrangian, the existence results obtained in the locally compact setting are applied in order to deduce the existence of minimizers in the case that $\mathcal{F}$ is a non-locally compact Polish subspace of an infinite-dimensional separable Banach space (the so-called "non-locally compact setting"). In particular, the existence result derived in the non-locally compact setting applies to the regular Banach manifold considered in [**67**]. In both the locally compact and the non-locally compact setting, the corresponding EL equations will be outlined. The existence theory of minimizers for causal variational principles in the homogeneous setting, on the other hand, is not accomplished in this thesis. Instead, the structure of the resulting EL equations in momentum space is worked out under appropriate assumptions. A possible strategy to develop the existence theory of minimizers on whole momentum space in the class of negative definite measures is to essentially make use of the obtained EL equations by applying suitable scaling arguments. A detailed analysis of the existence theory in the homogeneous setting under physically reasonable assumptions shall be postponed to future publications.

The thesis is organized as follows. In Chapter 2, we give a short introduction to the theory of causal fermion systems. After putting the theory of causal fermion systems into the historical context (Section 2.2), we recall fundamental physical preliminaries (Section 2.3). Next, underlying ideas concerning the principle of the fermionic projector are outlined (Section 2.4). Afterwards, we present the main structures of the theory of causal fermion systems in a concise way (Section 2.5). The reader not interested in the physical background of the theory of causal fermion systems may skip this chapter.

In Chapter 3, we analyze causal variational principles in the locally compact setting. We first give a short introduction (Section 3.1). Next, we recall the physical motivation as well as the mathematical preliminaries (Section 3.2). Afterwards, we outline causal variational principles in the σ-locally compact setting (Section 3.3). We then prove the existence of minimizers for Lagrangians of compact range (Section 3.4). Afterwards, the

results are extended to Lagrangians decaying in entropy (Section 3.5). In the appendix it is shown that the obtained minimizers are non-zero (Appendix A.1). More precisely, the goal of Chapter 3 is to prove the existence of minimizers of causal variational principles on second-countable, locally compact Hausdorff spaces. Furthermore, the corresponding Euler-Lagrange equations are derived. The method is to first prove the existence of minimizers of the causal variational principle restricted to compact subsets for a lower semi-continuous Lagrangian. Exhausting the underlying topological space by compact subsets and rescaling the corresponding minimizers, we obtain a sequence of measures which converges vaguely to a regular Borel measure of possibly infinite total volume. It is shown that, for continuous Lagrangians of compact range (see Definition 3.3.3), this regular Borel measure solves the Euler-Lagrange equations (Theorem 3.4.2). Moreover, we prove that the constructed Borel measure is a minimizer under variations of compact support (Theorem 3.4.9). Under additional assumptions, it is proven that this measure is a minimizer under variations of finite volume (Theorem 3.4.10). We finally extend our results to continuous Lagrangians decaying in entropy (see Definition 3.5.1). The results of this chapter were published in the article [**66**].

In Chapter 4, the existence results obtained in [**66**] are extended to causal variational principles in the non-locally compact setting. We first give a short introduction (Section 4.1) and recall the mathematical preliminaries (Section 4.2). Afterwards, causal variational principles in the non-locally compact setting are introduced (Section 4.3). After constructing a global measure (Section 4.4), the existence of minimizers for Lagrangians of bounded range is proven (Section 4.5). The results are then extended to Lagrangians vanishing in entropy (Section 4.6). Finally, we analyze topological properties of spacetime (Section 4.7). In the appendix, we first study topological properties of causal fermion systems (Appendix A.2); afterwards, the support of locally finite measures on Polish spaces is analyzed (Appendix A.3). From the conceptual point of view, we provide a method for constructing (possibly trivial) measures on non-locally compact Polish subspaces of infinite-dimensional separable Banach spaces which, under suitable assumptions, are minimizers of causal variational principles in the non-locally compact setting (see Definition 4.3.4). Furthermore, for non-trivial minimizers the corresponding Euler-Lagrange equations are derived. The method is to exhaust the underlying Banach space by finite-dimensional subspaces and to prove existence of minimizers of the causal variational principle restricted to these finite-dimensional subsets of the Polish space under suitable assumptions on the Lagrangian. This gives rise to a corresponding sequence of minimizers. Restricting the resulting sequence to countably many compact subsets of the Polish space, by considering the resulting diagonal sequence we are able to construct a regular measure on the Borel algebra over the whole topological space (Theorem 4.4.3). For continuous Lagrangians of bounded range (see Definition 4.3.7) it can be shown that, under suitable assumptions, the obtained measure is a (possibly non-trivial) minimizer under variations of compact support. Under additional assumptions, we prove that the constructed measure is a minimizer under variations of finite volume (Theorem 4.5.6) and solves the corresponding Euler-Lagrange equations (Theorem 4.5.7). Afterwards, we extend our results to continuous Lagrangians vanishing in entropy (see Definition 4.6.1). Finally, assuming that the obtained measure is locally finite, topological properties of spacetime are worked out and a connection to dimension theory is established (Theorem 4.7.3). The results were published in the paper [**110**].

In Chapter 5, we study causal variational principles in the homogeneous setting in more detail. To this end, we first give a short introduction (Section 5.1) and recall

variational principles in infinite spacetime volume (Section 5.2). Next, causal variational principles in the homogeneous setting are introduced (Section 5.3). Afterwards, we derive estimates for minimizing sequences of causal variational principles in the homogeneous setting (Section 5.4). Unfortunately, although these estimates do imply boundedness of minimizing sequences (with respect to the total variation), the lack of uniform tightness does not allow for employing Prohorov's theorem. This is the motivation for proceeding in the spirit of [**65**] and [**66**] by first deriving the Euler-Lagrange equations. Apart from being of physical interest, the hope is to employ the Euler-Lagrange equations in order to establish the existence theory afterwards. For this reason, we first analyze the causal action in the homogeneous setting (Section 5.5). Next, we consider auxiliary variational principles (Section 5.6), which allows us to derive the Euler-Lagrange equations in momentum space under appropriate assumptions (Section 5.7). Finally, we put our results into the physical context of causal fermion systems (Section 5.8). In the appendix we give a justification of the side conditions under consideration (Appendix A.4). More specifically, in Chapter 5 we derive Euler-Lagrange equations for causal variational principles in the homogeneous setting for minimizers in the class of strictly negative definite measures in the strong sense (see Definition 5.7.3). Our method is to proceed in several steps: Under the assumption that the Lagrangian is continuously differentiable (see Assumption 5.5.4), the initial step is to compute the first variation of the causal action. Afterwards, we rewrite the first variation of the causal action in terms of momentum space (Lemma 5.5.13). Analyzing the resulting expressions by means of auxiliary variational principles, we then derive a convenient representation of the first variation of the causal action in terms of Lagrange multipliers (Theorem 5.6.8). By restricting attention to appropriate variations of minimizers in the class of strictly negative definite measures in the strong sense, we are finally in the position to deduce the desired Euler-Lagrange equations in momentum space (Theorem 5.7.10).

In Appendix A, we compile some supplementary results to the previous chapters. In Appendix A.1 we show that the measure constructed in Chapter 3 is non-zero. In Appendix A.2 we derive some topological properties of causal fermion systems, and in Appendix A.3 we analyze the support of locally finite measures on Polish spaces. Moreover, in Appendix A.4 we give a justification of the side conditions under consideration in Chapter 5.

In Appendix B we collect selected mathematical definitions and results required in this thesis. For the sake of convenience, we compile some basic definitions and results concerning topology (Appendix B.1), measure theory (Appendix B.2), functional analysis (Appendix B.3) and differential geometry (Appendix B.4).

CHAPTER 2

Introduction to Causal Fermion Systems

ABSTRACT. We give a short introduction to the theory of causal fermion systems. After putting the theory of causal fermion systems into the historical context, we recall fundamental physical preliminaries. Afterwards, we enter the underlying ideas of the principle of the fermionic projector and clarify its connection to causal fermion systems. We finally outline the main structures of the theory of causal fermion systems.

2.1. Introduction

The theory of causal fermion systems is a new approach to describe theoretical physics by means of a variational principle. It aims to provide a unified description of quantum mechanics and general relativity. Based on Dirac's original concept of describing the vacuum in terms of a completely filled "Dirac sea," the theory of causal fermion systems evolved from the "principle of the fermionic projector," an approach which originated by the attempt to resolve some of the shortcomings of relativistic quantum field theory. The variational principle at the heart of the theory of causal fermion systems is known as "causal action principle." Spacetime and the structures therein are described by minimizers of the causal action principle. The underlying picture is to think of "macroscopic" spacetime as a certain limiting case (the so-called "continuum limit") of a more fundamental "microscopic" structure which is described mathematically in terms of a so-called "regularization." In the continuum limit (that is, when the regularization is removed), the resulting Euler-Lagrange equations give rise to an effective interaction which corresponds to classical gravity as well as the strong and electroweak gauge fields of the Standard Model.

Let us first recall some essential facts concerning modern physics. At present, there are four known fundamental forces in nature, which are *gravitation, electromagnetic interaction, weak interaction* and *strong interaction*. These four fundamental forces are described by *Einstein's theory of general relativity* (gravitation) and the *Standard Model of elementary particle physics* (electromagnetic, strong and weak interaction). All elementary particles known at present are either *fermions* or *bosons*. Fermions are the subatomic constituents of matter, comprising *quarks* and *leptons*. The fundamental bosons, on the other hand, are the mediating particles of the various interactions: *photons* (electromagnetism), $W^{\pm}$ *and* Z *bosons* (weak interaction) and *gluons* (strong interaction). Furthermore, there might exist not yet detected *gravitons* (gravitation). The Standard Model successfully incorporates all known properties of strong, weak and electromagnetic forces, whereas gravitational interactions are absent in the Standard Model. Thus it is considered one of the great physical problems of this century to unify general relativity and quantum mechanics [**73, 94, 170**].

In order to motivate the underlying ideas of the theory of causal fermion systems, let us point out the following drawback of the Standard Model of elementary particle

physics. The description of the Standard Model rests upon relativistic quantum field theory. Unfortunately, one of the serious complications in quantum field theory is that the theory is naively divergent in the ultraviolet region. More precisely, it turns out that many physical quantities of interest are not integrable in momentum space due to unbounded contributions for large momenta, that is, for high energies. In order to bypass these divergences, physicists developed the so-called *renormalization program*, which is a set of prescriptions for how to make sense of divergent integrals [**94, 129**]. To this end, one commonly introduces an "ultraviolet cutoff" in momentum space, thereby "cutting off" the region of high energies, a procedure known as "regularization." Despite its considerable success, this method of "regularization" seems artificial from the physical point of view, and it is unjustified from the mathematical one. Or, to put it in Dirac's own words (cf. [**103**, p. 184]),

> "I must say that I am very dissatisfied with the situation because the so-called 'good theory' does involve neglecting infinities in its equations, ignoring them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves disregarding a quantity when it is small — not neglecting it just because it is infinitely great and you do not want it!"

The need for a cutoff in momentum space indicates that large momenta are not taken into account in the correct way in quantum field theory. More specifically, the cutoff in momentum space, introduced in order to render divergent integrals finite, is often associated to Planck energy $E_P \approx 1.22 \cdot 10^{28}$ eV.[1] In view of Heisenberg's uncertainty principle, physicists infer a correspondence between large momenta (high energies) and small distances in position space. Because of that, in quantum field theory one disregards distances which are smaller than the Planck length $\ell_P \approx 1.6 \cdot 10^{-35}$ m. Consequently, the microstructure of spacetime is completely unknown. Unfortunately, at present there is no consensus on what the correct mathematical model for "Planck scale physics" should be.

The simplest and maybe most natural approach is to assume that on the Planck scale, spacetime is no longer a continuum but becomes in some way "discrete." This is the starting point in the monograph [**53**], which later evolved to the theory of causal fermion systems [**59**]. The objective of this chapter is to outline the main structure of the theory of causal fermion systems and to clarify its relation to physics. To begin with, let us first state the formal definition of causal fermion systems according to [**59**].

DEFINITION 2.1.1. *Let $\mathcal{H}$ be a separable complex Hilbert space together with a scalar product $\langle .|. \rangle_{\mathcal{H}}$, and denote by $\mathrm{L}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$. Moreover, given a parameter $n \in \mathbb{N}$ (the so-called* **"spin dimension"***), we let $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ be the set of all self-adjoint operators on $\mathcal{H}$ of finite rank, which (counting multiplicities) have at most n positive and at most n negative eigenvalues. Furthermore, let $d\rho$ be a positive measure on $\mathcal{F}$ (defined on a σ-algebra of subsets of $\mathcal{F}$), the so-called* **universal measure**. *We refer to $(\mathcal{H}, \mathcal{F}, d\rho)$ as a* **causal fermion system**.

In order to single out the physically admissible causal fermion systems, one must formulate physical equations. Following the principle of least action, we impose that the universal measure should be a minimizer of the causal action principle, which we now introduce. For any $x, y \in \mathcal{F}$, the product xy is an operator of rank at most $2n$. Denoting

[1] For clarity, one *electron volt* is approximately $1 \text{ eV} \approx 1.60 \cdot 10^{-19}$ kg m^2/s^2.

its eigenvalues (counting algebraic multiplicities) by $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy} \in \mathbb{C}$ and introducing the *spectral weight* $|\,.\,|$ of an operator as the sum of the absolute values of its eigenvalues, the spectral weight of the operator products xy and $(xy)^2$ are given by

$$|xy| = \sum_{i=1}^{2n} \left|\lambda_i^{xy}\right| \qquad \text{and} \qquad \left|(xy)^2\right| = \sum_{i=1}^{2n} \left|\lambda_i^{xy}\right|^2 .$$

We introduce the *Lagrangian* by

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R} , \qquad \mathcal{L}(x,y) = \left|(xy)^2\right| - \frac{1}{2n} |xy|^2$$

and the *causal action* by

$$\mathcal{S}(\rho) = \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d\rho(x)\, d\rho(y) . \tag{2.1.1}$$

The *causal action principle* is to minimize $\mathcal{S}$ by varying the universal measure $d\rho$ under the following constraints:

$$\textit{volume constraint:} \qquad \rho(\mathcal{F}) = \text{const} \tag{2.1.2}$$

$$\textit{trace constraint:} \qquad \int_{\mathcal{F}} \mathrm{tr}(x)\, d\rho(x) = \text{const} \tag{2.1.3}$$

$$\textit{boundedness constraint:} \qquad \mathcal{T}(\rho) := \iint_{\mathcal{F}\times\mathcal{F}} |xy|^2\, d\rho(x)\, d\rho(y) \leq C , \tag{2.1.4}$$

where C is a given parameter (and tr denotes the trace of a linear operator on $\mathcal{H}$).

In order to specify the class of measures in which to vary $d\rho$, on $\mathcal{F}$ we consider the topology induced by the operator norm

$$\|A\| := \sup \left\{ \|Au\|_{\mathcal{H}} \text{ with } \|u\|_{\mathcal{H}} = 1 \right\} .$$

The *Borel σ-algebra* $\mathcal{B}(\mathcal{F})$ is generated by the open sets of $\mathcal{F}$. Its elements are called *Borel sets*. A *regular Borel measure* is a measure on the Borel sets with the property that it is continuous under approximations by compact sets from inside and by open sets from outside (for basics see [**76, 87**] or Appendix B.2). The right prescription is to vary the universal measure $d\rho$ within the class of regular Borel measures on $\mathcal{F}$. The constraints (2.1.2)–(2.1.4) are needed to avoid trivial minimizers and in order for the variational principle to be well-posed.

It is of central importance to observe that the causal action principle essentially depends on the eigenvalues of the operator product xy with $x, y \in \mathcal{F}$. Instead of working with the operator product xy, however, it is convenient to proceed as follows. For every $x \in \mathcal{F}$, the corresponding *spin space* $S_x := x(\mathcal{H})$ is a subspace of $\mathcal{H}$ of dimension at most $2n$. Denoting the orthogonal projection onto the subspace $x(\mathcal{H})$ by π_x, for all $x, y \in \mathcal{F}$ one defines the *kernel of the fermionic projector* $P(x,y)$ by

$$P(x,y) := \pi_x y|_{S_y} : S_y \to S_x . \tag{2.1.5}$$

In order to express the eigenvalues of the operator xy, one then introduces the *closed chain* A_{xy} by

$$A_{xy} := P(x,y)P(y,x) : S_x \to S_x . \tag{2.1.6}$$

Repeating the arguments after [**59**, eq. (1.1.10)], it turns out that the eigenvalues of the closed chain A_{xy} coincide with the non-trivial eigenvalues $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy}$ of xy.

At this stage, however, the reader who is unfamiliar with the theory of causal fermion systems might wonder how the causal action principle comes about and in which way it is connected to physics. This introduction is intended to hopefully settle these issues; more precisely, the goal of this chapter is to put the theory of causal fermion systems into the physical context, to give a terse summary of its main concepts and to clarify the specific form of the causal action principle. Following [**153**], we shall restrict attention to the crucial constructions in a non-rigorous way without entering most of the details.

The chapter is organized as follows. In order to put the theory of causal fermion systems into the historical context, in Section 2.2 we give a succinct overview of the history of physics, thereby summarizing particularly significant physical achievements. In Section 2.3 we recall the fundamental physical preliminaries required for the theory of causal fermion systems: classical mechanics (§2.3.1), electrodynamics (§2.3.2), special relativity (§2.3.3), general relativity (§2.3.4) as well as quantum mechanics (§2.3.5). After recalling some aspects of group theory in physics (§2.3.6), we give a brief introduction to elementary particles (§2.3.7) and outline the Standard Model of elementary particle physics (§2.3.8). We conclude this section with a glimpse of quantum field theory (§2.3.9). After these preparations, in Section 2.4 we outline the principle of the fermionic projector, from which the theory of causal fermion systems eventually evolved. To this end, we first sketch the derivation of local gauge freedom (§2.4.1) and then enter the principle of the fermionic projector (§2.4.2). Afterwards, the connection to causal fermion systems is established (§2.4.3). In Section 2.5, we finally introduce the theory of causal fermion systems. We first present the general strategy (§2.5.1) and clarify the connection to Dirac's original concept of a Dirac sea (§2.5.2). We then recall the external field problem (§2.5.3) and describe the construction of the "auxiliary" fermionic projector (§2.5.4). In order to explain the connection to modern physics, we outline the so-called "light-cone expansion" (§2.5.5) and describe the formalism of the continuum limit in a few words (§2.5.6). This allows us to formulate the Euler-Lagrange equations (§2.5.7) in a concise way, from which finally the classical field equations can be derived (§2.5.8) and the connection to the Standard Model of particle physics may be established (§2.5.9).

2.2. A Brief History of Physics

The notion of "physics" was coined by one of Aristotle's (384–322 BC) major works [**141**]. Back to the ancient Greeks, the idea of an atomic structure of matter was postulated by Democritus (460–371 BC), and the Athenian philosopher Plato (428/427–348/347 BC) developed a first approach towards cosmology [**141**]. At the beginning of modern ages, Nikolaus Kopernikus (1473–1543) described in his principal work *De revolutionibus orbium coelestium* (1543) a heliocentric view of the world [**18**], which was affirmed by astronomical observations by Galileo Galilei (1564–1641) and calculations of his contemporary, Johannes Kepler (1571–1630). The foundations of modern mechanics were developed and formulated in Isaac Newton's magnum opus *Philosophiae Naturalis Principia Mathematica* (1687) in order to explain Kepler's laws by a universal law of gravitation [**124**]. The principle of least action goes back to Pierre Louis Moreau de Maupertuis (1698–1759). Employing variational principles, Leonhard Euler (1707–1783) and Joseph-Louis Lagrange (1736–1813) made fundamental contributions to the further development of mechanics.

At the end of the 18th century, Charles-Augustin de Coulomb (1736–1806) formulated the laws of electrostatics. Observations concerning heat conduction by Joseph Fourier (1768–1830) led to the fundamental method of *Fourier analysis*, whereas the

foundations of electrodynamics were laid by André-Marie Ampère (1775–1836) and Michael Faraday (1791–1867). The further development of electrodynamics as well as thermodynamics in the 19th century is mainly due to James Prescott Joule (1818–1889), Lord Kelvin (1824–1907), Hermann von Helmholtz (1821–1894), James Clerk Maxwell (1831–1879) and Ludwig Boltzmann (1844–1906), reaching its peak in *Maxwell's equations*. In honor of William Rowan Hamilton (1805–1865), the principle of least action is also referred to as *Hamilton's principle*. Moreover, Oliver Heaviside (1850–1925) and Heinrich Hertz (1857–1894) made important contributions for developing the theory of electromagnetism [**18, 80, 81, 127**].

At the beginning of the 20th century, the two fundamental pillars of modern physics were discovered: quantum theory by Max Planck (1858–1947) as well as the theory of relativity by Albert Einstein (1879–1955). Both theories substantially changed the physical view of the world [**18**]. Preparatory works by Hendrik Antoon Lorentz (1853–1928) and Henri Poincaré (1854–1912) led to Einstein's special relativity (1905), a theory which is solely based on two postulates, the principle of relativity and constancy of the speed of light. Subsequently, Hermann Minkowski (1864–1909) developed the formal basis of four-dimensional spacetime in special relativity theory. The foundations of general relativity go back to Albert Einstein as well. Originating from the principle of equivalence, he applied differential geometry as developed by Carl Friedrich Gauss (1777–1855), Bernhard Riemann (1826–1866), Elwin Bruno Christoffel (1829–1900), Gregorio Ricci-Curbastro (1853–1925) and Tullio Levi-Civita (1873–1941) in order to formulate gravitation as a geometric property of curved four-dimensional spacetime [**149**]. The *Einstein field equations* form the fundamental equations of general relativity. David Hilbert (1862–1943) was the first one to recognize a particularly elegant derivation of the Einstein equations by varying the *Einstein-Hilbert action*, thereby employing the principle of least action. The *Schwarzschild metric*, discovered in 1916 by Karl Schwarzschild (1873–1916) was the first known exact solution of the Einstein equations [**144**].

Quantum theory owes its origins to Planck's quantum hypothesis [**130**], according to which radiation is not emitted continuously, but in form of discrete "quanta." Based on Planck's hypothesis, Einstein deduced the particle character of light (photons). In order to resolve the resulting contrast to the wave theory of light, the French physicist Louis de Broglie (1892–1987) postulated *wave-particle duality* to be the fundamental principle of nature. In the mid-1920's, Werner Heisenberg (1901–1976), Max Born (1882–1970), Pascual Jordan (1902–1980), Wolfgang Pauli (1900–1958) and Erwin Schrödinger (1887–1961) developed quantum mechanics [**156**]. In 1928, the British physicist Paul Dirac (1902–1984) formulated the *Dirac equation* [**32**], thereby postulating the existence of anti-matter. Together with the *Klein-Gordon equation* (due to Oskar Klein (1894–1977) and Walter Gordon (1893–1939)), the Dirac equation admits a relativistic formulation of quantum theory. The discovery of spin, a fundamental property of all particles, led to a subdivision of matter into *bosons*[2] and *fermions*[3]. Max Born introduced the statistical interpretation of wave functions, which was expanded by the *Copenhagen interpretation of quantum theory* by Niels Bohr (1885–1962) and Werner Heisenberg in 1928. By eliminating determinism, quantum physics drastically changed the paradigms of physics [**18**].

[2]Particles with integer spin, in honor of the Indian physicist Satyendra Nath Bose (1894–1974).

[3]Particles with half-integer spin, named after the Italian physicist Enrico Fermi (1901–1954).

The 1930's were characterized by developments in nuclear physics, supported by the invention of particle accelerators. For instance, the neutron was discovered, which in the sequel led to the discovery of weak interaction. The further development of particle accelerators after the Second World War led to the discovery of a whole "zoo" of elementary particles. At the end of the 1940's, quantum field theory (QFT) emerged [**145, 146, 156**]. The development of quantum electrodynamics (QED) is mostly due to Shinichiro Tomonaga (1906–1979), Julian Schwinger (1918–1994) as well as Richard Feynman (1918–1988) and Freeman Dyson (1923–2020). Originally introduced by Hermann Weyl (1885–1955), *gauge theories* turned out to be of central importance for the emerging *Standard Model of elementary particles*, in particular *Yang-Mills theories* [**164**] (due to Chen-Ning Yang (1922–) and Robert Mills (1927–1999)). Unifying the electromagnetic and weak interaction to the electroweak theory was accomplished by Abdus Salam (1926–1996), Sheldon Glashow (1932–) and Steven Weinberg (1933–2021) in the 1960's. The roots of quantum chromodynamics (QCD), a Yang-Mills theory for the strong interaction, go back to the 1970's. Important contributions concerning the strong interaction are due to Murray Gell-Mann (1929–2019). Since the end of the 1970's, elementary particle physics was governed by string theory [**165**], aiming for unifying gravity and quantum theory, which also is the main objective of other approaches like quantum gravity [**138**]. The origins of what later evolved to the theory of causal fermion systems go back to the 1990's [**48, 49, 51**].

2.3. Physical Preliminaries

The aim of this subsection is to outline fundamental physical preliminaries which are of essential importance for the theory of causal fermion systems. For an introduction to concepts of modern physics, the interested reader is referred to [**11**]. For a non-technical introduction to physics we refer to [**45, 46, 47**] as well as [**80**]. The explanations in this subsection mostly follow the non-rigorous style common in physics textbooks, and we usually shall employ the physicists' conventions. Concerning the physicists' notation, we highly recommend [**73**]. For instance, physicists like to indicate the dimensionality of their integrals explicitly by writing the volume element on $\mathbb{R}^n$ as $d^n x$ (see [**73**]). Moreover, vectors in Euclidean space $\mathbb{R}^3$ are usually written as $\mathbf{x}$ or $\vec{x}$. In order to symbolically distinguish between conventional physics and and the theory of causal fermion systems, we shall first employ the bold-face notation and switch to the vector notation afterwards.

Let us finally address the subject of physical units. The fundamental aspects of the universe to which all physical measurements relate are mass, length or distance and time (cf. [**73**]). All quantities in physics come with "dimensions" that can be expressed in terms of these three basic ones. Very often is is customary and convenient to use so-called *natural units* $\hbar = c = 1$ (where $\hbar$ is reduced Planck's constant and c denotes the speed of light). In natural units, we have $[\text{length}] = [\text{time}] = [\text{energy}]^{-1} = [\text{mass}]^{-1}$, and the mass of a particle is equivalent to its rest energy mc^2, and also to its inverse Compton wavelength $mc/\hbar$ (see [**96, 129**]). Factors of $\hbar$ or c, which are not displayed, can be reinserted by "getting the dimension right."

2.3.1. Classical Mechanics. All physical theories are based on certain fundamental laws, which are referred to as *laws of nature*. In classical mechanics, *Newton's axioms* are regarded as being these fundamental laws. Based on Newton's axioms, classical mechanics provides the foundations of theoretical physics, and its methods, including the Hamilton and Lagrangian formalism, can be regarded as fundamental principles of

theoretical physics. An alternative approach is based on the principle of least action, in which case the laws of nature in classical mechanics are formulated in terms of an action principle (cf. [**70**]). In preparation for the theory of causal fermion systems, let us enter the Lagrangian formalism in classical mechanics in some more detail. The subsequent explanations are due to [**73**] and [**106**].

The Equations of Motion. The main subject of classical mechanics is to study time-dependent motions in three-dimensional Euclidean space $\mathbb{R}^3$. One of the fundamental concepts of classical mechanics is that of a *particle*. The position of a particle in $\mathbb{R}^3$ is indicated by its vector $\mathbf{r} = \mathbf{r}(t)$, depending on time t, whose components are its Cartesian coordinates $x = x(t), y = y(t), z = z(t)$. The derivative $d\mathbf{r}/dt$ of $\mathbf{r}$ with respect to time t is called *velocity* of the particle, and the second derivative $d^2\mathbf{r}/dt^2$ is referred to as *acceleration*. It is customary to denote differentiation with respect to time by $\mathbf{v} = \dot{\mathbf{r}}$.

To define the position of a system of N particles in space, one needs to specify N radius vectors, thus giving rise to $3N$ coordinates. Accordingly, if no constraints are involved, the number of *degrees of freedom* of such a system is $3N$, which in general is defined as the number of independent quantities required in order to uniquely specify the position of any system. In general, any s quantities $q_1, \ldots, q_s$ which completely define the position of a system with s degrees of freedom are called *generalized coordinates* of the system, and the derivatives $\dot{q}_i$ are called its *generalized velocities*. The manifold described by generalized coordinates is known as *configuration space*. The relations between the coordinates, velocities and accelerations are referred to as *equations of motion*.

The *principle of least action* (or *Hamilton's principle*), according to which every mechanical system is characterized by a definite function $L(q_1, \ldots, q_s, \dot{q}_1, \ldots, \dot{q}_s, t)$, or briefly $L(q, \dot{q}, t)$, is the most general formulation of the law governing the motion of mechanical systems. Given a path $t \mapsto q(t)$ in configuration space with $t_1 \leq t \leq t_2$, the *action* is introduced by

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt\,,$$

and the problem is to *minimize the action over all paths $q(t)$ that begin at $q(t_1)$ and end at $q(t_2)$*. The function L is called the *Lagrangian* of the system concerned. The necessary condition for S to have an extremum is that its *first variation* equals zero. Thus the principle of the least action may be expressed by imposing that

$$\delta S = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt = 0\,.$$

We then obtain s equations of the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \qquad \text{for all } i = 1, \ldots, s\,. \tag{2.3.1}$$

These are the equations of motions, also known as *Euler-Lagrange equations*.[4]

Let us now consider a system of N particles which may interact with one another, but are not subject to external interactions. This is called a *closed system*. In this case,

[4]Introducing the *momentum p conjugate to q* by $p \equiv \partial L/\partial \dot{q}$, and the *Hamiltonian* by the Legendre transformation $H(p, q) = p\dot{q} - L(q, \dot{q})$, the equations of motion (2.3.1) become

$$\frac{\partial H}{\partial p} = \dot{q} \qquad \text{and} \qquad -\frac{\partial H}{\partial q} = \dot{p}\,.$$

the Lagrangian can be written as

$$L = \sum_{a=1}^{N} \frac{1}{2} m_a v_a^2 - U(\mathbf{r}_1, \ldots, \mathbf{r}_N) \,, \tag{2.3.2}$$

where $\mathbf{r}_a$ is the radius vector of the ath particle ($a = 1, \ldots, N$). This is the general form of the Lagrangian for a closed system in classical mechanics, where $T \equiv \sum_{a=1}^{N} \frac{1}{2} m_a v_a^2$ is called the *kinetic energy*, and U is referred to as the *potential energy*.

Conservation Laws. In order to describe mechanical phenomena mathematically one needs to choose a *reference frame*. In classical mechanics, it is sensible to assume that space is *homogeneous* as well as *isotropic* and time is *homogeneous* (where homogeneity means that no point in space and time is distinguished, and isotropy means that no spatial direction is distinguished, see e.g. [**70**] and [**45**, Chapter 52]).

We now consider a mechanical system with s degrees of freedom whose dynamics respects the symmetries of space and time; more precisely, we assume its Lagrangian to be invariant under translations in space and time as well as rotations in space. To such symmetries correspond the following conservation laws. The first conservation law resulting from the *homogeneity of time* asserts that the quantity

$$E \equiv \sum_{i=1}^{s} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L$$

remains constant during the motion of a closed system; it is called the *energy* of the system. Mechanical systems whose energy is conserved are said to be *conservative*. A second conservation law follows from the *homogeneity of space*. More explicitly, in a closed mechanical system the vector

$$\mathbf{p} \equiv \sum_{a=1}^{N} \frac{\partial L}{\partial \mathbf{v}_a} \,,$$

known as the *momentum* of the system, remains constant in time. Differentiating the Lagrangian (2.3.2), in terms of the velocities of the particles the momentum reads

$$\mathbf{p} = \sum_{a=1}^{N} m_a \mathbf{v}_a \,,$$

where m_a denotes the *mass* of the ath particle ($a = 1, \ldots, N$). A third conservation law follows from the *isotropy of space*. In this case, the vector

$$\mathbf{M} \equiv \sum_{a=1}^{N} \mathbf{r}_a \times \mathbf{p}_a \,,$$

which is called the *angular momentum* of the system, is a conserved quantity of a closed system. In greater generality, Noether's theorem states that symmetries of a system give rise to conserved quantities (see e.g. [**81**, Chapter 4] or [**93**, Section 1.5]).

2.3.2. Electrodynamics. The interaction of particles in classical electrodynamics is described by means of *fields* [**105**].[5] Electromagnetic phenomena are governed by Maxwell's equations [**92**], specifying the behavior of an electric field $\mathbf{E}$ and a magnetic

[5] According to [**46**, Section 1-2], a "field" is any physical quantity which takes on different values at different points in space. It is precisely because $\mathbf{E}$ and $\mathbf{B}$ can be specified at every point in space that they are called "fields."

field **B**. Following [**73**, Section 2.4], the electric field **E** and the magnetic field **B** are vector-valued functions of position $\mathbf{x} \in \mathbb{R}^3$ and time t that enter the *Lorentz force law*: The force on a particle with charge q moving with velocity $\mathbf{v}$ is

$$\mathbf{F} = q\mathbf{E} + \frac{q}{c}\mathbf{v} \times \mathbf{B}\,,$$

where c is the speed of light, and the fields **E** and **B** are evaluated at the location of the particle (further details can be found in [**81**, Section 7.4]).

The behavior of **E** and **B** is governed by *Maxwell's equations*. Using the Heaviside-Lorentz convention (see [**73**, §1.1]), they read

$$\nabla \cdot \mathbf{E} = \rho \tag{2.3.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.3.4}$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 \tag{2.3.5}$$

$$\nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c}\mathbf{j}\,, \tag{2.3.6}$$

where ρ is the charge density and $\mathbf{j}$ is the current density, both of which are functions of position and time. The quantities ρ and $\mathbf{j}$ are not independent; indeed, they satisfy the so-called *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0\,, \tag{2.3.7}$$

which describes the conservation of electric charge.

For the homogeneous Maxwell equations, formula (2.3.4) is equivalent to the fact that **B** can be written as the curl of a *vector potential* **A**,

$$\mathbf{B} = \nabla \times \mathbf{A}\,. \tag{2.3.8}$$

With this, formula (2.3.5) becomes

$$\nabla \times \left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) = 0\,,$$

which is equivalent to the statement that the expression in brackets, $\mathbf{E} + (1/c)(\partial \mathbf{A}/\partial t)$, can be written as the gradient of a *scalar potential* ϕ,

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\,. \tag{2.3.9}$$

Note that **A** and ϕ are not uniquely determined; adjustments of **A** and ϕ which do not affect **E** and **B** are called *gauge transformations* (for details see [**73**, Chapter 9]). A frequently imposed gauge condition is the *Landau gauge* (or *Lorentz gauge*),

$$\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0\,. \tag{2.3.10}$$

Denoting the *wave operator* or *d'Alembertian* by $\Box$,

$$\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\,,$$

in the presence of the Lorentz gauge condition (2.3.10) Maxwell's equations read[6]

$$\Box\phi = \rho\,, \qquad \Box\mathbf{A} = \mathbf{j}\,.$$

[6]For a relativistic formulation of electrodynamics we refer to the end of the next subsection.

In case that ρ and $\mathbf{j}$ vanish, the Maxwell equations imply that the electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ satisfy the *wave equation*

$$\nabla^2 \psi = \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} \,,$$

which describes waves propagating at the speed of light c. Of special interest are waves that are periodic in space and time, which are referred to as *plane waves*. Periodicity in space is characterized by the *wave length* λ, while periodicity in time is represented by the *time period* T. The related quantities $k = 2\pi/\lambda$ and $\omega = 2\pi/T$ are called *wave number* and *angular frequency*, respectively. Such waves travels at a speed $c = \omega/k$. In three dimensions, a plane wave ψ at time t and position $\mathbf{r}$, moving in $\mathbf{k}$-direction (where $\mathbf{k}$ is called *wave vector*), can be written as

$$\psi(t, \mathbf{r}) = a e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}$$

with (real) *amplitude* a and $\omega = |\mathbf{k}| c$. The quantity $(\mathbf{k}\cdot\mathbf{r} - \omega t)$ is known as *phase*. For details see [**107**, §6] and [**148**, Section 3.1].

2.3.3. Special Relativity. Special relativity is based on two fundamental postulates: the principle of relativity and the constancy of the speed of light. The basic difficulty is to merge both facts, each of them being confirmed experimentally. Einstein's brilliant contribution [**37**] amounts to the insight that to this end, the common *notion of simultaneity* has to be abandoned. In simple terms, Einstein claims that *there is no absolute simultaneity*, thereby radically changing our understanding of space and time [**18**].

In order to describe processes taking place in nature, one considers a *reference frame*, that is, a system of coordinates in order to indicate the position of particles in space, together with a clock fixed in this system in order to indicate time. A reference frame, in which a particle which is not acted upon by external forces proceeds with constant velocity, is called an *inertial frame*. Classical mechanics is governed by the so-called *principle of relativity*, according to which all laws of nature are identical in all inertial systems of reference [**45, 105**]. On the other hand, it is found that the speed of light, usually denoted by c, is the *same* in *all* inertial systems of reference.[7]

Based on preparatory works by Lorentz [**116**], Einstein's solution to the problem raised is to claim that space and time become *relative*. More specifically, assume that S and S' are inertial frames, where S' moves with uniform velocity of magnitude v in x-direction with respect to S. Denoting the coordinates of space and time in S and S' by x, y, z, t and x', y', z', t', respectively, the so-called *Lorentz transformations* establish the following connection,

$$x' = \gamma(x - vt)\,, \qquad y' = y\,, \qquad z' = z\,, \qquad t' = \gamma(t - \frac{v}{c^2} x)\,, \tag{2.3.11}$$

where $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$. The Lorentz transformations have a number of immediate consequences: the *relativity of simultaneity*, *Lorentz contraction*, *time dilation* and the *addition of velocity* (for details we refer to [**81**, Chapter 3]).

It is well-known that the velocity of light has the numerical value

$$c = 2.998 \times 10^8 \text{ m/s}\,.$$

It is precisely the large value of c which explains why in most cases of practical interest classical mechanics appears to be sufficiently accurate. In other words, classical mechanics

[7] Concerning the connection to the famous Michelson-Morley experiment, carried out by Albert Michelson (1852–1931) and Edward Morley (1838–1923) in 1887, we refer to [**147**] and [**45**].

may be regarded as the limiting case of relativistic mechanics: Whenever the velocity v is small compared to the speed of light, i.e. $v \ll c$, the Lorentz transformations (2.3.11) go over to the *Galilean transformations*

$$x' = x - vt\,, \qquad y' = y\,, \qquad z' = z\,, \qquad t' = t$$

of classical mechanics (see e.g. [**105**, eq. (4.1)]).

Four-Vectors. In order to obtain a convenient framework, it is customary to introduce so-called *four-vectors* $x \in \mathbb{R}^4$ by

$$x^0 = ct\,, \qquad x^1 = x\,, \qquad x^2 = y\,, \qquad x^3 = z\,.$$

Introducing the abbreviation $\beta \equiv v/c$, the Lorentz transformations take the form

$$x^{0\prime} = \gamma(x^0 - \beta x^1)\,, \qquad x^{1\prime} = \gamma(x^1 - \beta x^0)\,, \qquad x^{2\prime} = x^2\,, \qquad x^{3\prime} = x^3\,. \tag{2.3.12}$$

Moreover, introducing the matrix Λ by

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,,$$

the Lorentz transformations (2.3.12) of a four-vector x can be written as $x' = \Lambda x$. It is common, however, in the physics literature to denote the vector whose components are $x^0, \ldots, x^3$ by x^μ rather than x (see [**73**]); thus in usual physics notation, this transformation of a four-vector x^μ is written as

$$x^{\mu\prime} = \Lambda^\mu_\nu x^\nu\,,$$

where in the last expression we employed the *Einstein summation convention*, stating that in any product of vectors and tensors[8] in which an index appears once as a subscript and once as a superscript, that index is to be summed from 0 to 3 (cf. [**73**]).

In other words, the matrix Λ allows us to describe transformations from S to S'. It is found that for any Lorentz transformation the following expression holds,

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x^{0\prime})^2 - (x^{1\prime})^2 - (x^{2\prime})^2 - (x^{3\prime})^2\,.$$

Introducing the *Minkowski metric* $\eta = \eta_{\mu\nu}$ by

$$\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)\,,$$

the *Minkowski inner product* given by

$$\langle x, y\rangle \equiv \eta_{\mu\nu} x^\mu y^\nu$$

is Lorentz invariant. The Euclidean space $\mathbb{R}^4$ endowed with $\langle ., . \rangle$ is called *Minkowski space*; in the sequel, we usually denote Minkowski space by $\mathscr{M}$. Defining the *covariant* four-vector x_μ by

$$x_\mu \equiv \eta_{\mu\nu} x^\nu$$

(the "original" four-vector x^μ is said to be *contravariant*), the Minkowski inner product of two four-vectors x^μ and y^μ can be written as

$$x \cdot y \equiv x_\mu y^\mu = x^0 y^0 - \mathbf{x} \cdot \mathbf{y}\,.$$

[8] Concerning tensor calculus, we refer the interested reader to [**115**].

For simplicity, the dot is sometimes omitted, i.e. $xy \equiv x \cdot y$. Moreover,

$$x^2 \equiv x \cdot x = (x^0)^2 - \mathbf{x}^2 .$$

It is important to observe that x^2 need not be positive. Indeed, a four-vector x is said to be

$$\begin{aligned} \textit{timelike} \qquad & \text{if } x^2 > 0 \\ \textit{spacelike} \qquad & \text{if } x^2 < 0 \\ \textit{lightlike} \qquad & \text{if } x^2 = 0 . \end{aligned}$$

Lightlike vectors are also referred to as *null* vectors. The subsets

$$L := \{x \in \mathscr{M} : x^2 = 0\} , \qquad I := \{x \in \mathscr{M} : x^2 > 0\} , \qquad J := \{x \in \mathscr{M} : x^2 \geq 0\}$$

are called *light cone*, *interior light cone* and *closed light cone*, respectively. They give rise to the following decomposition of Minkowski space $\mathscr{M}$,

$$\mathscr{M} = L \,\dot\cup\, I \,\dot\cup\, \mathscr{M} \setminus J .$$

For an illustration of the light cone we refer to [**89**, Figure 8] or [**123**, Figure 1.3.1]. In this framework, the notation of derivatives on $\mathbb{R}^4$ is

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \qquad \text{for all } \mu = 0, \ldots, 3 ,$$

so that, with respect to traditional coordinates $\mathbf{x}$ and t,

$$(\partial_0, \ldots, \partial_3) = (c^{-1}\partial_t, \nabla_{\mathbf{x}}) , \qquad (\partial^0, \ldots, \partial^3) = (c^{-1}\partial_t, -\nabla_{\mathbf{x}}) . \tag{2.3.13}$$

Further details can be found in [**53, 73, 81, 105**]. We point out that the light cone plays a central role in the theory of causal fermion systems.

Energy and Momentum. In what follows, we consider a particle of rest mass m, which is described by the four-vector x^μ. Then its *four-velocity* u^μ is introduced by

$$u^\mu \equiv \frac{dx^\mu}{ds} ,$$

where $ds = c\gamma\, dt$ (with γ defined above), and its *four-momentum* $p = p^\mu$ is defined by

$$p^\mu \equiv m u^\mu .$$

Making use of the fact that $u^\mu = (\gamma c, \mathbf{v})$, the four-momentum p^μ takes the form

$$p^\mu = (p^0, \mathbf{p}) = \gamma(mc, m\mathbf{v}) .$$

Defining the *relativistic energy* E by (cf. [**81**, eq. (3.41)])

$$E \equiv \gamma mc^2 = \frac{mc^2}{\sqrt{1 - v^2/c^2}} ,$$

we conclude that $p^0 = E/c$ and $p_\mu p^\mu = m^2 c^2$. From this we deduce that

$$E^2 - \mathbf{p}^2 c^2 = m^2 c^4 , \tag{2.3.14}$$

implying that $E = mc^2$ is the *rest energy* of the particle under consideration. The set of admissible four-momenta p^μ satisfying (2.3.14) forms a *hyperboloid* in $\mathbb{R}^4$.

The space of four-momenta $p^\mu \in \mathbb{R}^4$, which for clarity we denote by $\hat{\mathscr{M}}$, is known as *momentum space*. Identifying momentum space $\hat{\mathscr{M}}$ with Minkowski space $\mathscr{M}$, the Minkowski inner product can be considered as a mapping

$$\langle .,. \rangle : \hat{\mathscr{M}} \times \mathscr{M} \to \mathbb{R}\,, \qquad (p, x) \mapsto \langle p, x \rangle = g_{\mu\nu} p^\mu x^\nu\,.$$

Given $m > 0$, one introduces the so-called *mass shell* by

$$C_m := \{p \in \hat{\mathscr{M}} : p^2 = m^2\}\,.$$

Furthermore, one often distinguishes its subsets

$$C_m^+ := \{p \in \hat{\mathscr{M}} : p^2 = m^2,\, p_0 > 0\} \qquad \text{and} \qquad C_m^- := \{p \in \hat{\mathscr{M}} : p^2 = m^2,\, p_0 < 0\}\,,$$

which are known as *upper* and *lower mass shell*, respectively. The sets $C_m^\pm$ form the two parts of a hyperbola in momentum space. In the case $m = 0$, the corresponding sets C_0^+ and C_0^- are called *forward* and *backward light cone*, respectively (cf. [**73**, Section 1.3]). These constructions will be of crucial importance in the theory of causal fermion systems.

Relativistic Formulation of Electrodynamics. Applying the notation of special relativity to electrodynamics gives rise to particularly concise formulas. More specifically, introducing the *electromagnetic 4-potential* by

$$A^\mu = (\phi, \mathbf{A}) \qquad \text{or} \qquad A_\mu = (\phi, -\mathbf{A})\,,$$

the electric and magnetic fields are combined into the *electromagnetic field tensor*

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,, \qquad \text{or} \qquad F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu\,.$$

Thus the Maxwell equations (2.3.4) and (2.3.5) become

$$\partial_\kappa F_{\mu\nu} + \partial_\mu F_{\nu\kappa} + \partial_\nu F_{\kappa\mu} = 0\,,$$

and the equations (2.3.3) and (2.3.6) take the form

$$\partial_\mu F^{\mu\nu} = j^\nu\,,$$

where the *4-current density* j^μ is defined by

$$j^\mu \equiv (\rho, \mathbf{j})\,.$$

The Lorentz gauge condition (2.3.10) reads

$$\partial_\mu A^\mu = 0\,,$$

and the continuity equation (2.3.7) can be written as

$$\partial_\mu j^\mu = 0\,.$$

The previous relations may also be expressed in terms of differential forms, making use of the Hodge star operator; for details we refer to [**73**, Section 2.4] or [**170**].

Let us finally point out that the Maxwell equations (2.3.3)–(2.3.6) can be obtained from a variational principle. To this end, one considers the action (cf. [**105**, eq. (28.6)])

$$S = -\frac{1}{c^2} \int A_\mu j^\mu \, d\Omega - \frac{1}{16\pi c} \int F_{\mu\nu} F^{\mu\nu} \, d\Omega\,,$$

where $d\Omega = c\, dt\, dx\, dy\, dz$ (see [**105**, eq. (27.4)]). Following [**105**, §30], by varying the electromagnetic potential one obtains Maxwell's equations (2.3.3) and (2.3.6), whereas the equations (2.3.4) and (2.3.5) are a consequence of (2.3.8) and (2.3.9) by [**105**, §26].

2.3.4. General Relativity. The main reason for the fact that solely Einstein's name is associated with the theory of relativity is due to the further development: So far, the principle of relativity is restricted to inertial frames; influenced by the Austrian physicist and philosopher ERNST MACH (1838–1916), however, Einstein was guided by the postulate that the "true" laws of physics should hold in any reference frame. The starting point for his considerations was the insight that *no physical experiment can distinguish between gravitational and inertial forces*. This is known as *Einstein's principle of equivalence* [**18, 149**]. The resulting "general theory of relativity," including a new theory of gravity, changed our ideas of the universe in a fundamental way.

In short, Einstein's results can be summarized as follows: In the theory of general relativity, spacetime is described mathematically by a four-dimensional Lorentzian manifold (M, g), which is a semi-Riemannian manifold either of signature $(-,+,+,+)$ or $(+,-,-,-)$.[9] At each point $p \in M$, the tangent space T_pM can be identified with Minkowski space, and special relativity can be regarded as general relativity of Minkowski spacetime (for details see [**10, 125**]). The famous *Einstein equations* read

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} \,. \tag{2.3.15}$$

where $G_{\mu\nu}$ denotes the *Einstein gravitational tensor*, $T_{\mu\nu}$ is known as *stress-energy tensor* (or *energy-momentum tensor*), and κ denotes Newton's gravitational constant. The Einstein field equations (2.3.15) describe the behavior of curvature of spacetime in the presence of matter. For clarity we point out that $G_{\mu\nu}$ crucially depends on the metric tensor g, whereas $T_{\mu\nu}$ is to be determined from physical observations dealing with the distribution of matter and energy (cf. [**89**, Chapter 3] or [**120**, Chapter 5]). In simple terms [**160, 125**], the Einstein equations often are summarized by saying that

> "Spacetime tells matter how to move,
> matter tells spacetime how to curve."

As being of relevance for the theory of causal fermion systems, let us consider basic objects on Lorentzian manifolds in some more detail.[10] More generally, let (M, g) be a semi-Riemannian manifold. Denoting the set of all smooth vector fields on M by $\mathfrak{X}(M)$ and the set of all smooth real-valued functions on M by $\mathfrak{F}(M)$, a *connection* is a function

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$$

with the properties that

$$\begin{aligned} \nabla_V (X+Y) &= \nabla_V X + \nabla_V Y \,, \\ \nabla_{fV+hW}(X) &= f \nabla_V X + h \nabla_W X \,, \\ \nabla_V (fX) &= f \nabla_V X + V(f) X \end{aligned}$$

for all $f, h \in \mathfrak{F}(M)$ and all $X, Y, V, W \in \mathfrak{X}(M)$. The vector field $\nabla_X Y$ is said to be the *covariant derivative* of Y with respect to X. The *Lie bracket* of the ordered pair of vector fields X and Y is a vector field $[X, Y]$ which acts on a smooth function f by

$$[X, Y](f) = X(Y(f)) - Y(X(f)) \,.$$

Next, the *torsion tensor* T of ∇ is the function $T : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ given by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \,.$$

[9] The two sign conventions are related by considering the metric g or $-g$, respectively.

[10] Our explanations are due to [**10**] with signature $(-,+,+,+)$, as is common in many textbooks on general relativity; concerning the sign conventions we also refer to [**125**].

The *curvature* $R(.,.)$ of ∇ is a function which assigns to each pair $X, Y \in \mathfrak{X}(M)$ the f-linear map $R(X,Y) : \mathfrak{X}(M) \to \mathfrak{X}(M)$ given by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \,.$$

Thus curvature provides a measure of the non-commutativity of ∇_X and ∇_Y. Given a semi-Riemannian manifold (M, g), there is a unique connection ∇ on M which is *metric compatible*, i.e.

$$Z(g(X,Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) \,,$$

as well as *torsion free*, i.e.

$$[X,Y] = \nabla_X Y - \nabla_Y X$$

for all $X, Y, Z \in \mathfrak{X}(M)$. This connection ∇ is called *Levi-Civita connection*. The trace of the curvature tensor is the *Ricci curvature*; for each $p \in M$, the Ricci curvature may be interpreted as the bilinear map $\mathrm{Ric}_p : T_pM \times T_pM \to \mathbb{R}$. In terms of the *Ricci tensor* $R_{\mu\nu}$, Ricci curvature may be represented as

$$\mathrm{Ric}_p(v,w) = R_{\mu\nu} v^\mu w^\nu \,,$$

and the trace of the Ricci curvature is the *scalar curvature* R, i.e. $R = R^\mu{}_\mu$. Then the Einstein tensor $G_{\mu\nu}$ is given by (also see [**22**, eq. (4.29)]

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \,,$$

and it satisfies $\nabla^\mu G_{\mu\nu} = 0$. Just as Maxwell's equations govern how the electric and magnetic fields respond to charges and currents, Einstein's field equations determine how the metric responds to energy and momentum.

In his paper [**38**], Einstein suggests to modify his field equations by introducing an additional constant (which is referred to as *cosmological constant* Λ). In this case, the *vacuum field equations* read

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 0 \,. \tag{2.3.16}$$

In 1915, Hilbert [**90**] observed that the Einstein equations can be recovered by means of a variational principle. More precisely, varying the so-called *Einstein-Hilbert action*

$$S_{\mathrm{EH}} = \int_M \left[\frac{1}{2\kappa} (R - 2\Lambda) \right] \, d\mu$$

with respect to the metric g yields (2.3.16), where $d\mu := \sqrt{-\det g}\, d^4x$ (for details we refer to [**22, 83**]). Actually, introducing an action S in the fashion of [**53**, eq. (1.1.9)], it is possible to recover the equations of motion, the Maxwell equations as well as the Einstein equations by varying the action S with respect to the spacetime curve, the electromagnetic potential and the metric, respectively (cf. [**53**, Section 1.1]). The non-quantized field equations derived so far are also referred to as "classical."

2.3.5. Quantum Mechanics. From the conceptual point of view, it is instructive to motivate quantum mechanics by illustrating the famous *double-slit experiment* (for details see [**148**, Chapter 3]) in order to expose the inadequacy of classical physics. More precisely, performing the double-slit experiment for single photons, one observes that each photon carries the same energy E and the same momentum $\mathbf{p}$, and its energy is given by $E = |\mathbf{p}|c$. Applying the relativistic equation $E^2 = \mathbf{p}^2 c^2 + m^2 c^4$ (see (2.3.14)), one infers that each photon is a particle of mass $m = 0$. Moreover, varying the (angular) frequency ω of the light source, one discovers that $E = \hbar\omega$ and $p = \hbar k$ (where $k = |\mathbf{k}|$

is the absolute value of the wave vector $\mathbf{k}$ and $\hbar$ is *reduced Planck's constant*), implying that energy and momentum are *quantized* [**130**].

From these facts Born drew the following conclusion: Each particle has associated with it a wave function $\psi(t, \mathbf{x})$ such that $|\psi(t, \mathbf{x})|^2$ gives the probability of finding it at a point $\mathbf{x}$ at time t. This is known as *wave-particle duality*. The dynamics of the particle is described in terms of its wave function $\psi(t, \mathbf{x})$, which in non-relativistic quantum mechanics is determined by *Schrödinger's equation*. The resulting quantum theory is then based on a few fundamental postulates, which take into account these experimental findings (see [**148**, Chapter 4]). As $|\psi(t, \mathbf{x})|^2$ is interpreted as probability of finding a particle with wave function ψ at time t at a point $\mathbf{x}$, for each time t the integral over the whole space must be normalized,

$$\int_{\mathbb{R}^3} |\psi(t, \mathbf{x})|^2 \, d^3x = 1 \,.$$

Thus for every t, a wave function $\psi(t, \cdot)$ is contained in the Hilbert space of square-integrable functions $L^2(\mathbb{R}^3, \mathbb{C})$, endowed with the inner product $\langle \,.\,|\,.\,\rangle_{L^2(\mathbb{R}^3, \mathbb{C})}$. This property is reflected in the first postulate of quantum mechanics, according to which the *state* of a particle is represented by a normalized vector in a complex Hilbert space, and physically measurable quantities (referred to as *observables*) correspond to self-adjoint operators on the Hilbert space (thereby establishing a connection to spectral calculus, for details see [**31, 152**]). Instead of entering the postulates of quantum mechanics in more detail, let us stick to the most important consequences concerning the theory of causal fermion systems.

Besides the classical "orbital" angular momentum, a specific feature of quantum mechanics is that every elementary particle comes with a certain "intrinsic" angular momentum which is independent of its motion in space [**107**]. This intrinsic angular momentum is called *spin*. Denoting the spin of a particle by s, it is remarkable that s only takes integer or half-integer values, i.e. $s \in \{0, 1/2, 1, 3/2, \ldots\}$. Particles with integer spin are called *bosons*, whereas particles with half-integer spin are known as *fermions* [**81, 156**]. Concerning fermions with spin $s = 1/2$, there are $(2s + 1) = 2$ possible spin orientations (see [**81**, Section 4.2]); they are usually referred to as "up" and "down" (for details see [**142**]).

As mentioned before, the dynamics in non-relativistic quantum mechanics obeys Schrödinger's equation, in which case relativistic effects are not taken into consideration. However, as the principles of special relativity are generally accepted, a correct quantum theory should satisfy the requirement of relativity: laws of motion valid in one inertial system must hold in all inertial systems. Stated mathematically, relativistic quantum theory must be formulated in a Lorentz covariant form [**14**]. In other words, Schrödinger's equation is unsatisfactory from the relativistic point of view, as it treats space and time on a different footing. In relativistic mechanics, momentum $\mathbf{p}$ and energy E of a free particle combine to a four-vector $p^\mu = (E, \mathbf{p})$ (in "natural units," in which Planck's constant $\hbar$ and the speed of light c both equal one, $\hbar = c = 1$). The recipe for quantizing energy and momentum amounts to replacing $p^\mu \to i\partial^\mu$, where ∂^μ is given by (2.3.13). Applying this replacement rule to the free-particle energy-momentum relation $p^2 = m^2$ yields the relativistic equation

$$(\Box - m^2)\phi = 0 \,,$$

known as *Klein-Gordon equation* (where $\Box \equiv \partial_\mu \partial^\mu$ is the wave operator). Complex-valued solutions of this equation describe a scalar particle (that is, a particle without spin) of

mass m. Unfortunately, the scalar component ρ which appears in the corresponding continuity equation is not positive definite and therefore cannot be understood as a probability density. For further details we refer to [**53, 73, 81, 146, 152**].

Dirac's aim was to avoid this difficulty of negative probability densities, taking into account relativistic covariance. Dirac's strategy was to "factorize" the relativistic energy-momentum relation $p^2 = m^2 c^2$ (cf. [**81**]). To this end, he introduced the so-called *Dirac matrices* $\gamma^\mu \in \mathbb{C}^{4\times 4}$ ($\mu = 0, \ldots, 3$), satisfying the anti-commutation relations

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} \qquad \text{for all } \mu, \nu = 0, \ldots, 3$$

(with Minkowski metric $\eta^{\mu\nu}$). This approach finally led him to set up the *Dirac equation* in Minkowski space, which in natural units takes the form

$$(i\partial\!\!\!/ - m)\psi = 0\,, \tag{2.3.17}$$

where $\partial\!\!\!/ \equiv \gamma^\mu \partial_\mu$ is known as *Dirac slash* (cf. [**81, 156**]). The Dirac equation is Lorentz covariant [**14**]. The solutions ψ of the Dirac equation are referred to as *Dirac spinors*, which are mappings $\psi : \mathscr{M} \to \mathbb{C}^4$ from Minkowski space $\mathscr{M}$ to spinor space $\mathbb{C}^4$; they describe particles of spin 1/2. Introducing the *adjoint spinor* $\overline{\psi} \equiv \psi^\dagger \gamma^0$ (where by $\psi^\dagger$ we denote the adjoint of ψ with respect to the Euclidean inner product on $\mathbb{C}^4$) gives rise to an indefinite inner product $\overline{\psi}\phi$ on $\mathbb{C}^4$ of signature $(2,2)$, which will be of central relevance in the theory of causal fermion systems. The success of Dirac's intuition lies in the fact that the scalar component ρ of the corresponding continuity equation is given by $\psi^\dagger \psi = |\psi|^2$, and thus non-negative. For a probabilistic interpretation of the wave functions, one requires the *normalization condition*

$$\int_{\mathbb{R}^3} (\psi^\dagger \psi)(t, \mathbf{x})\, d^3\mathbf{x} = 1\,, \tag{2.3.18}$$

thus giving rise to the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3)^4$, endowed with the scalar product

$$(\psi \mid \phi) := \int_{\mathbb{R}^3} (\overline{\psi}\gamma^0 \phi)(t, \mathbf{x})\, d^3\mathbf{x} \qquad \text{for all } \psi, \phi \in H \tag{2.3.19}$$

for any time t (cf. [**53**, Section 1.2], [**73**, Section 4.1] and [**152**, Section 1.3]).

In what follows we make use of the *energy function* ω, defined by

$$\mathbb{R}^3 \ni \mathbf{k} \mapsto \omega(\mathbf{k}) := \sqrt{\mathbf{k}^2 + m^2} \in \mathbb{R}\,.$$

Introducing for any $\mathbf{k} \in \mathbb{R}^3$ the orthogonal projections $p_\pm(\mathbf{k}) : \mathbb{C}^4 \to \mathbb{C}^4$ by

$$p_\pm(\mathbf{k}) := \frac{k\!\!\!/ + m}{2k^0} \gamma^0 \Big|_{k^0 = \pm\omega(\mathbf{k})}\,,$$

the spinor space $\mathbb{C}^4$ decomposes into two orthogonal subspaces,

$$\mathbb{C}^4 = \mathcal{W}_\mathbf{k}^+ \oplus \mathcal{W}_\mathbf{k}^-\,,$$

where $\mathcal{W}_\mathbf{k}^\pm = p_\pm(\mathbf{k})(\mathbb{C}^4)$ for any $\mathbf{k} \in \mathbb{R}^3$, and a general solution of the Dirac equation takes the form (cf. [**31**, eq. (9.24)])

$$\psi(t, \mathbf{x}) = \int_{\mathbb{R}^3} \frac{d\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \left(e^{-i\omega(\mathbf{p})t} \psi_+(\mathbf{p}) + e^{i\omega(\mathbf{p})t} \psi_-(\mathbf{p}) \right)$$

with $\psi_\pm(\mathbf{p}) \in \mathcal{W}_\mathbf{p}^\pm$ for all $\mathbf{p} \in \mathbb{R}^3$. It is an important consequence that the solution space $\mathcal{H}$ of the Dirac equation splits into a positive and negative energy subspaces,

$$\mathcal{H} = \mathcal{H}^+ \oplus \mathcal{H}^-\,,$$

where $\mathcal{H}^\pm = \{\psi \in \mathcal{H} : \psi_\mp = 0\}$ (for details see [**31, 126**]). Solutions of *positive* energy are interpreted as *particles*; the physical interpretation of negative energy solutions, on the other hand, is due to Dirac [**34**]. Based on the *Pauli exclusion principle*, according to which no two fermions can occupy the very same state, Dirac proposed that "all the states of negative energy are occupied except perhaps a few of small velocity" in order to prevent particles of positive energy to fall down into negative energy states. In accordance with Dirac's interpretation [**34**], these few vacant states, regarded as "holes" in the sea of negative energy particles, were later discovered as *anti-particles* (cf. [**156**]). This concept, also referred to as a "Dirac sea," became unpopular in quantum field theory;[11] however, Dirac's original idea of a sea of particles was revived in the theory of causal fermion systems (see [**57**]).

Whenever a wave function $\psi = \psi(t, \mathbf{x})$ is given, the four-vector $x = (t, \mathbf{x}) \in \mathbb{R}^4$ is said to be in "position space." Denoting position space by $\mathscr{M} \simeq \mathbb{R}^4$ and assuming that $\psi \in L^1(\mathscr{M})$, its Fourier transform is given by (for details see [**73**])

$$\widehat{\psi}(k) = \int_{\mathscr{M}} e^{ikx} \psi(x)\, d^4x \qquad \text{for all } k \in \mathbb{R}^4\,,$$

where $k \in \mathbb{R}^4$ is said to be in "momentum space." Accordingly, we refer to $\hat{\mathscr{M}} \simeq \mathbb{R}^4$ as *momentum space*. For mathematical details we refer to [**71**].

Let us finally deal with the famous uncertainty principle, one of the fundamental principles of quantum mechanics, discovered by Heisenberg in 1927 (cf. [**107**]). Due to the axioms of quantum mechanics, to each quantum mechanical system is associated a Hilbert space, and every measurable quantity (referred to as an "observable," such as energy, position or momentum) is represented by a self-adjoint operator [**152**]. Given a self-adjoint operator A on a Hilbert space $(H, \langle .,. \rangle)$, for any unit vector $\psi \in H$ the *expectation value* of A in the state ψ is defined by $\langle A \rangle_\psi = \langle \psi, A\psi \rangle$, and the *uncertainty of* A, denoted by $\Delta_\psi A$, is given by $(\Delta_\psi A)^2 = \langle A^2 \rangle_\psi - (\langle A \rangle_\psi)^2$ (cf. [**86**, Section 3.6]). Introducing the *position operators* $(X^i)_{i=1,2,3}$ by

$$(X^i \psi)(\mathbf{x}) = x^i \psi(\mathbf{x}) \qquad \text{for all } i = 1, 2, 3\,,$$

which amounts to a "multiplication by x^i," the expectation value of X^i in the state ψ is then $\langle X^i \rangle_\psi = \langle \psi, X^i \psi \rangle$ for all $i = 1, 2, 3$. Moreover, one can introduce the *momentum operators* $(P^i)_{i=1,2,3}$ such that $[X^i, P^j] = \delta_{ij} i\hbar$ for all $i, j = 1, 2, 3$ (cf. [**86**, Chapter 3]). Making use of the fact that the position and momentum operators X^i and P^i do not commute for all $i = 1, 2, 3$, for any state ψ the *Heisenberg uncertainty principle* reads (for details see [**86**, Chapter 12] or [**73, 82, 107, 148**])

$$(\Delta_\psi X^i)(\Delta_\psi P^i) \geq \frac{\hbar}{2\pi} \qquad \text{for all } i = 1, 2, 3\,,$$

implying that position and momentum cannot be measured simultaneously. In other words, in order to probe *small distances* one requires *high energies* [**81**].

2.3.6. Group Theory in Physics. The importance of group theory in physics was recognized soon after the discovery of quantum mechanics [**151**], based on fundamental works by Weyl [**158, 159**]. In the sequel, group theory evolved to a substantial part of modern theoretical physics. For this reason, let us briefly compile some basics from group

[11] To quote Julian Schwinger (cf. [**156**, p. 14]), "the picture of an infinite sea of negative energy electrons is now best regarded as a historical curiosity, and forgotten."

theory required in the following. To begin with, we recall that, for any $n \in \mathbb{N}$, the *unitary group* $\mathrm{U}(n)$ is given by

$$\mathrm{U}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^\dagger = A^{-1}\}\,,$$

where $\mathrm{GL}(n,\mathbb{C}) = \{A \in \mathbb{C}^{n\times n} : \det(A) \neq 0\}$. The *special unitary group* is defined by

$$\mathrm{SU}(n) = \{A \in \mathrm{U}(n) : \det(A) = 1\}\,.$$

Clearly, the norm of a spinor $\psi(x) \in \mathbb{C}^n$ is invariant under unitary transformations

$$\psi'(x) = U\psi(x) \qquad \text{with } U \in \mathrm{U}(n)\,.$$

We recall that in general, a *Lie group* G is a differentiable manifold which is also a group such that the mapping

$$G \times G \to G\,, \qquad (\sigma,\tau) \mapsto \sigma\tau^{-1}$$

is smooth (see e.g. [**155**]). Whenever a Lie group G is a subgroup of $\mathrm{GL}(n,\mathbb{C})$, it is also referred to as *matrix Lie group* [**85**]. Both $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are (compact) matrix Lie groups. In physical applications one is typically interested in a fixed Lie group G (the "gauge group") which represents a symmetry of the theory (cf. [**118**, Section 6.3]).

As $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are the most important Lie groups in particle physics, let us restrict attention to matrix Lie groups in what follows. According to [**85**, Chapter 2], the *exponential map* is defined by

$$\exp : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}\,, \qquad X \mapsto e^X := \sum_{n=0}^{\infty} \frac{X^n}{n!}\,.$$

Whenever G is a matrix Lie group, the *Lie algebra of* G, denoted by $\mathfrak{g}$, is defined as the set of all matrices X such that $e^{tX} \in G$ for all $t \in \mathbb{R}$. For clarity we point out that the *physicists' convention* is to consider the map $X \mapsto e^{iX}$ instead of $X \mapsto e^X$. Accordingly, physicists are accustomed to think of the Lie algebra of G as the set of all matrices X such that $e^{itX} \in G$ for all $t \in \mathbb{R}$. Moreover, physics literature does not always distinguish clearly between a Lie group and its Lie algebra.

In general, a *Lie algebra* $\mathfrak{g}$ over $\mathbb{R}$ is a real vector space $\mathfrak{g}$ together with a bilinear operator $[.,.] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ (called the *Lie bracket*) such that for all $X, Y, Z \in \mathfrak{g}$,

(i) $[X, X] = 0$ for all $X \in \mathfrak{g}$, and
(ii) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ for all $X, Y, Z \in \mathfrak{g}$ (*Jacobi identity*).

Note that each Lie group is closely related to a corresponding Lie algebra (cf. [**155**]); indeed, whenever G is a Lie group with unit element e, the vector space $\mathfrak{g} := T_eG$ is called the *Lie algebra* of G (cf. [**20**, Definition (2.4)]). It becomes a Lie algebra with the *Lie product* (cf. [**20**, eq. (2.13)])

$$[X, Y] = XY - YX\,.$$

Whenever G is a matrix Lie group with Lie algebra $\mathfrak{g}$, and $X_1, \ldots, X_n$ is a basis for $\mathfrak{g}$ (as a vector space), then for all $i, j \in \{1, \ldots, n\}$ the Lie bracket $[X_i, X_j]$ can be written uniquely in the form

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ijk} X_k\,.$$

The constants c_{ijk} are called the *structure constants* of $\mathfrak{g}$ (see [**85**, §2.8.1]). In physics literature, the basis vectors $X_1, \ldots, X_n$ are known as *generators* (cf. [**3, 127, 156**]).

2.3.7. Introduction to Elementary Particles. To begin with, elementary particles are thought of as objects which are not made up of more fundamental ones; one thus can say that elementary particles are the fundamental constituents of all objects in the universe.

For a historical introduction to elementary particles we refer the interested reader to [**81**, Chapter 1] and [**127**, Chapter 9]. Nevertheless, it might be instructive to motivate our present understanding of elementary particles as follows. The idea that matter is composed of "indivisible" atoms goes back to the ancient Greeks [**141**]. However, our knowledge of the detailed structure of atoms originates from scattering experiments due to Ernest Rutherford (1871–1937), according to which the positive charge as well as most of the mass are concentrated in the center of the atom, the *nucleus*. In the 20th century it turned out that the nucleus itself is made up of *neutrons* and *protons*. The electrons, already discovered in 1897 by Joseph John Thompson (1856-1940), were supposed to "orbit" around the nucleus. This picture summarizes the period of classical elementary particle physics [**81**].

In the mean time, further experimental findings revealed that neutrons and protons themselves consist of even more fundamental constituents, so-called *quarks*, which are denoted by the first letters of their names. More precisely, there are six kinds of quarks, known as *flavors*, referred to as up (u), down (d), strange (s), charm (c), bottom (b) and top (t), and each quark comes in three *colors* (red, green, blue). They fall into doublets (called "families" or "generations"); a first generation (u, d), a second generation (c, s) and a third generation (t, b). It should be noticed that free quarks have never been observed (referred to as *quark confinement*). Electrons, on the other hand, belong to the class of *leptons*. There are six types of leptons, called *flavors*, comprising three charged leptons (*electron* e, *muon* μ and *tauon* τ) and three neutral leptons (*electron neutrino* ν_{e}, *muon neutrino* ν_μ and *tau neutrino* ν_τ). By contrast to quarks, leptons can exist as free particles. Similarly as quarks, leptons come in three generations: a first generation (ν_{e}, e), a second generation (ν_μ, μ) and a third generation (ν_τ, τ). This comprises all known elementary particles, which therefore can be summarized in the following table (where the arrows $\uparrow$ and $\downarrow$ refer to "isospin up" and "isospin down," respectively; for details concerning weak isospin see below):

		1^{st} generation	2^{nd} generation	3^{rd} generation	I_3
quarks	red	u	c	t	$\uparrow$
	green	u	c	t	$\uparrow$
	blue	u	c	t	$\uparrow$
	red	d	s	b	$\downarrow$
	green	d	s	b	$\downarrow$
	blue	d	s	b	$\downarrow$
leptons	neutral	ν_{e}	ν_μ	ν_τ	$\uparrow$
	charged	e	μ	τ	$\downarrow$

(2.3.20)

Bearing in mind that to each particle there is a corresponding *anti-particle*, there are 48 elementary particles in total, which are all Dirac particles (that is, particles of spin 1/2 which obey the Dirac equation). Thus one can say that on the fundamental level, all matter is described by the Dirac equation. In the theory of causal fermion systems, Table (2.3.20) will play an important role in the construction of the fermionic projector [**53**]. The fact that all particles in Table (2.3.20) are fermions motivates to speak of "fermionic" projector and causal "fermion" systems.

Concerning the terminology of subatomic particles [**73**], there are actually two basic dichotomies: all particles are either bosons, with integer spin, or fermions, with half-integer spin; and all particles are either *hadrons*, which participate in the strong interaction, or non-hadrons. The fundamental fermions comprise quarks, which are hadrons, and leptons, which are not. Quarks combine in triplets to make *baryons*, and in quark/anti-quark pairs to make *mesons*. Baryons are fermions, whereas mesons are bosons. The most familiar baryons are the proton and the neutron.

Elementary particles are characterized by several *quantum numbers* which are often said to be "intrinsic" to them: *electric charge* (Q), *weak isospin* (I), *baryon number* (B) and *lepton number* (L). Note that, in natural units, electric charge is dimensionless, and the unit of charge in Heaviside-Lorentz units is simply the number 1. The basic physical constant is then the fundamental unit of charge, the absolute value of the charge of an electron, which usually is denoted by e (cf. [**73**]). The electron, muon and tauon have electric charge $-e$, while each neutrino has electric charge zero [**81, 96**]. Down, strange and bottom quarks carry electric charge $-1/3e$, while up, charm and top quarks have electric charge $2/3e$.

As mentioned before, elementary particles are either *hadrons* or *leptons*, depending on whether they respond to the strong interaction (hadrons) or not (leptons). Now to each lepton is associated the *lepton number* $L = 1$, and $L = -1$ to all anti-leptons; all other particles have $L = 0$. In a similar fashion, to each baryon is assigned the *baryon number* $B = 1$, and $B = -1$ to all anti-baryons, while all other particles have $B = 0$. Taking into account that baryons are composed of three quarks, the baryon number of a quark is $B = 1/3$ (see [**11**]). Next, there are particles which in experiments behaved so unexpectedly that they were called "strange particles." Accordingly, the so-called *strangeness number* S was introduced; in particular, the strange quark has strangeness $S = -1$, whereas all other quarks have strangeness $S = 0$ (cf. [**11**, Table 13.4]). A related quantity, called *hypercharge*, is just the sum of strangeness S and baryon number B, i.e. $Y = S + B$ (cf. [**39**]). There is one more property, *(weak) isospin* (short for "isotopic spin"), which has mathematical properties similar to those of spin, but which has no direct physical relationship to spin (cf. [**39**, Section 17-3] and [**127**] for its connection to group theory). Of particular interest is its third component, usually denoted by I_3, which satisfies the relation $Y = 2(Q - I_3)$ (cf. [**73**] and [**81**, eq. (9.139)]). Similarly to spin, isospin either takes integer or half-integer values, i.e. $I_3 \in \{0, 1/2, 1, \ldots\}$; accordingly, a representation of isospin $I_3 = 1/2$ contains $(2I_3 + 1) = 2$ states, referred to as "isospin up" and "isospin down," thus giving rise to *isospin doublets* which coincide with the above generations of fermions (cf. [**39**, Section 18-8] and [**3**, Chapter 12]). For a group theoretical treatment we refer the interested reader to [**151**, Chapter 5]. Moreover, a comprehensive overview of quantum numbers can be found in [**39**, Table 17-1].

The reason for introducing the adjoint spinor $\overline{\psi} = \psi^\dagger \gamma^0$ is that $\overline{\psi}\psi$ is a relativistic invariant. Moreover, considering the *parity transformation*

$$P : \mathbb{R}^3 \to \mathbb{R}^3 , \qquad \mathbf{x} = (x, y, z) \mapsto (-x, -y, -z) = -\mathbf{x} ,$$

the quantity $\overline{\psi}\psi$ is invariant under P. For this reason, it is called a *scalar*. Introducing the so-called *pseudoscalar matrix* (cf. [**131**, eq. (A.2.18)])

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 ,$$

the quantity $\overline{\psi}\gamma^5\psi$ changes sign under P, i.e. $P(\overline{\psi}\gamma^5\psi) = -\overline{\psi}\gamma^5\psi$; for this reason, it is referred to as *pseudoscalar*. In a similar fashion, for all $\mu = 0, \ldots, 3$ one refers to the

quantity $\overline{\psi}\gamma^\mu\psi$ as a *vector*, whereas $\overline{\psi}\gamma^\mu\gamma^5\psi$ is said to be a *pseudovector* (or *axial vector*). Introducing the *bilinear covariants* $\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ gives rise to the *anti-symmetric tensor* $\overline{\psi}\sigma^{\mu\nu}\psi$. Making use of the fact that $\mathbb{1}, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5$ and $\sigma^{\mu\nu}$ constitute a basis of $\mathbb{C}^{4\times 4}$, any 4×4-matrix can be written as a linear combination of the resulting 16 terms (for details see [**81**, Section 4.4 and Section 7.3] or [**129**]). From the vector and pseudovector matrices, one can form the *(Dirac) current* j^μ and the *axial vector current* $j^{\mu 5}$ by (cf. [**129**, eq. (3.73)])

$$j^\mu = \overline{\psi}\gamma^\mu\psi \qquad \text{and} \qquad \gamma^{\mu 5} = \overline{\psi}\gamma^\mu\gamma^5\psi\,,$$

respectively. Whenever ψ satisfies the Dirac equation, the current j^μ is a conserved quantity, that is, $\partial_\mu j^\mu = 0$. In the case $m = 0$, the axial vector current $j^{\mu 5}$ is also conserved, that is, $\partial_\mu j^{\mu 5} = 0$ (cf. [**129**, Section 3.4]).

Prior to 1956, it was taken for granted that the laws of physics are invariant with respect to parity transformations in the sense that the mirror image of any physical process also represents a perfectly possible physical process. However, as experiments carried out by Chien-Shiung Wu (1912–1997) revealed, weak interactions are *not* invariant under parity [**163**]. Accordingly, neutrinos observed in nature are *left-handed*, whereas the observed anti-neutrinos are *right-handed* (see [**81**]). These properties are summarized by the notion of *chirality*. For a mathematical description of chirality (cf. [**53**, Section 1.2] and [**96**, Chapter 5]), one introduces so-called *chiral projectors*,

$$\chi_{L/R} = \frac{1 \mp \gamma^5}{2}\,.$$

These are projection operators, i.e. $\chi_{L/R}^2 = \chi_{L/R}$, satisfying the properties

$$\chi_L\chi_R = 0\,, \qquad \chi_L + \chi_R = \mathbb{1}\,, \qquad \gamma^5\,\chi_L = -\chi_L\,, \qquad \gamma^5\chi_R = \chi_R\,, \qquad \chi_L^* = \chi_R$$

(with respect to the spin scalar product $\overline{\psi}\phi$). For any spinor ψ, the projections $\chi_L\psi$ and $\chi_R\psi$ are referred to as *left-* and *right-handed* components of ψ, respectively. We point out that a matrix is said to be *even* or *odd* if it commutes or anti-commutes with γ^5, respectively. It is straightforward to verify that the Dirac matrices are odd,

$$\gamma^\mu\chi_{L/R} = \chi_{R/L}\gamma^\mu \qquad \text{for all } \mu = 0, \ldots, 3\,.$$

Introducing $\psi_{L/R} = \chi_{L/R}\psi$, this gives rise to the relation

$$\overline{\psi}\gamma^\mu\psi = \overline{\psi}_L\gamma^\mu\psi_L + \overline{\psi}_R\gamma^\mu\psi_R\,. \tag{2.3.21}$$

The summands in (2.3.21) are called *left-* and *right-handed currents*, respectively.[12]

Experiments have shown that neutrinos may convert from one flavor to another (for instance, $\nu_e \leftrightarrow \nu_\mu$). These phenomena, referred to as *neutrino oscillations*, suggest that neutrinos have non-zero mass, which in turn motivates to also take *right*-handed neutrinos into consideration [**81, 131**].[13] The neutrino oscillations can mathematically be described by the MNS matrix (or *neutrino mixing matrix*, cf. [**81**, Section 11.5]). For quarks, a similar effect is described by the CKM matrix (or *quark mixing matrix*, see [**81**, Section 9.5] or [**131**, Chapter 7] for details). We point out that currently there is no experimental reason for not suspecting the quarks and leptons in Table (2.3.20) to be the ultimate elementary particles [**131**]; in particular, it is unlikely to expect a fourth generation of fermions [**11**].

[12] For the relation to weak isospin and hypercharge we refer to [**81**, Chapter 9].

[13] These experimental observations are taken into account in the theory of causal fermion systems.

2.3.8. The Standard Model of Particle Physics. The quarks and leptons, which are the basic particles of matter, are subject to several interactions (or "forces"). These forces between fundamental fermions are mediated via the exchange of bosons (so-called "gauge bosons"). At present, there are four known fundamental interactions, comprising electromagnetic, weak and strong interaction as well as gravity. The first three of them are described by the *Standard Model of particle physics*. More explicitly, electromagnetic and weak interaction are combined into electroweak theory, which is dealt with in quantum electrodynamics (QED), whereas the theory of strong interaction is quantum chromodynamics (QCD). The gauge bosons of electroweak theory are the photon and the $W^{\pm}$ and Z bosons, while the gauge bosons of the strong force are (eight) gluons. In order to understand the number of gluons as well as its mathematical description in the Standard Model in some more detail, let us point out that modern theory of particle interactions is based on gauge theories [**127**]. Without describing its precise structure we note that, in mathematical terms, gauge theory deals with principal bundles, connections on them, and the curvatures of these connections [**121, 170**].

Before entering the basic idea of gauge theories, let us briefly outline some central structures of quantum field theory. In analogy to classical physics, it has become conventional to formulate particle physics in terms of a Lagrangian. More precisely, one usually introduces an *action* S by

$$S = \int L\, dt \qquad \text{with} \qquad L = \int \mathcal{L}\, d^3x\,,$$

where L is referred to as *Lagrangian* and $\mathcal{L}$ as *Lagrangian density* (which often is also called "Lagrangian"). In this formalism, the dynamics is determined by a single function $\mathcal{L}$; whenever the Lagrangian $\mathcal{L}$ is Lorentz invariant (that is, invariant under Lorentz transformations), the whole theory is Lorentz invariant [**96**]. For instance, the Dirac equation can be derived from the Lagrangian

$$\mathcal{L} = i\overline{\psi}\partial\!\!\!/\psi - m\overline{\psi}\psi\,,$$

where m denotes mass and $\overline{\psi}$ the adjoint spinor; this Lagrangian is invariant under the transformations $\psi \mapsto e^{i\theta}\psi$ for $\theta \in \mathbb{R}$ (cf. [**73**, Section 4.1]). Varying the action with respect to $\overline{\psi}$, treated as independent of ψ, one obtains the Dirac equation (2.3.17) (cf. [**128**, Section 1.5]).

The modern viewpoint of gauge invariance is due to quantum theory [**96**]. Since observables depend on $|\psi|^2$, the structure of the theory is invariant under *global gauge transformations* (cf. [**96**, eq. (3.6)])

$$\psi \to \psi' = e^{-i\theta}\psi\,,$$

where $\theta \in \mathbb{R}$ is a constant. In case that $\theta = \theta(t, \mathbf{x})$ depends on each spacetime point, the corresponding transformation (cf. [**96**, eq. (3.7)])

$$\psi(t, \mathbf{x}) \to \psi'(t, \mathbf{x}) = e^{-i\theta(t,\mathbf{x})}\psi(t, \mathbf{x})$$

is called a *local gauge transformation* (or, in this case, a *local phase transformation*). Making use of the fact that group elements of the unitary group U(1) are expressed by $e^{i\theta}$ with $\theta \in \mathbb{R}$, the above (local) gauge transformations give rise to the "gauge" group U(1).

In order to prepare for Yang-Mills theory and to explain gauge theories in some more detail, let us return to electrodynamics, which is considered as both the simplest gauge

theory and the most familiar [**131**]. In classical electrodynamics, the fields $\mathbf{B}$ and $\mathbf{E}$ are related to the vector potential $\mathbf{A}$ by (see e.g. [**96**])

$$\mathbf{B} = \nabla \times \mathbf{A}\,, \qquad \mathbf{E} = -\nabla \phi - \partial_t \mathbf{A}\,. \tag{2.3.22}$$

Whenever Λ is an arbitrary (smooth) real-valued function, it is evident that the equations (2.3.22) remain unchanged with respect to the transformations

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla \Lambda\,, \qquad \phi \to \phi' = \phi - \partial_t \Lambda\,. \tag{2.3.23}$$

Combining $\mathbf{A}$ and ϕ into a four-vector $A^\mu = (\phi, \mathbf{A})$, the transformations (2.3.23) read

$$A^\mu \to A^{\mu\prime} = A^\mu - \partial^\mu \Lambda\,.$$

In other words, the electromagnetic potential A_μ is not uniquely defined but only modulo the gauge transformations $A_\mu \mapsto A_\mu + \partial_\mu \Lambda$, where Λ is an arbitrary (smooth) real-valued function. The resulting freedom to choose A_μ is called *gauge invariance* in electrodynamics (see [**73**, Section 4.2] and [**131**, Chapter 3]). Concerning the notion of "gauge invariance" we refer to [**142**] and [**73**].

Let us also note the following invariance property: Whenever ψ and A_μ satisfy the Dirac equation $(i\partial\!\!\!/ - eA\!\!\!/ - m)\psi = 0$, then $\psi' = \exp(ie\Lambda)\psi$ and $A'_\mu = A_\mu + \partial_\mu \Lambda$ satisfy the equation $(i\partial\!\!\!/ - eA\!\!\!/' - m)\psi' = 0$. In other words, the Dirac equation $(i\partial\!\!\!/ - eA\!\!\!/ - m)\psi = 0$ is invariant under the simultaneous transformations $\psi \to \exp(ie\Lambda)\psi$, $A_\mu \to A_\mu + \partial_\mu \Lambda$ (cf. [**73**, Section 4.2 and Section 9.1]), thus combining the two gauge transformations mentioned before. Accordingly, we obtain a gauge-invariant Lagrangian

$$\mathcal{L} = \overline{\psi}(i\partial_\mu \gamma^\mu - A_\mu \gamma^\mu - m)\psi\,;$$

its physical interpretation is that the matter field ψ is coupled to a "gauge field" A_μ.

On the other hand, starting with the free Dirac Lagrangian $\mathcal{L}_0 = \overline{\psi}(i\partial\!\!\!/ - m)\psi$, one observes that $\mathcal{L}_0$ is invariant under the transformation $\psi \mapsto e^{i\chi}\psi$ in case that χ is a (real) constant. The natural question to ask is how the Lagrangian $\mathcal{L}_0$ needs to be modified in order to be invariant under the transformation $\psi \mapsto e^{i\chi}\psi$ for an *arbitrary* smooth function $\chi : \mathbb{R}^4 \to \mathbb{R}$. Namely, in the case that χ is *not* constant, the derivative ∂_μ does not commute with multiplication by $e^{i\chi}$. Indeed, one needs to replace the derivative ∂_μ by a "covariant derivative" $D_\mu = \partial_\mu + iA_\mu$ for some gauge field A_μ, thus giving rise to the "interacting" Lagrangian $\mathcal{L} = \overline{\psi}(i\partial\!\!\!/ - A\!\!\!/ - m)\psi$.[14] The important point to notice is that the gauge field A_μ appearing in electrodynamics is closely related to the local gauge transformations $\psi \to e^{i\chi}\psi$ with $e^{i\chi}$ in the Lie group $G = \mathrm{U}(1)$; more precisely, the gauge field A_μ takes values in the corresponding Lie algebra $\mathfrak{g} = \mathfrak{u}(1)$ (cf. [**73**, Section 9.1]).

Thus the transition from a global symmetry (constant χ) to a local symmetry which depends on spacetime coordinates (arbitrary χ) requires to introduce "compensating gauge fields" by means of the corresponding covariant derivative (for further details see [**39**, Section 18-6]); the gauge field is interpreted as interacting in a specific way, thus giving rise to a dynamical theory. These types of dynamical theories, based on local invariance principles, are called gauge theories [**2**].

By contrast to electrodynamics, in which case the gauge group U(1) is *abelian*, quantum chromodynamics and the electroweak theory are built on a generalization of this gauge principle, in which case the "phase factors" become matrices, which in general do not commute with each other. Accordingly, the associated symmetry is called a *non-abelian* one. The transition from a global symmetry to a local one in the non-abelian

[14] Replacing the derivative ∂ by the covariant derivative D is known as "minimal coupling" [**129**].

setting was first accomplished by Yang and Mills [**164**] in order to describe the weak interaction between quarks and leptons. More specifically, to obtain a local symmetry from the global symmetry of isospin invariance, Yang and Mills considered gauge transformations $\psi \to U\psi$ with $U \in \mathrm{SU}(2)$. The corresponding Lie algebra $\mathfrak{su}(2)$ has three generators; the corresponding $\mathfrak{su}(2)$-valued gauge fields A^a_μ $(a = 1, 2, 3)$ are referred to as *Yang-Mills fields*, and the resulting Euler-Lagrange equations are called *Yang-Mills equations* [**73**]. In a similar fashion, the strong interaction between quarks is described by quantum chromodynamics, a non-abelian gauge theory of the Yang-Mills type corresponding to the gauge group SU(3); the gauge potentials take values in the gauge algebra $\mathfrak{su}(3)$ (cf. [**3, 73**]). For a concise summary of Yang-Mills theory we refer to [**129**, Chapter 15].

Thus in summary, the interactions in the Standard Model are described by Yang-Mills type equations [**2**]. The weak and electromagnetic theory usually are combined into *electroweak theory* with gauge group $\mathrm{U}(1) \times \mathrm{SU}(2)$; henceforth, the gauge group of the Standard Model is given by

$$\mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)\,.$$

Let us finally establish the connection to the number of gluons involved in a gauge theory. For given $n \in \mathbb{N}$, consider the Lie group $\mathrm{SU}(n)$ and let $\mathfrak{su}(n)$ be its Lie algebra. We point out that $\mathfrak{su}(n)$ has n^2-1 generators, denoted by T_a $(a = 1, \ldots, n^2-1)$. Considering gauge transformations $\psi \to \psi' = U\psi$, in order to obtain local gauge invariance one is led to introduce the covariant derivative

$$D_\mu = \partial_\mu + igT_aA^a_\mu\,,$$

where the "coupling constant" g determines the strength of the interaction and A^a_μ $(a = 1, \ldots, n^2-1)$ are the *gauge fields*, taking values in $\mathfrak{su}(n)$. The number of these fields is equal to the number of generators, and the particles generated by the gauge fields are called *gauge bosons* (for details see [**127**]).

2.3.9. Quantum Field Theory. In short, quantum field theory (QFT) arose from the necessity to combine special relativity with quantum mechanics. More precisely, the world of everyday life is governed by classical mechanics. For classical objects that travel very fast compared to the speed of light, the laws of classical mechanics are modified by special relativity. For objects which are small (compared to the size of atoms, roughly speaking), classical mechanics is superseded by quantum mechanics. In order to describe objects that are both fast *and* small, one requires a theory that incorporates relativity and quantum mechanics: quantum field theory [**81, 167**]. It is essential for understanding the current state of elementary particle physics [**129**]. An explanation for the necessity of the *field* viewpoint can be found in [**129**, Section 2.1].

In order to deal with electroweak and strong interactions, quantum field theory is subdivided into quantum electrodynamics (QED) and quantum chromodynamics (QCD), where QED deals with the interaction of electrically charged particles and QCD describes the interaction of quarks by the exchange of gluons [**80**].

In analogy to the Lagrangian formalism in classical mechanics (it is instructive to follow the explanations in [**15**]), quantum field theory is usually written in terms of an action S, being the time integral over some Lagrangian L, which in turn is given as the spatial integral over a Lagrangian density $\mathcal{L}$ (where $\mathcal{L}$ again is referred to as "Lagrangian"). More specifically, "quantization" in quantum mechanics amounts to replacing

the (generalized) coordinate q of a classical particle by a Hermitian operator and its conjugate momentum $p \equiv \partial L/\partial \dot{q}$ (where $\dot{q} = \partial q/\partial t$) by $-i\partial/\partial q$ in such a way that the commutator relation (cf. [**15**, eq. (11.6)])

$$[p, q] = -i$$

is satisfied. More generally, given a system with n degrees of freedom, the dynamics of a quantum mechanical system is determined by imposing the commutator relations

$$[p_i(0), q_j(0)] = -i\delta_{ij} \,, \qquad [p_i(0), p_j(0)] = 0 \,, \qquad [q_i(0), q_j(0)] = 0$$

for all $i, j \in \{1, \ldots, n\}$ at time $t = 0$ (cf. [**15**, Section 11.2]). This method, treating q_i and p_i as quantized variables, is referred to as *first quantization*. In quantum field theory, on the other hand, one is mainly interested in quantizing *fields* which have an *infinite* number of degrees of freedom (corresponding to the limit $n \to \infty$). This process is called *second quantization* [**94, 129**]. For an example we refer to the quantization of the Klein-Gordon field ϕ (see [**129**, Section 2.3]). In general, the idea is to start with a classical field theory and then "quantize" it, that is, reinterpret dynamical variables as operators which obey canonical commutation relations. This procedure, describing the transition from classical field theory to quantum field theory, is known as "second quantization" in order to distinguish the resulting field equation (in which ϕ is an operator) from the old one-particle field equation (in which ϕ was a wave function). In analogy to classical mechanics, the action in QFT is given as an integral over a Lagrangian $\mathcal{L}(\phi, \partial_\mu \phi)$ depending on some field ϕ and its derivative $\partial_\mu \phi$,

$$S = \int L \, dt \qquad \text{with} \qquad L = \int \mathcal{L}(\phi, \partial_\mu \phi) \, d^3x \,,$$

where L is integrated between initial and final times [**94**].

For convenience, let us review some important QFT Lagrangians. To begin with, the Maxwell equations $\partial^\mu F_{\mu\nu} = 0$ can be obtained from the Lagrangian

$$\mathcal{L}_{\text{Maxwell}} = -\frac{1}{4}(F_{\mu\nu})^2 \,,$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor and A^μ is the vector potential (see [**129**, eq. (3.7)]). Moreover, the Lorentz-invariant Dirac Lagrangian reads

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi \,.$$

Next, the QED Lagrangian $\mathcal{L}_{\text{QED}}$, given by

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{Maxwell}} + \mathcal{L}_{\text{int}}$$

(with $\mathcal{L}_{\text{int}}$ describing the interaction), can be written in the simpler form

$$\mathcal{L}_{\text{QED}} = \overline{\psi}(i\not{D} - m)\psi - \frac{1}{4}(F_{\mu\nu})^2 \,,$$

where D_μ is the *gauge covariant derivative*,

$$D_\mu \equiv \partial_\mu + ieA_\mu(x)$$

with electron charge e (in order to describe a fermion of charge Q, replace e by Q). Varying $\overline{\psi}$, the resulting Euler-Lagrange equations for ψ take the form

$$(i\not{\partial} + ie\not{A} - m)\psi(x) = (i\not{D} - m)\psi(x) = 0 \,,$$

which is just the Dirac equation coupled to the electromagnetic field (with coupling constant e). The Euler-Lagrange equations for A_ν are given by

$$\partial_\mu F^{\mu\nu} = e\overline{\psi}\gamma^\nu\psi = ej^\nu \,,$$

which are the inhomogeneous Maxwell equations with the current density $j^\nu = \overline{\psi}\gamma^\nu\psi$ (see [**129**, eqs. (4.3)–(4.8)]). The QCD Lagrangian, on the other hand, reads

$$\mathcal{L}_{\mathrm{QCD}} = \sum_f \overline{\psi}^f \big(i\partial\!\!\!/ - \alpha A\!\!\!/ - m^f\big)\psi^f - \frac{1}{4}(F_{\mu\nu})^2$$

(cf. [**73**, Section 9.2]), where $f \in \{u, d, s, c, t, b\}$ are the quark flavors, m^f are the quark masses and α is the strong coupling constant (the same for all quark flavors). For the Yang-Mills Lagrangian, which can be obtained by adding the gauge field Lagrangian to the usual Dirac Lagrangian (with the ordinary derivative replaced by the covariant derivative), we refer to [**129**, Section 15.2].

Note that the symmetry of a Lagrangian may be violated, in which case the field theory is said to have a *hidden* or *spontaneously broken* symmetry. One important consequence of spontaneous symmetry breaking is *Goldstone's theorem* which states that for every spontaneously broken continuous theory there is a massless particle (known as *Goldstone boson*, see [**129**, Section 11.1]). Applying Goldstone's theorem to gauge theories, however, one finds that these goldstone bosons convert massless gauge bosons into massive ones. This is the so-called "Higgs mechanism" (we refer the interested reader to [**94**, Section 10.2] and [**129**, Section 20.1]).

Applying the effect of spontaneous symmetry breaking to the gauge theory of weak interactions gives rise to the so-called *GWS model*, introduced by Glashow, Weinberg and Salam (see [**94**, Section 10.4] and [**129**, Section 20.2]). This model yields a unified description of weak and electromagnetic interactions. More explicitly, starting with a theory with SU(2) gauge symmetry, one ends up with one massless gauge boson (photon), whereas the remaining three gauge bosons ($W^\pm, Z$) acquire masses from the Higgs mechanism. In the GWS model, only the left-handed components of the quark and lepton fields couple to the W bosons. As far as weak interactions are concerned, the resulting electroweak theory thus arranges the quarks and leptons as follows,

$$\begin{pmatrix}\nu_{\mathrm{e}}\\ e\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L \qquad \text{and} \qquad \begin{pmatrix}u\\ d\end{pmatrix}_L, \quad \begin{pmatrix}c\\ s\end{pmatrix}_L, \quad \begin{pmatrix}t\\ b\end{pmatrix}_L .$$

These are all weak isospin doublets; the right-handed components are weak isospin singlets (cf. [**39**, Section 18-8] and [**129**, eq. (20.75)]). For this reason, physicists often denote the SU(2) gauge group of weak interactions by $\mathrm{SU}(2)_L$.

The Standard Model of particle physics outlined above evolved as a consequence of experiments and advances in theoretical physics in the last century. Nevertheless, the predictions of a physical theory have to fit experimental results. Concerning particle physics, so-called *scattering experiments* are of crucial importance. Roughly speaking, in those experiments a beam of "incoming" particles is aimed at a target, and thereby "scattered" into "outgoing" particles. A typical quantity measured in the laboratory is the *scattering cross section*, which can be thought of the effective "size" of each target particle as seen by an incoming beam. More precisely, the cross section is calculated in terms of the rate of collisions in a scattering experiment; it gives the probability that a collection of particles in some initial state will decay or scatter into another collection of particles in some final state. For a mathematical description of such processes, the

incoming particles are assigned an *i*nitial state $|i\rangle$ at initial time $t = -\infty$, whereas the outgoing particles are in some *f*inal state $|f\rangle$ at time $t = \infty$. In order to calculate the probability of the transition from the initial state to the final state, one introduces the so-called "scattering matrix" S (cf. [**94**]). Then the transition from an initial state $|i\rangle$ to a final state $|f\rangle$ is given by the probability amplitude

$$S_{fi} = \langle f \mid i \rangle \,,$$

which can be thought of as one entry of the S matrix (see [**15**, Chapter 16]). Thus the task is to compute the S matrix for a scattering process. To this end, we recall that in quantum mechanics, a state ψ is a (normed) vector in a Hilbert space $\mathcal{H}$. Following the axioms of quantum mechanics, the time evolution of a system is given by a one-parameter group of unitary operators $U(t)$ on Hilbert space $\mathcal{H}$ such that

$$\psi(t) = U(t) \psi(0) \,,$$

where $\psi(t)$ denotes the state at time t. By Stone's theorem (see [**31**, Theorem 1.13]), the time evolution $U(t)$ is generated by a self-adjoint operator H,

$$U(t) = e^{-itH} \,.$$

The operator H, called the *Hamiltonian*, corresponds to the energy of the system (cf. [**31**, Section 3.1]). In order to calculate the S matrix by means of perturbation theory (for details see [**156**, Section 3.5] and [**73**, Chapter 6]), one assumes that the Hamiltonian H splits into a free-particle Hamiltonian H_0 and an interaction term V,

$$H = H_0 + V \,.$$

Introducing the operator

$$U(\tau, \tau_0) \equiv \exp(i H_0 \tau) \exp(-i H (\tau - \tau_0)) \exp(-i H_0 \tau_0)$$

for all $\tau, \tau_0 \in \mathbb{R}$, the S matrix is given by

$$S = U(+\infty, -\infty) \,. \tag{2.3.24}$$

Starting from (2.3.24) and defining

$$V(t) \equiv \exp(H_0 t) V \exp(-i H_0 t)$$

for all $t \in \mathbb{R}$, one obtains the following perturbation expansion for S,

$$\begin{aligned} S = \mathbb{1} - i \int_{-\infty}^{\infty} V(t_1)\, dt_1 + (-i)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{t_1} V(t_1) V(t_2)\, dt_2\, dt_1 \\ + (-i)^3 \int_{-\infty}^{\infty} \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} V(t_1) V(t_2) V(t_3)\, dt_3\, dt_2\, dt_1 + \cdots \,. \end{aligned} \tag{2.3.25}$$

In terms of the *time-ordered product* T (cf. [**156**, Section 3.5]), expression (2.3.25) reads

$$S = \mathbb{1} + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} T \left\{ V(t_1) \cdots V(t_n) \right\} \, dt_1\, dt_2 \cdots\, dt_n \,. \tag{2.3.26}$$

This expression, known as *Dyson series*, is also written as *time-ordered exponential*,

$$S = T \exp \left(-i \int_{\infty}^{\infty} V(t)\, dt \right) \,.$$

Making use of *Wick's theorem*, each time-ordered summand of the perturbation expansion (2.3.26) can be written as a sum of *normal-ordered* summands (cf. [**129**, Section 4.3] or [**94**, Section 5.6]). The resulting summands may be reexpressed diagrammatically in

terms of so-called *Feynman diagrams*. This yields a convenient method for evaluating the S matrix to a given order by employing the corresponding *Feynman rules*. In most calculations, it is preferable to express the Feynman rules in terms of momenta, thus giving rise to *momentum-space Feynman rules* (for details see [**129**]).

One of the serious complications found in quantum field theory is that the theory is naively divergent. When higher-order corrections are calculated for QED, one finds that the integrals diverge in the ultraviolet region, that is, for large momentum p. These divergences, which are related to the experimentally not accessible high-energy region, reflect our ignorance about the nature of physics at extremely small distances. In order to avoid these divergences, one "cuts off" momentum space at some large momentum Λ, thereby obtaining finite quantities. This manipulation is known as "regularization." Afterwards one considers the limit $\Lambda \to \infty$. A theory is said to be *renormalizable* if the physical quantities turn out to be independent of Λ (cf. [**129**]). For instance, it was proven by GERARD 'T HOOFT (1946–) that spontaneously broken Yang-Mills theory is renormalizable [**94**]. Following [**167**], quantum field theory thus should be regarded "as an effective low energy theory. (...) It is thought that as we go to higher and higher energies the whole edifice of quantum field theory will turn out to be an approximation to a theory whose identity we don't know yet."

Many different renormalization procedures have been proposed, but they all share some basic physical features. The essential idea is to assume a set of "bare" or "naked" parameters that are divergent, such as the coupling constants and masses.[15] By contrast to "physical" parameters, these bare parameters are unmeasurable. The divergences of these parameters are chosen so that they cancel against the ultraviolet infinities which probe the small-distance behavior of the theory. After these divergences have been absorbed by the bare parameters, we are left with the physical, renormalized, or "dressed" parameters that are indeed measurable [**94**]. We point out that, motivated by the renormalization program, the concept of bare parameters will show up in the theory of causal fermion systems.

Note that most conventional regularization schemes are based on the perturbative Feynman expansion. A non-perturbative approach is to work on spacetime lattices instead of a spacetime continuum [**3, 25**]. In this setting, one considers a discrete set of lattice points, separated by a minimum distance: the lattice spacing a. Regarding the lattice merely as an ultraviolet cutoff, when removing the regularization observable quantities should approach their physical values in the limit $a \searrow 0$, the so-called "continuum limit" [**77**]. These explanations may serve as a motivation for the notion of "continuum limit" which we encounter in the theory of causal fermion systems.

As being of relevance for the theory of causal fermion systems, we finally outline the generalization of the Dirac equation to curved spacetime, modelled by a (globally hyperbolic) pseudo-Riemannian manifold (M, g) (for details see [**128**]). In Minkowski spacetime $\mathscr{M}$ (with Minkowski metric $\eta_{\mu\nu}$), the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0$$

is obtained from the Lagrangian

$$\mathcal{L} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi \,,$$

[15] Concerning the bare and effective masses of quarks we refer to [**81**, Table 4.4].

and the γ-matrices satisfy the anti-commutation relations

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \,. \tag{2.3.27}$$

In curved spacetime, on the other hand, formula (2.3.27) is generalized to

$$\underline{\gamma}^\mu(x)\underline{\gamma}^\nu(x) + \underline{\gamma}^\nu(x)\underline{\gamma}^\mu(x) = 2g^{\mu\nu}(x) \,, \tag{2.3.28}$$

where $g^{\mu\nu}$ is the inverse of $g_{\mu\nu}$; the underline is used to distinguish the spacetime-dependent $\underline{\gamma}$-matrices from the constant γ-matrices in (2.3.27). In order to represent the matrices $\underline{\gamma}_\mu(x)$ in terms of the matrices γ_μ, one introduces a so-called *vierbein* $b^\alpha{}_\mu(x)$ of vector fields (or *tetrad*, see [**26**] for details), defined by

$$g_{\mu\nu}(x) = \eta_{\alpha\beta} b^\alpha{}_\mu(x) b^\beta{}_\nu(x) \,.$$

Vierbein indices are lowered with $\eta_{\alpha\beta}$, whereas spacetime indices are lowered with the metric $g_{\mu\nu}$. In terms of the Dirac matrices γ^μ in Minkowski space, the matrices $\underline{\gamma}^\mu(x)$ in (2.3.28) can be written in the form

$$\underline{\gamma}^\mu(x) = b_\alpha{}^\mu(x) \gamma^\alpha \,.$$

Next, the spinorial affine connections $\Gamma_\mu(x)$ (for details see [**74, 102, 112**]) are defined by the vanishing of the covariant derivative of the $\underline{\gamma}$-matrices,

$$\nabla_\mu \underline{\gamma}_\nu \equiv \partial_\mu \underline{\gamma}_\nu - \Gamma^\lambda{}_{\mu\nu} \underline{\gamma}_\lambda - \Gamma_\mu \underline{\gamma}_\nu + \underline{\gamma}_\nu \Gamma_\mu = 0 \,,$$

where $\underline{\gamma}_\mu = g_{\mu\nu} \underline{\gamma}^\nu$. Introducing the covariant derivative acting on a spinor field ψ by

$$\nabla_\mu \psi \equiv (\partial_\mu - \Gamma_\mu) \psi \,,$$

the generally covariant Dirac equation in curved spacetime reads

$$(i \underline{\gamma}^\mu(x) \nabla_\mu - m) \psi(x) = 0 \,.$$

2.4. The Principle of the Fermionic Projector

The theory of causal fermion systems, as outlined in [**59**], evolved from the principle of the fermionic projector [**53**]. For this reason, it seems a good starting point to first present the underlying ideas of the fermionic projector approach in more detail. To this end, we point out that the monograph [**53**] itself has its origins in [**48**], which in turn is based on previous considerations summarized in the articles [**49, 51**]. In [**49**] it was suggested to link the physical gauge principle with non-relativistic quantum mechanical measurements of the position variable. The basic idea in [**51**] is to extend this concept to relativistic quantum mechanics and to explain local gauge freedom for Dirac spinors by a local $\mathrm{U}(2,2)$ symmetry, which allows for a natural description of both electrodynamics and general relativity as a classical gauge theory [**51**].

Actually, the basic ideas in [**49, 51**] may be considered as the starting point for the development of the theory of causal fermion systems. For this reason, let us outline the underlying concepts, which are essential in order to understand how the specific form of the causal action principle comes about, in more detail. This procedure eventually leads to the principle of the fermionic projector, which in turn allows us to formulate a variational principle in "discrete" spacetime.

2.4.1. Derivation of Local Gauge Freedom. To begin with, let us recall that in quantum mechanics one usually considers an abstract Hilbert space H endowed with a scalar product $\langle . | . \rangle$. The physical observables in non-relativistic quantum mechanics correspond to self-adjoint operators on H with respect to the scalar product $\langle . | . \rangle$, and measurements with respect to a state $\psi \in H$ correspond to calculating the expectation value $\langle \psi | \mathcal{O} | \psi \rangle$. Moreover, the wave function $\psi(\vec{x})$ is introduced by

$$\psi(\vec{x}) = \langle \vec{x} \mid \psi \rangle \,. \tag{2.4.1}$$

It is of crucial importance to note that the observables for space are of particular interest as they determine the geometry of the physical system under consideration. Usually, they are given by mutually commuting operators $(X^i)_{i=1,\ldots,3}$ such that

$$X^i \psi(x) = x^i \psi(x) \qquad \text{for } i = 1, 2, 3 \,.$$

Introducing the wave function $\psi(\vec{x})$ by (2.4.1) is known as "position representation." In bra/ket notation,[16] this is done by choosing an "eigenvector basis" $|\vec{x}\rangle$ of the position operators in such a way that

$$\vec{X} |\vec{x}\rangle = \vec{x} |\vec{x}\rangle \,, \qquad \langle \vec{x} \mid \vec{y} \rangle = \delta^3(\vec{x} - \vec{y}) \,.$$

At this point it is central to observe that the position representation is not unique, as the "eigenvectors" $|\vec{x}\rangle$ are only determined up to a phase; more precisely, they can be transformed according to

$$|\vec{x}\rangle \to e^{ie\Lambda(\vec{x})} |\vec{x}\rangle \tag{2.4.2}$$

with a real function $\Lambda(\vec{x})$. This corresponds to a local phase transformation

$$\psi(\vec{x}) \to e^{-ie\Lambda(\vec{x})} \psi(\vec{x}) \tag{2.4.3}$$

of the wave functions. This arbitrariness of the local phase of the wave functions can also be understood from the fact that the wave function itself is not an observable, but only its absolute square $|\psi|^2$ has a physical interpretation as probability density. Starting point in [**51**] is to interpret the local phase transformations (2.4.2), (2.4.3) as U(1) gauge transformations. Generalizing the previous considerations to wave functions with several components, for suitable unitary matrices $U(\vec{x})$ one arrives at

$$\psi(\vec{x}) \to U(\vec{x}) \psi(\vec{x}) \,.$$

If these local transformations $U(\vec{x})$ could be identified with physical gauge transformations, then the local gauge principle would no longer be an a-priori principle in physics. Instead, it would be a consequence of a quantum mechanical "measurement principle," namely the description of space with observables X^i on an abstract Hilbert space. This idea might make it possible to describe additional interactions (like gravitation or the weak and strong forces). Furthermore, the local gauge group could no longer be chosen arbitrarily. These observations may be regarded as the starting point for the development of causal fermion systems.

In order to generalize this "measurement principle" to the relativistic setting, it is convenient to introduce "observables" for space and time as multiplication operators with

[16] Given a Hilbert space H with scalar product $\langle . | . \rangle$, in physics it is customary to "split up" the inner product $\langle . | . \rangle$ into a *ket* $| . \rangle$ and a *bra* $\langle . |$ (this terminology is due to [**33**]). Note that the formal bra/ket notation can be made mathematically precise using spectral measures [**49**].

the coordinate functions by

$$X^i \psi(x) = x^i \psi(x) \qquad \text{for all } i = 0, \ldots, 3\,,$$

where $\psi : \mathscr{M} \to \mathbb{C}^4$ is regarded as a Dirac spinor on Minkowski space $\mathscr{M}$. Motivated by the "spin scalar product" $\overline{\psi}\phi$ of quantum mechanics, we let $\mathbb{C}^4$ be endowed with an indefinite inner product $\prec . \mid . \succ$ of signature (2,2). Unfortunately, the usual positive scalar product (2.3.18),

$$(\psi \mid \phi) = \int_{\mathbb{R}^3} \psi^\dagger(t, \vec{x})\, \phi(t, \vec{x})\, d\vec{x}\,, \tag{2.4.4}$$

where the spinors are integrated over a space-like hypersurface at a constant time t, is not compatible with time measurement. More precisely, the expectation value

$$(\psi \mid X^0 \mid \psi) = t(\psi \mid \psi)$$

depends on the choice of the hypersurface, whereas for solutions of the Dirac equation, current conservation implies that (2.4.4) is independent of t. This motivates to introduce a different scalar product where the spinors are also integrated over the time variable,

$$< \psi \mid \phi > = \int_{\mathbb{R}^4} \prec \psi(x) \mid \phi(x) \succ \; d^4x\,. \tag{2.4.5}$$

In contrast to (2.4.4), however, the inner product (2.4.5) has no immediate physical interpretation. Introducing an "eigenvector basis" of the time and position operators with respect to $< . \mid . >$, one again obtains local gauge freedom of the form

$$\psi(x) \to U(x)\psi(x) \qquad \text{with } U(x) \in \mathrm{U}(2,2)\,, \tag{2.4.6}$$

where $\mathrm{U}(2,2)$ denotes the set of unitary matrices on $\mathbb{C}^4$ with respect to $\prec . \mid . \succ$.

In order to clarify the connection of these "gauge transformations" to interactions between fermions, it is illustrative to consider the free Dirac equation (cf. (2.3.17))

$$(i\partial\!\!\!/ - m)\psi = 0$$

with $\partial\!\!\!/ = \gamma^j \partial_j$, where γ^j $(j = 0, \ldots, 3)$ are the usual Dirac matrices. Due to (2.4.6), we are given the freedom to modify ψ according to $\psi \to U^{-1}\psi$ with $U \in \mathrm{U}(2,2)$. Multiplying by U from the left, this leads us to consider the modified Dirac equation

$$(U(i\partial\!\!\!/)U^{-1} - m)\psi = 0\,.$$

Thus the gauge transformation (2.4.6) yields a transformation of the Dirac operator,

$$i\partial\!\!\!/ \to G := U(i\partial\!\!\!/)U^{-1} = iG^j \frac{\partial}{\partial x^j} + B$$

with

$$G^j(x) = U(x)\gamma^j U(x)^{-1}\,, \qquad B(x) = iU(x)\gamma^j(\partial_j U(x)^{-1})\,.$$

The resulting operator G is Hermitian with respect to the inner product $< . \mid . >$ and Lorentz invariant. Furthermore, there exist gauge transformations such that G locally coincides with the original Dirac operator $i\partial\!\!\!/$ (see [**51**]).

In order to motivate the generalization to curved spacetime, we point out that the Minkowski metric η can be derived from the Dirac matrices γ^j by $2\eta^{ij}\mathbb{1} = \{\gamma^i, \gamma^j\}$ for all $i, j \in \{0, \ldots, 3\}$. For the generalization to curved spacetime, one replaces Minkowski space $\mathscr{M}$ by a four-dimensional smooth manifold M. Given an abstract Hilbert space H

endowed with an indefinite inner product $< . \,|\, . >$, the previous observations motivate to introduce the Dirac operator as a first order differential operator G on H by

$$G = iG^j \frac{\partial}{\partial x^j} + B$$

with (4×4)-matrices $G^j(x), B(x)$ in such a way that G^j locally coincides with γ^j for a specific gauge in a specific chart. The time and position operators, which are defined by means of spectral measures $(dE_x)_{x \in M}$, admit to introduce wave functions. Moreover, the Lorentzian metric g on M is given by

$$g^{jk}(x) \mathbb{1} = \frac{1}{2} \left\{ G^j(x), G^k(x) \right\} .$$

As a consequence, the Lorentzian metric allows us to construct further objects like the Levi-Civita connection or curvature tensors. It should be noticed that neither $B(x)$ nor the $\mathrm{U}(2,2)$ gauge symmetry enter these constructions. Indeed, the matrix $B(x)$ can be used for describing additional interactions (like electromagnetism). In this way, the Dirac operator is considered as the basic object on the manifold; all required objects can be constructed from the Dirac operator [**48**]. In particular, the gauge potentials are implicitly contained in the Dirac operator.

In short, the above constructions can be summarized as follows [**51**]: Adapting the ideas in [**49**] to the relativistic context yields a local $\mathrm{U}(2,2)$ gauge symmetry of the Dirac equation. In order to describe the physical interactions with this gauge symmetry, it is necessary to consider the Dirac operator as the basic object of the theory. By constructing the spin derivative (for details see [**51**]), the gauge potentials (which are implicitly contained in the Dirac operator) can be recovered as describing the electromagnetic and gravitational field. In this way, the concept of measurability of spacetime gives a fundamental explanation for local $\mathrm{U}(2,2)$ gauge symmetry. In particular, this description has the advantage that both the Dirac theory and classical field theory are developed from few a-priori given objects: The fermionic particles correspond to vectors of an abstract Hilbert space H endowed with an indefinite scalar product $< . \,|\, . >$, and spacetime is described by spectral measures $(dE_x)_{x \in M}$ on a manifold M. The Dirac operator yields the gauge potentials and determines the interaction between the fermions and the gauge fields. This description, which is conceptually simple, is the starting point for further constructions which eventually lead to the "principle of the fermionic projector" as introduced in [**53**] and finally give rise to causal fermion systems [**59**].

2.4.2. The Principle of the Fermionic Projector. After these preliminaries, we are in the position to enter the principle of the fermionic projector and to clarify its relation to the causal action principle. For simplicity, let $M = \mathscr{M}$ be Minkowski space. In generalization of [**51**], the starting point for the "discretization of spacetime" is an abstract Hilbert space H together with an indefinite inner product $< . \,|\, . >$ of signature $(2N, 2N)$ for some $N \in \mathbb{N}$. Following the explanations in [**53**, Section 3.1], each vector $\psi \in H$ gives rise to a wave function $\psi(x)$ with $x \in M$, and in analogy to §2.4.1 one obtains local gauge freedom of the form

$$\psi(x) \to U(x) \psi(x) \qquad \text{with } U(x) \in \mathrm{U}(2N, 2N) .$$

Now every particle is described by its wave function $\psi(x)$, or, in a gauge-independent way, by a vector $\psi \in H$. For this reason, it seems convenient to consider the complex subspace $< \psi >$ in H spanned by ψ in order to describe the corresponding particle. Accordingly, a finite number of particles $\psi_1, \ldots, \psi_f$ spans a finite-dimensional

complex subspace $Y := < \psi_1, \ldots, \psi_f >$ in H. One thus obtains an indefinite inner product space $(Y, < . \,|\, . >)$. Just as in positive definite scalar product spaces, every non-degenerate subspace $Y \subset H$ uniquely determines a projector $P_Y : H \to Y$ onto this subspace, characterized by the conditions $P_Y^2 = P_Y = P_Y^*$ and $P_Y(H) = Y$, where the star denotes the adjoint with respect to the inner product $< . \,|\, . >$ (for details see [**53**, Section 3.2]). Instead of working with the subspace Y, it is more convenient to consider the corresponding projector $P = P_Y$ in order to describe the particles of the system. Having Dirac particles in mind, which are fermions, the projector $P \in \mathrm{L}(H)$ is called *fermionic projector*, and

$$f = \dim P(H)$$

is referred to as the *number of particles*. Since P is a projector, one has $f = \mathrm{tr}(P)$, where tr denotes the trace of a linear operator (cf. [**56**, Section 3]).

Next, in order to present the underlying ideas for deriving a variational principle in discrete spacetime, we again let H be an abstract Hilbert space together with an indefinite inner product $< . \,|\, . >$ of signature $(2N, 2N)$. Then in analogy to [**51**], one defines mutually commuting time and position operators X^i $(i = 0, \ldots, 3)$. In order to establish the connection to discrete spacetime, one replaces the time and position operators X^i by mutually commuting operators X^i with a *purely discrete spectrum*. The joint spectrum of these operators is then regarded as discrete spacetime M, and the joint eigenspaces $(e_x)_{x \in M}$ are $4N$-dimensional subspaces of H. The corresponding projectors E_x on e_x for every $x \in M$ are uniquely characterized as the spectral projectors of the operators X^i. The projectors $(E_x)_{x \in M}$ are called *spacetime projectors*, and the resulting structure $(H, < . \,|\, . >, (E_x)_{x \in M})$ is referred to as *discrete spacetime* (for details see [**53**, Section 3.3] and [**54**]). Roughly speaking, the underlying idea is to think of discrete spacetime as a "loose set of points" [**48**] of "mean distance" $\varepsilon \lesssim \ell_P$ (for some "regularization length" $\varepsilon > 0$). In order to take the principle of general coordinate invariance appropriately into account, spacetime M should merely be considered as an index set for the spectral projectors, whereas the projectors $(E_x)_{x \in M}$ are regarded as the basic objects describing the geometry of spacetime [**53**].

The *principle of the fermionic projector* states that a physical system is completely described by the fermionic projector in discrete spacetime. In particular, the physical equations should be formulated exclusively with the operators P and $(E_x)_{x \in M}$ on H (see [**53**, Section 3.5]). Thus in order to formulate physical equations, the projectors P and $(E_x)_{x \in M}$ need to be combined in a mathematically interesting way. Guided by the Lagrangian formalism of quantum field theory, a promising strategy is to define an "action principle" in terms of a real-valued "Lagrangian." A possible strategy towards this aim is to form endomorphisms and then to consider the eigenvalues thereof. To this end, one introduces the "discrete kernel" of the fermionic projector $P(x, y)$ by $P(x, y) \equiv E_x P E_y$. Following the reasoning in [**48, 53**], it turns out that the so-called "closed chain" A_{xy} is of particular interest, which for all $x, y \in M$ is defined by

$$A_{xy} \equiv P(x, y) P(y, x) : E_x(H) \to E_x(H) \,.$$

Given a real-valued functional $\mathcal{L}$ on the endomorphisms of $E_x(H) \subset H$ (the so-called "Lagrangian"), the expression $\mathcal{L}[A_{xy}]$ depends on the two spacetime points $x, y \in M$. Summing over $x, y \in M$, the ansatz for the "action" is

$$S = \sum_{x, y \in M} \mathcal{L}[P(x, y) P(y, x)] \,.$$

An obvious way to form a *positive* functional is then to add up the absolute values of the eigenvalues of A_{xy} (counted with multiplicities). Denoting the eigenvalues of A_{xy} by λ_i, the resulting quantity is referred to as "spectral weight,"

$$|A| = \sum_{i=1}^{2N} |\lambda_i| \,.$$

It turns out that it is preferable to formulate a variational principle which aspires to equalize the absolute values of all eigenvalues. This can be accomplished by combining the expressions $|A^2|$ and $|A|^2$. More precisely, it is reasonable to minimize $|A^2|$, keeping $|A|^2$ fixed. This is the motivation for considering the variational principle

$$\text{minimize} \qquad S[P] = \sum_{x,y \in M} \left| A_{xy}^2 \right| \tag{2.4.7}$$

by varying P in the class of linear operators $\mathrm{L}(H)$ under the constraint

$$T[P] := \sum_{x,y \in M} |A_{xy}|^2 = C \tag{2.4.8}$$

for some constant $C > 0$. The corresponding *Euler-Lagrange equations* read

$$[P, Q] = 0$$

for some operator Q (cf. [**53**, eqs. (3.5.9)–(3.5.10) and (3.5.20)–(3.5.21)]). The variational principle (2.4.7)–(2.4.8) was first analyzed mathematically in [**54**]. In order to clarify its connection to the causal action principle in the theory of causal fermion systems, it is instructive to recall an existence result in [**54**]. More explicitly, in view of [**54**, Theorem 2.8] there exists a minimizer of the variational principle (2.4.7)–(2.4.8) in the class of operators $P \in \mathrm{L}(H)$ with the property that $\mathrm{tr}(P) = f$, $\mathrm{rank}(P) \leq f$ and $(-P)$ is positive in the sense that

$$< u \,|\, (-P) u > \geq 0 \qquad \text{for all } u \in H \,. \tag{2.4.9}$$

The fact that P has a negative definite image makes it possible to introduce a Hilbert space $(\mathcal{H}, \langle \,.\,|\,.\,\rangle_{\mathcal{H}})$ by setting $\langle \,.\,|\,.\,\rangle_{\mathcal{H}} = < \,.\,|\,(-P)\,.\,>$ and dividing out the null space. This construction, which was first outlined in [**61**], gives rise to an underlying Hilbert space structure. For similar constructions we refer to [**109**].

2.4.3. Connection to Causal Fermion Systems. According to [**53**, Preface to the second online edition], the step from indefinite inner product spaces to Hilbert spaces is regarded as "maybe the most important change in the mathematical setup." Actually, this transition from indefinite inner product spaces to Hilbert spaces is closely related to introducing a "regularization," one of the fundamental concepts of causal fermions systems. In order to outline this transition in a mathematical convincing way, let us explain the idea of introducing a "regularization" in some more detail. The following explanations are due to [**59**, Section 1.2] and [**53**, Section 1.5].

We let $\mathscr{M}$ be Minkowski space, endowed with the standard volume measure $d\mu$. We define the *spinor bundle* $S\mathscr{M}$ as a vector bundle over $\mathscr{M}$ with fiber $\mathbb{C}^4$. Sections in the spinor bundle are called *spinors* or wave functions. By $C(\mathscr{M}, S\mathscr{M})$ we denote the continuous Dirac wave functions (that is, the continuous sections of the spinor bundle,

not necessarily solutions of the Dirac equation). The spinor space at a point $x \in \mathscr{M}$ is denoted by $S_x\mathscr{M}$, so that a wave function ψ takes values in

$$\psi(x) \in S_x\mathscr{M} \simeq \mathbb{C}^4 .$$

The spinor space at x is endowed with an indefinite inner product of signature $(2,2)$, which we denote by $\prec . \,|\, . \succ$. We then consider solutions of the Dirac equation

$$(i\partial\!\!\!/ - m)\psi = 0 ,$$

where $\partial\!\!\!/ = \gamma^j \partial_j$. For a solution ψ, the function $(\overline{\psi}\gamma^0\psi)(t,\vec{x})$ has the interpretation as the probability density of the Dirac particle at time t to be at the position $\vec{x}$. The spatial integral of this probability density is time independent, and the corresponding bilinear form corresponding to this probability integral gives the scalar product (cf. (2.3.19))

$$(\psi \mid \phi) = \int_{\mathbb{R}^3} (\overline{\psi}\gamma^0\phi)(t,\vec{x})\, d^3x . \tag{2.4.10}$$

In order to ensure that the integral in (2.4.10) is well-defined and finite, we first consider solutions which at time t are smooth and have compact support. Taking the completion, the solution space becomes a separable Hilbert space (see e.g. [**72**, Proposition 8.17]). Next, we choose $\mathcal{H}$ as a closed subspace of this Hilbert space with the induced scalar product $\langle . \,|\, . \rangle_{\mathcal{H}} := (. \,|\, .)|_{\mathcal{H}\times\mathcal{H}}$. Then $\mathcal{H}$ is again a separable Hilbert space. However, as Dirac solutions $\psi, \phi \in \mathcal{H}$ are in general not continuous, the pointwise evaluation

$$- \prec \psi(x) \mid \phi(x) \succ \qquad \text{for all } \psi, \phi \in \mathcal{H}$$

at $x \in \mathscr{M}$ makes in general no mathematical sense. This is the reason for introducing an *ultraviolet regularization*. Such regularizations are performed most conveniently with so-called "regularization operators," which we now define. Consider a family of linear operators $(\mathfrak{R}_\varepsilon)_{\varepsilon\in(0,\varepsilon_{\max})}$ with $\varepsilon_{\max} > 0$ which map $\mathcal{H}$ to the continuous wave functions,

$$\mathfrak{R}_\varepsilon \,:\, \mathcal{H} \to C(\mathscr{M}, S\mathscr{M}) . \tag{2.4.11}$$

Whenever the technical conditions [**59**, eqs. (1.2.6)–(1.2.8)] in [**59**, Definition 1.2.3] hold, the family $(\mathfrak{R}_\varepsilon)_{\varepsilon\in(0,\varepsilon_{\max})}$ is said to be a family of *regularization operators*. A regularized wave function $\mathfrak{R}_\varepsilon\psi$, however, need not again be a solution of the Dirac equation. Given a family of regularization operators $(\mathfrak{R}_\varepsilon)_{\varepsilon\in(0,\varepsilon_{\max})}$ for some $\varepsilon_{\max} > 0$, the connection to causal fermion systems is accomplished as follows. Fixing $\varepsilon \in (0, \varepsilon_{\max})$, for any $x \in \mathscr{M}$ we consider the well-defined bilinear form

$$b_x^\varepsilon \,:\, \mathcal{H}\times\mathcal{H} \to \mathbb{C} , \qquad b_x^\varepsilon(u,v) = - \prec (\mathfrak{R}_\varepsilon\, u)(x) \mid (\mathfrak{R}_\varepsilon\, v)(x) \succ .$$

Making use of the Fréchet-Riesz theorem (see for instance [**113**, Section 6.3]), there is a unique bounded linear operator $F^\varepsilon(x) : \mathcal{H} \to \mathcal{H}$ such that

$$\langle u \mid F^\varepsilon(x)\, v\rangle_{\mathcal{H}} = b_x^\varepsilon(u,v) = - \prec (\mathfrak{R}_\varepsilon\, u)(x) \mid (\mathfrak{R}_\varepsilon\, v)(x) \succ \qquad \text{for all } u, v \in \mathcal{H} .$$

Taking into account that the inner product $\prec . \,|\, . \succ$ on the Dirac spinors at x has signature $(2,2)$, we infer that $F^\varepsilon(x)$ has at most 2 positive and at most 2 negative eigenvalues. Defining $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ as the set of self-adjoint operators with at most 2 positive and at most 2 negative eigenvalues, we thus obtain a mapping

$$F^\varepsilon : \mathscr{M} \to \mathcal{F} .$$

The operators $F^\varepsilon(x)$ are referred to as *local correlation operators* (as they furnish a connection between the underlying measure space and the structure $\mathcal{F}$ which plays a

crucial role in the setting of causal fermion systems). It is important to observe that on $\mathcal{F}$ we are given a measure $d\rho$ (referred to as "universal measure") defined by

$$d\rho^\varepsilon := F^\varepsilon_* \, d\mu$$

as the push-forward measure of $d\mu$ under F^ε (more precisely, for any Borel set $\Omega \subset \mathcal{F}$ we have $(F^\varepsilon)_* \mu(\Omega) := \mu((F^\varepsilon)^{-1}(\Omega))$). The resulting structure $(\mathcal{H}, \mathcal{F}, d\rho^\varepsilon)$ is a causal fermion system of spin dimension $n = 2$ (see Definition 2.1.1), and the support of the universal measure $d\rho^\varepsilon$ is referred to as *spacetime*. Note that in the context of causal fermion systems, the physical picture behind the ultraviolet regularization is that only the *regularized* objects are regarded as the fundamental physical objects. In particular, the regularization has a physical significance as describing the microscopic structure of physical spacetime. Removing the regularization by taking the limit $\varepsilon \searrow 0$ in a certain way (see [**59**, Section 3.5]) is referred to as "continuum limit."

In this way, according to [**53**, Preface to the second online edition], when working out the existence theory of the variational principle (2.4.7)–(2.4.8), "it became clear that instead of using the kernel of the fermionic projector, the causal action principle can be formulated equivalently in terms of the local correlation operators," thus giving rise to the causal action principle as stated in Section 2.1.

2.5. The Theory of Causal Fermion Systems

Having derived the specific form of the causal action principle as outlined in Section 2.1, it remains to clarify the connection of causal fermion systems to the Standard Model of particle physics as well as general relativity. This is precisely the aim of this section. In order to clarify the underlying ideas, it seems best to first outline the general strategy (§2.5.1).

2.5.1. The General Strategy. The objective of this subsection is to present the general strategy behind the theory of causal fermion systems. It is worth noting that the basic procedure outlined in this subsection is closely related to the ideas presented in §2.4.1. In short, the main ingredient of the theory can be regarded a class $(M, g, \mathcal{B})$ of four-dimensional, globally hyperbolic Lorentzian manifolds M (for details see [**10**]) endowed with Lorentzian metric g and an external potential $\mathcal{B}$ on M with appropriate properties. The underlying idea is that M models (curved) physical spacetime according to general relativity; the corresponding free Dirac equation reads (for details see §2.3.9)

$$(i \underline{\gamma}^\mu(x) \nabla_\mu - m) \psi(x) = 0 \tag{2.5.1}$$

with spacetime-dependent matrices $\underline{\gamma}^\mu$ and covariant derivative ∇_μ. Introducing an interaction $\mathcal{B}$, the resulting Dirac equation takes the form

$$(i \underline{\gamma}^\mu(x) \nabla_\mu + \mathcal{B}(x) - m) \psi(x) = 0 \,. \tag{2.5.2}$$

Roughly speaking, the Dirac equation (2.5.2) can be regarded as the starting point for the theory of causal fermion systems. Albeit the theory is worked out only in Minkowski space, the general strategy is based on the following steps: The starting point of the theory is to consider a solution of the unperturbed Dirac equation (2.5.1). The next step is to introduce a potential $\mathcal{B}$ and to make use of the so-called "causal perturbation expansion" in order to construct a solution ψ of the Dirac equation (2.5.2). Following Dirac's concept of describing the vacuum (in presence of an external potential $\mathcal{B}$) by means of a "completely filled Dirac sea," by introducing particles and anti-particles it is possible to modify the solution ψ of the Dirac equation (2.5.2). As a consequence, the

Dirac equation (2.5.2) will no longer be satisfied. Thus in order for (2.5.2) to hold again, it is necessary to also replace the potential $\mathcal{B}$ by some other potential $\tilde{\mathcal{B}}$. The resulting potential $\tilde{\mathcal{B}}$ is then interpreted as "interaction" between fermions.

In order to establish a connection to the Standard Model of particle physics, it is reasonable to extend the previous considerations by introducing one Dirac sea for each of the fundamental fermions compiled in Table (2.3.20). This is done by introducing the so-called "auxiliary fermionic projector," which is defined as the direct sum of the Dirac seas corresponding to the fundamental fermions. Employing the ansatz of "massive neutrinos," it is necessary to break the chiral symmetry; to this end, massless right-handed neutrinos are taken into account by adding one more direct summand to the auxiliary fermionic projector. In order to enter the setting of causal fermion systems, a "regularization" is inserted. Introducing particles and anti-particles, by forming the so-called "sectorial projection" one finally obtains the "fermionic projector" (for further details see [**59**, Remark 4.2.3]).

Motivated by the variational principle (2.4.7)–(2.4.8), the underlying picture is that the fermionic projector is supposed to be of physical relevance if it is a minimizer of the causal action principle. In this way, the resulting Euler-Lagrange (EL) equations give rise to conditions for the fermionic projector. It is crucial to observe that the fermionic projector obtained by the just-mentioned method essentially depends on the external potential $\mathcal{B}$ and the underlying spacetime (M, g). Accordingly, the EL equations pose conditions on "admissible" manifolds (M, g) and potentials $\mathcal{B}$. In this way, only those external potentials $\mathcal{B}$ and manifolds (M, g), for which the EL equations are satisfied, are considered as physically relevant; the resulting potentials are interpreted as "bosonic interactions" between fundamental fermions. In this way, the causal action principle allows to recover the interactions of the Standard Model and gives rise to the Einstein field equations of general relativity in a suitable limiting case (the so-called "continuum limit"). More precisely, the basic idea to obtain a connection to the Standard Model and general relativity is to rewrite the resulting EL equations as corresponding to an "effective action" $\mathcal{S}_{\text{eff}}$ of the form (cf. [**59**, eq. (5.4.1)])

$$\mathcal{S}_{\text{eff}} = \int_{\mathscr{M}} (\mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{EH}}) \sqrt{-\det g}\, d^4x$$

for a suitable choice of the Dirac Lagrangian $\mathcal{L}_{\text{Dirac}}$, the Yang-Mills Lagrangian $\mathcal{L}_{\text{YM}}$ and the Einstein-Hilbert Lagrangian $\mathcal{L}_{\text{EH}}$. Moreover, under appropriate assumptions one recovers the effective gauge group of the Standard Model,

$$\mathcal{G} = \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)\,.$$

The general strategy how to obtain these results shall be outlined in what follows.

2.5.2. The Kernel of the Fermionic Projector. Before entering the theory of causal fermion systems in more detail, let us briefly summarize some conceptual ideas which enter the theory from the very beginning. To begin with, taking into account that the causal action principle evolved from the principle of the fermionic projector [**53**], the following physical principles implicitly enter the theory of causal fermion systems: Einstein's principle of relativity, the Pauli exclusion principle and, of course, the principle of least action. Moreover, the principle of causality is built in in such a way that spacelike separated points do not enter the Lagrangian and thus the causal action. Moreover, the local gauge principle arises as explained in §2.4.1.

Next, only the fermionic states are considered as the basic physical objects. The bosonic fields, on the other hand, are merely regarded as auxiliary objects used for describing the behavior of the fermionic states. For this reason, the fundamental fermions summarized in Table (2.3.20) will play a central role in what follows. Motivated by the renormalization program, the concept of "bare" or "naked" masses and coupling constants is employed throughout; the "physical" masses which can be measured in experiments are obtained from the naked masses by self-interaction.

In order to take this fundamental character of fermions into account, the theory of causal fermion systems essentially employs the concept of a "Dirac sea" which was introduced by Paul Dirac in his paper [**34**] (cf. §2.3.5). Based on the Pauli exclusion principle, Dirac assumed that

> "(...) *all the states of negative energy are occupied except perhaps a few of small velocity.* (...) *Only the small departure from exact uniformity, brought about by some of the negative-energy states being unoccupied, can we hope to observe.* (...) We are therefore led to the assumption that *the holes in the distribution of negative-energy electrons are the [positrons].*"

Dirac made this picture precise in his paper [**35**] by introducing a relativistic density matrix $R(t, \vec{x}; t', \vec{x}')$ with $(t, \vec{x}), (t', \vec{x}') \in \mathbb{R} \times \mathbb{R}^3$ defined by

$$R(t, \vec{x}; t', \vec{x}') = \sum_{l \text{ occupied}} \Psi_l(t, \vec{x})\, \overline{\Psi_l(t', \vec{x}')}\,.$$

In analogy to Dirac's original idea, in [**50**] the kernel of the fermionic projector is introduced as the sum over all occupied wave functions

$$P(x, y) = -\sum_{l \text{ occupied}} \Psi_l(x)\, \overline{\Psi_l(y)}$$

for spacetime points $x, y \in \mathscr{M}$ (also see [**57**]). A straightforward calculation shows that (cf. [**60**, §4.1] and [**59**, Section 3.8]), up to an irrelevant constant, the relativistic density matrix $R(t, \vec{x}; t', \vec{x}')$ coincides precisely with kernel of the fermionic projector $P(x, y)$,

$$P(x, y) = \int_{\hat{\mathscr{M}}} \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)}\,, \tag{2.5.3}$$

where δ is Dirac's delta distribution and Θ denotes the Heaviside function, the slash in $\not{k} = k^j \gamma_j$ denotes contraction with the Dirac matrices γ_j $(j = 0, \ldots, 3)$ and the expression $k(x - y)$ is a short notation for the Minkowski inner product $k_j (x - y)^j$. The quantity $P(x, y)$ in (2.5.3) is called the (unregularized) *kernel of the fermionic projector of the vacuum.* We also refer to the expression (2.5.3) as a *completely filled Dirac sea*; it can be thought of as describing the "vacuum" with respect to *one* kind of elementary particles of mass m.

Note that particles are obtained by adding states to a completely filled Dirac sea, whereas anti-particles are described by removing states therefrom. This picture of a Dirac sea is made precise by considering the free Dirac equation in Minkowski space,

$$(i\partial\!\!\!/ - m)\psi = 0\,. \tag{2.5.4}$$

Assume that $\mathcal{H}$ is a separable Hilbert space of solutions of the Dirac equation (2.5.4), and let $(u_\ell)_{\ell \in \mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$. Dirac's original idea is then specified with

the help of regularization operators (2.4.11) by introducing the regularized kernel of the fermionic projector for all $x, y \in \mathscr{M}$ as (cf. [**59**, eq. (1.2.19)])

$$P^\varepsilon(x,y) = -\sum_\ell \big(\mathfrak{R}_\varepsilon u_\ell\big)(x)\, \overline{\big(\mathfrak{R}_\varepsilon u_\ell\big)(y)} = -\sum_\ell \big|\big(\mathfrak{R}_\varepsilon u_\ell\big)(x) \succ \prec \big(\mathfrak{R}_\varepsilon u_\ell\big)(y)\big| \,.$$

In the limit $\varepsilon \searrow 0$, the regularized kernel of the fermionic projector $P^\varepsilon(x,y)$ converges as a bi-distribution to the unregularized kernel $P(x,y)$ defined by

$$P(x,y) := -\sum_\ell u_\ell(x)\, \overline{u_\ell(y)} = -\sum_\ell |u_\ell(x) \succ \prec u_\ell(y)| \,, \tag{2.5.5}$$

as is made mathematically precise in [**59**, Proposition 1.2.7]. Furthermore, if $\mathcal{H}$ is the subspace of the solution space of the Dirac equation (2.5.4) which is spanned by all negative-frequency solutions, then the unregularized kernel of the fermionic projector as defined by (2.5.5) is the tempered bi-distribution (see [**59**, Lemma 1.2.8])

$$P(x,y) = \int \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k_0)\, e^{-ik(x-y)} \,.$$

In order to appreciate the relevance of the fermionic projector, it is instructive to outline the so-called "external field problem" which we recall now.

2.5.3. The External Field Problem. In order to understand the basic difficulty of introducing interactions into the Dirac equation in more detail, let us consider an interaction described by an external field $\mathcal{B}$ in Minkowski space. In this case, the Dirac wave functions are supposed to obey the modified Dirac equation

$$(i\partial\!\!\!/ + \mathcal{B} - m)\psi = 0 \,,$$

where $\mathcal{B}$ is assumed to be a smooth potential with suitable decay properties for large times and at spatial infinity (for details we refer to [**53**, Section 2.1] and [**59**, §2.1.2]). Whenever the potential $\mathcal{B}$ is *static* (that is, time-independent), one can separate the time-dependence of the wave function with a plane wave ansatz (cf. [**53**, eq. (2.1.2)]),

$$\psi(t, \vec{x}) = e^{-i\omega t} \psi_\omega(\vec{x}) \,, \tag{2.5.6}$$

where ω has the interpretation as the energy of the solution. In this way, the sign of ω gives rise to a *canonical splitting* of the solution space of the Dirac equation into subspaces of positive and negative energy, respectively. Unfortunately, the situation becomes much more difficult in the case that $\mathcal{B}$ is *time-dependent*. The basic difficulty in this setting is that the separation ansatz (2.5.6) no longer works. More precisely, one is no longer given a canonical splitting of the solution space of the Dirac equation. Therefore it is not clear which solutions can be interpreted as "negative-energy solutions" and thus correspond to the anti-particle states. This difficulty is referred to as *external field problem*. It is a common belief that in the presence of a general time-dependent external potential there exists no longer a *canonical* decomposition of the solution space into two subspaces, and the resulting arbitrariness in decomposing the solution space into two subspaces is often associated to an "observer." In this way, the interpretation of particles and anti-particles becomes *observer-dependent*. Nevertheless, one of the fundamental results in the theory of causal fermion systems is that *even in the presence of a time-dependent external potential there is a canonical decomposition of the solution space into two subspaces*. This method allows us to construct solutions of the modified Dirac equation (2.5.7) under suitable assumptions on $\mathcal{B}$, as shall be outlined in the following.

Starting point for the following constructions is a completely filled Dirac sea (2.5.3) of mass m (where the superscript "vac" clarifies Minkowski vacuum),

$$P_m^{\text{vac}}(x,y) = \int \frac{d^4k}{(2\pi)^4} \,(\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)}\,, \tag{2.5.7}$$

which is referred to as *kernel of the fermionic projector in the vacuum.* The basic idea in the theory of causal fermion systems in order to resolve the external field problem is to split up the fermionic projector of the vacuum in such a way that

$$P_m^{\text{vac}}(x,y) = \frac{1}{2}\big(p_m(x,y) - k_m(x,y)\big)\,, \tag{2.5.8}$$

where $p_m(x,y)$ and $k_m(x,y)$ are the distributions

$$p_m(x,y) = \int \frac{d^4k}{(2\pi)^4} (\not{k} + m)\, \delta(k^2 - m^2)\, e^{-ik(x-y)}\,,$$
$$k_m(x,y) = \int \frac{d^4k}{(2\pi)^4} (\not{k} + m)\, \delta(k^2 - m^2)\, \epsilon(k^0)\, e^{-ik(x-y)}\,,$$

and $\epsilon : \mathbb{R} \to \{-1, 1\}$ is the sign function defined by $\epsilon(x) = 1$ for $x \geq 0$ and $\epsilon(x) = -1$ otherwise. Remarkably, the distributions p_m and k_m can be rewritten in terms of Green's functions, as we now explain. For every $x \in \mathscr{M}$ let $J_x^\vee$ and $J_x^\wedge$ denote the points in the *causal future* respectively *past* of x,

$$J_x^\vee = \{y \in M \,|\, (y-x)^2 \geq 0,\ (y^0 - x^0) \geq 0\}\,,$$
$$J_x^\wedge = \{y \in M \,|\, (y-x)^2 \geq 0,\ (y^0 - x^0) \leq 0\}\,.$$

Denoting Dirac's delta distribution in Minkowski space by δ^4, due to [**59**, Chapter 2] the *advanced* and *retarded Green's functions* are uniquely defined by

$$(i\partial\!\!\!/ - m)\, s_m^\vee(x,y) = \delta^4(x-y) \qquad \text{with} \qquad \operatorname{supp} s_m^\vee(x,.) \subset J_x^\vee$$

and

$$(i\partial\!\!\!/ - m)\, s_m^\wedge(x,y) = \delta^4(x-y) \qquad \text{with} \qquad \operatorname{supp} s_m^\wedge(x,.) \subset J_x^\wedge\,,$$

respectively. Then the *causal fundamental solution* k_m can be recovered by

$$k_m(x,y) = \frac{1}{2\pi i}\big(s^\vee(x,y) - s^\wedge(x,y)\big)\,. \tag{2.5.9}$$

Introducing $J_x := J_x^\vee \cup J_x^\wedge$, the distribution k_m is causal in the sense that

$$\operatorname{supp} k_m(x,.) \subset J_x\,.$$

Making use of the identity (see [**59**, eq. (2.1.18) and Lemma 2.1.3])

$$k_m\, k_{m'} = \delta(m - m')\, p_m\,,$$

one can deduce p_m from k_m. In this way, the Green's functions $s_m^\vee$ and $s_m^\wedge$ allow us to introduce the fermionic projector of the vacuum P_m^{vac} by (2.5.8).

The idea is to extend these constructions to the interacting situation, that is, to the Dirac equation in presence of an external potential $\mathcal{B}$. According to [**59**, §2.1.4], the advanced and retarded Green's functions are uniquely defined even in the presence of an external potential (2.5.7),

$$(i\partial\!\!\!/ + \mathcal{B} - m)\psi(x) = 0\,. \tag{2.5.10}$$

More explicitly, the *advanced* and *retarded Green's functions* $\tilde{s}^\vee_m$ and $\tilde{s}^\wedge_m$ for the Dirac equation (2.5.10) are characterized by

$$\begin{aligned}
&(i\partial\!\!\!/ + \mathcal{B} - m)\, \tilde{s}^\vee_m(x,y) = \delta^4(x-y) \qquad \text{with} \qquad \operatorname{supp} \tilde{s}^\vee_m(x,.) \subset J^\vee_x \qquad \text{and} \\
&(i\partial\!\!\!/ + \mathcal{B} - m)\, \tilde{s}^\wedge_m(x,y) = \delta^4(x-y) \qquad \text{with} \qquad \operatorname{supp} \tilde{s}^\wedge_m(x,.) \subset J^\wedge_x\,,
\end{aligned}$$

respectively. The crucial point in what follows is that the Green's functions $\tilde{s}^\vee_m$ and $\tilde{s}^\wedge_m$ can be uniquely expressed in terms of $s^\vee_m$ and $s^\wedge_m$. More precisely, one has the following unique perturbation series (cf. [**59**, eq. (2.1.25)])

$$\tilde{s}^\vee_m = \sum_{n=0}^\infty \big(- s^\vee_m\, \mathcal{B} \big)^n\, s^\vee_m\,, \qquad \tilde{s}^\wedge_m = \sum_{n=0}^\infty \big(- s^\wedge_m\, \mathcal{B} \big)^n\, s^\wedge_m\,. \tag{2.5.11}$$

Its summands are referred to as "Feynman diagrams" (of n^{th} order).[17]

Having derived a perturbation series for the causal Green's functions, the *causal fundamental solution* is defined in generalization of (2.5.9) by

$$\tilde{k}_m := \frac{1}{2\pi i}(\tilde{s}^\vee_m - \tilde{s}^\wedge_m)\,.$$

Omitting the mass indices, for any complex number λ in the resolvent set of k and $\tilde{k}$, the unperturbed resolvent R_λ and the (perturbed) resolvent $\tilde{R}_\lambda$ are introduced by

$$R_\lambda := (k-\lambda)^{-1} \qquad \text{and} \qquad \tilde{R}_\lambda = \big(\tilde{k} - \lambda\big)^{-1}\,,$$

respectively (also see Appendix A.2). Rewriting the causal fundamental solution $\tilde{k}$ as

$$\tilde{k} = k + \Delta k\,,$$

the resolvent $\tilde{R}_\lambda$ can be written as a Neumann series (see e.g. [**157**, Satz II.1.11]),

$$\tilde{R}_\lambda = (k - \lambda + \Delta k)^{-1} = (1 + R_\lambda \cdot \Delta k)^{-1} \cdot R_\lambda = \sum_{n=0}^\infty (-R_\lambda \cdot \Delta k)^n \cdot R_\lambda\,.$$

Furthermore, the unperturbed resolvent can be expressed in terms of p and k,

$$R_\lambda = \frac{p+k}{2}\left(\frac{1}{1-\lambda}\right) + \frac{p-k}{2}\left(\frac{1}{-1-\lambda}\right) - \frac{\mathbb{1}-p}{\lambda}\,.$$

Choosing a contour $\Gamma_- \subset \mathbb{C}$ which encloses the point -1 in counter-clockwise direction and does not enclose the points 1 and 0, the fermionic projector P^{sea} is defined by

$$P^{\text{sea}} = -\frac{1}{2\pi i} \oint_{\Gamma_-} (-\lambda)\, \tilde{R}_\lambda\, d\lambda\,.$$

According to [**59**, Proposition 2.1.7], it has the desirable property that

$$(i\partial\!\!\!/ + \mathcal{B} - m)\, P^{\text{sea}} = 0\,. \tag{2.5.12}$$

Following the previous constructions, the fermionic projector P^{sea} only depends on $\mathcal{B}$ and the Green's functions $s^\vee_m$, $s^\wedge_m$. Next, particles and anti-particles are introduced as follows. Taking the closure of the solution space of the Dirac equation (2.5.10) with respect to the inner product (2.4.10), one obtains a Hilbert space $\mathcal{H}_m$ with the induced

[17] For the connection to Feynman diagrams in physics see [**59**, §3.8.4].

scalar product $(. \,|\, .)_m$. Denoting by $\mathcal{N}_t := \{t\} \times \mathbb{R}^3 \subset \mathscr{M}$ the spatial hyperplane at time t, one can define the mapping $\Pi^{\text{sea}} : C_0^\infty(\mathcal{N}_t, S\mathscr{M}) \to C^\infty(\mathscr{M}, S\mathscr{M})$ by

$$(\Pi^{\text{sea}}\psi)(x) = -2\pi \int_{\mathbb{R}^3} P^{\text{sea}}\big(x, (t, \vec{y})\big)\, \gamma^0\, \psi(\vec{y})\, d^3y \tag{2.5.13}$$

(where by $C_0^\infty(\mathcal{N}_t, S\mathscr{M})$ and $C^\infty(\mathscr{M}, S\mathscr{M})$ we denote the set of smooth wave functions on $\mathcal{N}_t$ with compact support and the set of smooth wave functions on $\mathscr{M}$, respectively). Denoting the orthogonal projection onto a subspace $A \subset \mathcal{H}_m$ by $\Pi_A : \mathcal{H}_m \to \mathcal{H}_m$, the operator Π^{sea} can be extended by continuity to a projection operator on $\mathcal{H}_m$, i.e.

$$\Pi^{\text{sea}} : \mathcal{H}_m \to \mathcal{H}_m \qquad \text{with} \qquad (\Pi^{\text{sea}})^* = \Pi^{\text{sea}} = (\Pi^{\text{sea}})^2 .$$

This allows us to introduce particles and anti-particles by

$$\Pi := \Pi^{\text{sea}} + \Pi_{\text{span}(\psi_1, \ldots, \psi_{n_{\text{p}}})} - \Pi_{\text{span}(\phi_1, \ldots, \phi_{n_{\text{a}}})}$$

with $n_{\text{p}}, n_{\text{a}} \in \mathbb{N}_0$, provided that the functions $\phi_l \in \mathcal{H}_m$ ($l = 1, \ldots, n_{\text{p}}$) lie in the image of Π^{sea}, while the vectors $\psi_k \in \mathcal{H}_m$ ($k = 1, \ldots, n_{\text{a}}$) are in the orthogonal complement of the image of Π^{sea}. Employing the normalization conditions in [**59**, eq. (2.1.72)], the operator Π reads

$$\Pi\psi := \Pi^{\text{sea}}\psi + \frac{1}{2\pi} \sum_{k=1}^{n_{\text{p}}} \psi_k\, (\psi_k|\psi)_m - \frac{1}{2\pi} \sum_{l=1}^{n_{\text{a}}} \phi_l\, (\phi_l|\psi)_m .$$

The projection operator Π can be written in the form (2.5.13) with the distribution[18]

$$P_m(x,y) = P_m^{\text{vac}}(x,y) - \frac{1}{2\pi} \sum_{k=1}^{n_{\text{p}}} \psi_k(x)\overline{\psi_k(y)} + \frac{1}{2\pi} \sum_{l=1}^{n_{\text{a}}} \phi_l(x)\overline{\phi_l(y)} ,$$

thus allowing for n_{p} particles and n_{a} anti-particles of mass m (with $n_{\text{p}}, n_{\text{a}} \in \mathbb{N}_0$).

2.5.4. The Auxiliary Fermionic Projector. The previous considerations can be extended to several Dirac seas of different masses by forming the so-called "auxiliary fermionic projector." In order to recover the Standard Model of elementary particle physics, one clearly needs to take into account all fundamental fermions as compiled in Table (2.3.20). As outlined in §2.3.7, the fundamental fermions consist of quarks and leptons, each of which obeys the corresponding Dirac equation. Quarks come in three colors (red, green, blue), each quark color as well as the leptons come in three generations, and each generation consists of a (weak) isospin doublet (isospin up and down). Accordingly, there are four "blocks" (three quark blocks and one lepton block), each of which consists of two "sectors" (corresponding to the isospin up and down component of the electroweak theory, respectively), and each sector comprises three generations. Furthermore, according to §2.3.7, the leptons split into a charged and a neutrino component, whereas all quarks carry an electric charge. In this way, all fundamental fermions either belong to a charged or neutrino component. We also point out that the theory of causal fermion systems does not distinguish the chirality of neutrinos in the lepton block and even allows for massive neutrinos. In order to take these facts into account, the (bare) mass of the i^{th} neutrino generation will be denoted by $\tilde{m}_i$, whereas the (bare) mass of the i^{th} generation in the charged sector is denoted by m_i ($i = 1, 2, 3$). Note that the

[18] For the normalization of the particle and anti-particle states see [**53**, Section 2.8] or [**59**, §3.4.3] and [**69**].

bare mass of all charged particles in each generation is the same. Thus for describing the elementary particles of the Standard Model, one requires 24 Dirac seas in total.

Moreover, as mentioned in §2.3.7, the theory of causal fermion systems takes into account massless right-handed neutrinos as well as the fact that they are not observed for low energies. The mass of the right-handed neutrino is given by $\tilde{m}_4 = 0$, and its chirality is distinguished by the chiral projector χ_R. It turns out that a suitable regularization of these right-handed states will break the chiral symmetry. This regularization will be chosen in such a way that it vanishes in the low-energy regime, thus taking into consideration that right-handed neutrinos have never been observed in nature.

In order to extend the constructions outlined in §2.5.3 to all fundamental fermions including right-handed massless neutrinos, it is useful to define the *auxiliary fermionic projector of the vacuum* P^{aux} by

$$P^{\text{aux}} = P^N_{\text{aux}} \oplus P^C_{\text{aux}}\,, \tag{2.5.14}$$

where

$$P^N_{\text{aux}} = \Big(\bigoplus_{\beta=1}^{3} P^{\text{vac}}_{\tilde{m}_\beta} \Big) \oplus 0 \qquad \text{and} \qquad P^C_{\text{aux}} = \bigoplus_{a=1}^{7} \bigoplus_{\beta=1}^{3} P^{\text{vac}}_{m_\beta} \tag{2.5.15}$$

(where the Dirac seas $P^{\text{vac}}_{\tilde{m}_\beta}$ and $P^{\text{vac}}_{m_\beta}$ are given in analogy to (2.5.7)). Thus P^{aux} is composed of 25 direct summands, four in the neutrino and 21 in the charged sector. The fourth direct summand of P^N_{aux} has the purpose of describing the right-handed massless neutrinos (or, in the terminology of [**59**], the "right-handed high-energy states").

Next, the *chiral asymmetry matrix* X and the *mass matrix* Y are introduced by

$$X = (\mathbb{1}_{\mathbb{C}^3} \oplus \tau_{\text{reg}}\, \chi_R) \oplus \bigoplus_{a=1}^{7} \mathbb{1}_{\mathbb{C}^3}\,, \tag{2.5.16}$$

$$mY = \text{diag}\big(\tilde{m}_1, \tilde{m}_2, \tilde{m}_3, 0\big) \oplus \bigoplus_{a=1}^{7} \text{diag}\big(m_1, m_2, m_3\big)\,, \tag{2.5.17}$$

where m is an arbitrary mass parameter and $\tau_{\text{reg}} \in (0,1]$ is a dimensionless parameter for which we always assume the scaling

$$\tau_{\text{reg}} = (m\varepsilon)^{p_{\text{reg}}} \qquad \text{with} \qquad 0 < p_{\text{reg}} < 2 \qquad \text{and} \qquad 0 < \varepsilon \lesssim \ell_P\,.$$

This allows us to rewrite the auxiliary fermionic projector as (cf. [**59**, eq. (5.2.5)])

$$P^{\text{aux}} = Xt \qquad \text{with} \qquad t := \bigoplus_{\beta=1}^{25} P^{\text{vac}}_{mY^\beta_\beta}\,,$$

where t is a solution of the Dirac equation

$$(i\partial\!\!\!/ - mY)\, t = 0\,.$$

In order to introduce an interaction, we insert an operator $\mathcal{B}$ into the Dirac equation,

$$(i\partial\!\!\!/ + \mathcal{B} - mY)\, \tilde{t} = 0\,. \tag{2.5.18}$$

The causal perturbation theory defines $\tilde{t} = \tilde{t}[\mathcal{B}]$ in terms of a unique perturbation series (see §2.5.3 or [**59**, Section 2.1]). Assuming the *causality compatibility condition*[19]

$$(i\partial\!\!\!/ + \mathcal{B} - mY)\, X = X^*\, (i\partial\!\!\!/ + \mathcal{B} - mY) \qquad \text{for all } \tau_{\text{reg}} \in (0,1]\,, \tag{2.5.19}$$

[19]The causality compatibility condition is required in order to obtain *bounded* line integrals, for details we refer to [**48**].

the (unregularized) interacting auxiliary fermionic projector is then introduced by $X\tilde{t}$.

The remainder of this section is devoted to establish the connection to physics. To this end, the procedure is as follows: The first step is to obtain a convenient form for $\tilde{t}$ by performing the so-called "light-cone expansion." The next step is to "regularize" the resulting expressions. Afterwards, one can introduce particles and anti-particles in the fashion of §2.5.3, thus giving rise to an interacting auxiliary fermionic projector $\tilde{P}^{\text{aux}}$. Forming the so-called "sectorial projection," one finally obtains a fermionic projector $\tilde{P}$. The general procedure is summarized in [**59**, Remark 4.2.3]. Assuming that the resulting operator $\tilde{P}$ is a minimizer of the causal action, the corresponding Euler-Lagrange (EL) equations give rise to conditions on the interaction $\mathcal{B}$ in a "weak evaluation on the light cone." The underlying picture is to regard precisely those potentials $\mathcal{B}$ as physically relevant for which the Euler-Lagrange equations are satisfied.

Before entering the light-cone expansion in more detail, it is necessary to specify the admissible form of the external potential $\mathcal{B}$ in the Dirac equation (2.5.18). To this end, we point out that the specific form of the interacting potential $\mathcal{B}$ is determined by the causal action principle. For this reason, one should choose $\mathcal{B}$ as general as possible, thereby even allowing for potentials which are usually not considered in physics. The following overview of potentials which are of relevance in the theory of causal fermion systems as worked out in [**59**] is due to [**59**, §3.4.5]. The most obvious choice is an electromagnetic potential,[20]

$$\mathcal{B} = \mathcal{A}\,.$$

More generally, one can choose *chiral potentials*, which may be non-diagonal in the generations,

$$\mathcal{B} = \chi_L\, \mathcal{A}_R + \chi_R\, \mathcal{A}_L\,,$$

To describe a *gravitational field*, one needs to choose $\mathcal{B}$ as a differential operator of first order; more precisely,

$$\mathcal{B} = \mathcal{D} - i\phi\!\!\!/\,,$$

where $\mathcal{D}$ is the Dirac operator in the presence of a gravitational field.

Apart from the above choices of $\mathcal{B}$ motivated from physics, one can also choose other physically less obvious operators, like *scalar* or *pseudoscalar potentials*,

$$\mathcal{B} = \Phi + i\Gamma\Xi$$

with $\Phi = \Phi^\alpha_\beta$, $\Xi = \Xi^\alpha_\beta$ and $\alpha, \beta = 1, \ldots, g$ for some $g \in \mathbb{N}$ (here $\Gamma \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ is the usual pseudoscalar matrix). Next, one can consider a *scalar differential operator*,

$$\mathcal{B} = i\Phi^j \partial_j\,,$$

or a higher order differential operator. Furthermore, *pseudoscalar differential potentials* turn out to be useful,

$$\mathcal{B} = \Gamma\left(v^j \partial_j + \partial_j v^j\right)$$

(see e.g. [**59**, §3.7.4]), as well as *vector differential potentials* (cf. [**59**, §3.7.5]). Further potentials and fields are discussed in [**59**, §3.9.3].

[20] Following the convention in [**59**], the coupling constant e is omitted in the Dirac equation. This convention can be obtained from the usual choice $\mathcal{B} = e\mathcal{A}\!\!\!/$ by the transformation $A \to e^{-1}A$.

For clarity, we point out that the situation in which the external potential $\mathcal{B}$ is composed of left- or right-handed potentials, i.e.

$$\mathcal{B} = \chi_L \,\mathcal{A}\!\!\!/_R + \chi_R \,\mathcal{A}\!\!\!/_L \,, \tag{2.5.20}$$

is of particular interest. The reason is that potentials of the form (2.5.20), which are referred to as *chiral potentials*, allow for the description of *gauge fields*. For instance, an electromagnetic field is described by choosing $A_L = A_R = A$, where A is the electromagnetic potential. Furthermore, as weak potentials only couple to the left-handed component of fundamental fermions, a left-handed potential is required for describing the weak interaction in the Standard Model. Note that in the case of *non-abelian* gauge fields, the potentials A_L and A_R take values in a Lie algebra (cf. [**59**, §2.2.3]).

2.5.5. The Light-Cone Expansion. After these preliminaries, let us enter the light-cone expansion in more detail. For simplicity, we first restrict attention to a single Dirac sea. Thus the starting point is a Dirac sea $P(x,y)$ of mass m as given by (2.5.7),

$$P(x,y) = \int \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)} \,. \tag{2.5.21}$$

It is useful to observe that $P(x,y)$ can be rewritten in the form

$$P(x,y) = (i\not{\partial}_x + m)\, T_{m^2}(x,y) \,, \tag{2.5.22}$$

where $T_{m^2}(x,y)$ is the scalar bi-distribution

$$T_{m^2}(x,y) := \int \frac{d^4k}{(2\pi)^4}\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)} \,. \tag{2.5.23}$$

The important point in what follows is that $T_{m^2}(x,y)$ can be reformulated as

$$\begin{aligned} T_{m^2}(x,y) = &-\frac{1}{8\pi^3} \left(\frac{\mathrm{PP}}{(y-x)^2} + i\pi \delta\big((y-x)^2\big)\, \epsilon\big((y-x)^0\big) \right) \\ &+ \frac{m^2}{32\pi^3} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!\,(j+1)!}\, \frac{\big(m^2 (y-x)^2\big)^j}{4^j} \\ &\times \Big(\log \big| m^2 (y-x)^2 \big| + c_j + i\pi\, \Theta\big((y-x)^2\big)\, \epsilon\big((y-x)^0\big) \Big) \end{aligned} \tag{2.5.24}$$

with real coefficients c_j (here Θ and ϵ are again the Heaviside and the sign function, respectively), where the distribution PP/ξ^2, denoted by *principal value*, is defined by evaluating weakly with a test function $\eta \in C_0^\infty(\mathcal{M})$ and by removing the positive and negative parts of the pole in a symmetric way,

$$\begin{aligned} \int \frac{\mathrm{PP}}{\xi^2}\, \eta(\xi)\, d^4\xi &= \lim_{\nu \searrow 0} \int \Theta\big(|\xi^2| - \nu\big)\, \frac{1}{\xi^2}\, \eta(\xi)\, d^4\xi \\ &= \lim_{\nu \searrow 0} \frac{1}{2} \sum_{\pm} \int \frac{1}{\xi^2 \pm i\nu}\, \eta(\xi)\, d^4\xi = \lim_{\nu \searrow 0} \frac{1}{2} \sum_{\pm} \int \frac{1}{\xi^2 \pm i\nu \xi^0}\, \eta(\xi)\, d^4\xi \end{aligned}$$

(where $\xi^2 \equiv \xi^j \xi_j$ is the Minkowski inner product). It is important to observe that the distribution $T_{m^2}(x,y)$ is smooth away from the light cone (i.e. for $\xi^2 \neq 0$ with $\xi = y - x$), whereas on the light cone

$$L = \big\{ \xi \in \mathcal{M} : \xi^2 = 0 \big\} \,,$$

the distribution $T_{m^2}(x,y)$ has singularities (see [**59**, Lemma 1.2.9]). As a consequence, the main contributions to the fermionic projector (2.5.21) arise from the contributions

on the light cone (also cf. [**48**, §2.1.4] and [**53**, Section 4.2]). A useful tool in order to analyze the singular structure of the fermionic projector on the light cone in more detail is the so-called "light-cone expansion." In particular, the expression (2.5.24) turns out to be a light-cone expansion, which is defined as follows (see [**59**, Definition 2.2.1]):

DEFINITION 2.5.1. *A distribution $A(x,y)$ is **of the order** $\mathcal{O}((y-x)^{2p})$ for $p \in \mathbb{Z}$ if*

$$(y-x)^{-2p}\, A(x,y)$$

is a regular distribution (i.e. a locally integrable function). An expansion of the form

$$A(x,y) = \sum_{j=g}^{\infty} A^{[j]}(x,y) \tag{2.5.25}$$

with $g \in \mathbb{Z}$ is called **light-cone expansion** *if $A^{[j]}(x,y) \in \mathcal{O}((y-x)^{2j})$ for all $j \geq g$ (that is, the $A^{[j]}(x,y)$ are distributions of the order $\mathcal{O}((y-x)^{2j})$ for every $j \geq g$) and if A is approximated by the partial sums in the sense that for all $p \geq g$,*

$$A(x,y) - \sum_{j=g}^{p} A^{[j]}(x,y) \qquad \text{is of the order } \mathcal{O}\big((y-x)^{2p+2}\big)\,. \tag{2.5.26}$$

The parameter g gives the leading order of the singularity of $A(x,y)$ on the light cone. We point out that we do not demand that the infinite series in (2.5.25) converges. Thus, similar to a formal Taylor series, the series in (2.5.26) is defined only via the approximation by the partial sums (2.5.26).

Due to the factors $(y-x)^{2j}$, the expression in (2.5.24) is a light-cone expansion. More precisely, the term with the leading singularity becomes integrable after multiplying by $(y-x)^2$, showing that $g = -1$. The light-cone expansion of the kernel of the fermionic projector $P(x,y)$ in (2.5.21) is readily obtained using the relation (2.5.22). To this end, one simply applies the differential operator $(i\partial\!\!\!/ + m)$ to the above series expansion of $T_{m^2}(x,y)$ and computes the derivatives term by term. Since differentiation increases the order of the singularity on the light cone by one, we thus obtain a light-cone expansion of the form (2.5.25) with $g = -2$.

In order to study the effect of an interaction $\mathcal{B}$ in more detail, our goal is to analyze the solution of the modified Dirac equation (2.5.10). Unfortunately, deriving a light-cone expansion of the fermionic projector P^{sea} in (2.5.12) is not quite straightforward (see formula (2.5.41) below). Since the unique perturbation series in (2.5.11) essentially enter the construction of P^{sea}, it is a promising strategy to develop a method for performing the light-cone expansion of each summand of the perturbation series in (2.5.11). Let us begin by considering the free advanced Green's function $s^{\vee}_m$ of the Dirac equation of mass m in position space: Similar to (2.5.22), it is again convenient to pull the Dirac matrices out of $s^{\vee}_m$ by setting

$$s^{\vee}_m(x,y) = (i\partial\!\!\!/_x + m)\, S^{\vee}_{m^2}(x,y)\,, \tag{2.5.27}$$

where $S^{\vee}_{m^2}(x,y)$ can be represented as (cf. [**59**, eq. (2.2.7)])

$$\begin{aligned} S^{\vee}_{m^2}(x,y) =& -\frac{1}{2\pi}\, \delta\big((y-x)^2\big)\, \Theta\big(y^0-x^0\big) \\ &+ \frac{m^2}{8\pi} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!\,(j+1)!}\, \frac{\big(m^2(y-x)^2\big)^j}{4^j}\, \Theta\big((y-x)^2\big)\, \Theta\big(y^0-x^0\big)\,. \end{aligned} \tag{2.5.28}$$

This computation shows that $S^\vee_{m^2}(x,y)$ has a $\delta((y-x)^2)$-like singularity on the light cone. Moreover, one sees that $S^\vee_{m^2}$ is a power series in m^2. The important point for what follows is that higher order contributions in m^2 contain more factors $(y-x)^2$ and are thus of higher order on the light cone. More precisely,

$$\left(\frac{d}{dm^2}\right)^n S^\vee_{m^2}(x,y)\Big|_{m=0} \qquad \text{is of the order } \mathcal{O}\big((y-x)^{2n-2}\big)$$

(where m is treated as a variable parameter). According to (2.5.27), the Dirac Green's function is obtained by computing the first partial derivatives of (2.5.28). Therefore, the Green's function $s^\vee_m(x,y)$ has an even $\sim \delta'((y-x)^2)$-like singularity on the light cone, and the higher order contributions in m are again of increasing order on the light cone. This means that we can view the Taylor expansion of (2.5.27) in m,

$$s^\vee_m(x,y) = \sum_{n=0}^\infty (i\partial\!\!\!/ + m)\, \frac{1}{n!} \left(\frac{d}{dm^2}\right)^n S^\vee_{m^2}(x,y)\Big|_{m=0}\,,$$

as a light-cone expansion of the free Green's function. These considerations motivate the strategy to first expand (2.5.11) with respect to the mass (see [**59**, Section 2.2]).

The expansion of (2.5.11) with respect to m gives a double sum over the orders in the mass parameter and in the external potential. It is convenient to combine these two expansions in a single perturbation series. To this end, we rewrite the Dirac operator as

$$i\partial\!\!\!/ + \mathcal{B} - m = i\partial\!\!\!/ + B \qquad \text{with} \qquad B := \mathcal{B} - m\,. \tag{2.5.29}$$

For the light-cone expansion of the Green's functions, we will always view B as the perturbation of the Dirac operator. This has the advantage that the unperturbed objects are massless. Thus expanding in powers of B gives the mass expansion and the perturbation expansion in one step. In order to simplify the notation, for the massless objects we usually omit the index m. Thus we write the Green's function of the massless Dirac equation in the Minkowski vacuum as

$$s^\vee(x,y) = i\partial\!\!\!/_x\, S^\vee_{m^2}(x,y)\big|_{m=0}\,, \qquad s^\wedge(x,y) = i\partial\!\!\!/_x\, S^\wedge_{m^2}(x,y)\big|_{m=0}\,.$$

Then the interacting Green's functions are given by the perturbation series

$$\tilde{s}^\vee = \sum_{k=0}^\infty (-s^\vee B)^k s^\vee\,, \qquad \tilde{s}^\wedge = \sum_{k=0}^\infty (-s^\wedge B)^k s^\wedge\,. \tag{2.5.30}$$

The following constructions (for details see [**59**, Chapter 2]) are exactly the same for the advanced and retarded Green's functions. In order to treat both cases at once, in the remainder of this subsection we will omit all superscripts '$\vee$', '$\wedge$'. The formulas for the advanced and retarded Green's functions are obtained by either adding '$\vee$' or '$\wedge$' to all factors s, S. Moreover, in order to carry out the mass expansion of S_{m^2}, we set $a = m^2$ and use the notation

$$S^{(l)} = \left(\frac{d}{da}\right)^l S_a\big|_{a=0}\,. \tag{2.5.31}$$

Furthermore, the combination $(y-x)_k\, S^{(-1)}(x,y)$ is defined by (cf. [**59**, eq. (2.2.24)])

$$(y-x)_k\, S^{(-1)}(x,y) := 2\,\frac{\partial}{\partial x^k} S^{(0)}(x,y)\,.$$

In the following, we restrict attention to chiral potentials (2.5.20),

$$\mathcal{B} = \chi_L\, A\!\!\!/_R + \chi_R\, A\!\!\!/_L\,. \tag{2.5.32}$$

Considering Dirac particles with (bare) rest masses $m_1, \ldots, m_g$ for some $g \in \mathbb{N}$ and introducing the *mass matrix* Y by

$$Y = \frac{1}{m} \operatorname{diag}(m_1, \ldots, m_g)$$

for some mass parameter m, we combine the mass term with the potential in analogy to (2.5.29) by setting

$$B = \chi_L \mathcal{A}_R + \chi_R \mathcal{A}_L - mY \,. \tag{2.5.33}$$

Then the perturbation expansion of the causal Green's functions can again be written in the form (2.5.30). In order to allow for scalar and pseudoscalar potentials it is most convenient to replace the mass matrix by a spacetime-dependent *dynamical mass matrix*

$$Y = Y(x) := \chi_L Y_L(x) + \chi_R Y_R(x) \,. \tag{2.5.34}$$

In what follows, we are mainly interested in analyzing the structure of solutions of the Dirac equation

$$(i\mathcal{\partial\!\!\!/} + \mathcal{B} - mY)\, \psi(x) = 0 \,. \tag{2.5.35}$$

In order to recall the structural results on the contributions to the light-cone expansion of the Green's functions, let us state the following definitions: For the line integrals, we introduce the short notation

$$\int_x^y [l, r \,|\, n]\, dz\; f(z) := \int_0^1 d\alpha\; \alpha^l\, (1-\alpha)^r\, (\alpha - \alpha^2)^n\; f(\alpha y + (1-\alpha) x) \,.$$

Furthermore, we abbreviate the following products with multi-indices,

$$\partial_z^J := \frac{\partial}{\partial z^{j_1}} \cdots \frac{\partial}{\partial z^{j_l}} \,, \qquad (y-x)^J := (y-x)^{j_1} \cdots (y-x)^{j_l} \,, \quad \gamma^J := \gamma^{j_1} \cdots \gamma^{j_l} \,,$$

where $J = (j_1, \ldots, j_l)$. Unfortunately, in order to clarify the detailed structure of the light-cone expansion it seems unavoidable to recall the following definitions and results. Let us begin with [**59**, Theorem 2.2.4]:

THEOREM 2.5.2. *In the presence of chiral potentials* (2.5.33)*, the light-cone expansion of the k^{th} order contribution $((-sB)^k\, s)(x,y)$ to the perturbation series* (2.5.30) *can be written as an infinite sum of expressions, each of which has the form*

$$\begin{aligned} \chi_{c_0}\, C\, (y-x)^I \int_x^y [l_1, r_1 \,|\, n_1]\, dz_1\; \partial_{z_1}^{I_1}\, \Box_{z_1}^{p_1}\, V^{(1)}_{J_1, c_1}(z_1) \int_{z_1}^y [l_2, r_2 \,|\, n_2]\, dz_2\; \partial_{z_2}^{I_2}\, \Box_{z_2}^{p_2}\, V^{(2)}_{J_2, c_2}(z_2) \\ \cdots \int_{z_{k-1}}^y [l_k, r_k \,|\, n_k]\, dz_k\; \partial_{z_k}^{I_k}\, \Box_{z_k}^{p_k}\, V^{(k)}_{J_k, c_k}(z_k)\; \gamma^J\; S^{(h)}(x,y) \,. \end{aligned} \tag{2.5.36}$$

In this formula, C is a complex number and the parameters l_a, r_a, n_a, and p_a are non-negative integers; the indices c and c_a can take the two values L or R. The functions $V^{(a)}_{J_a, c_a}$ (where J_a is a multi-index and $c_a \in \{L, R\}$ is a chiral index) coincide with any of the individual potentials in (2.5.33) *and* (2.5.34) *with chirality c_a, i.e.*

$$\begin{aligned} V^{(a)}_{c_a} &= A_{c_a} \qquad \text{(in which case } |J_a| = 1\text{)} \qquad \text{or} \\ V^{(a)}_{c_a} &= m Y_{c_a} \qquad \text{(in which case } |J_a| = 0\text{)} \,. \end{aligned} \tag{2.5.37}$$

The chirality c_a of the potentials is determined by the following rule:

(i) The chirality is reversed precisely at every mass matrix, i.e.

$$c_{a-1} \text{ and } c_a \quad \begin{cases} \text{coincide} & \text{if } V^{(a)}_{c_a} = A_{c_a} \\ \text{are opposite} & \text{if } V^{(a)}_{c_a} = m Y_{c_a} \end{cases}$$

for all $a = 1, \ldots, k$.

The tensor indices of the multi-indices in (2.5.36) *are all contracted with each other, according to the following rules:*

(a) No two tensor indices of the same multi-index are contracted with each other.

(b) The tensor indices of the factor γ^J are all contracted with different multi-indices, in the order of their appearance in the product (2.5.36) *(i.e., for $J = (j_1, \ldots, j_l)$ and $1 \leq a < b \leq l$, the multi-index with which j_a is contracted must stand to the left of the multi-index corresponding to j_b).*

The parameter h is given by

$$2h = k - 1 - |I| + \sum_{a=1}^{k} \Big(|I_a| + 2p_a \Big) .$$

The number of factors $(y-x)$ is bounded by

$$|I| \leq k + 1 - \sum_{a=1}^{k} |I_a| .$$

In a few words, Theorem 2.5.2 states that the light-cone expansion of the k^{th} order Feynman diagrams (cf. (2.5.11)) can be written with k nested line integrals. In particular, the number of summands (2.5.36) is finite to every order on the light cone. Therefore, the light-cone expansion of Theorem 2.5.2 makes mathematical sense in terms of Definition 2.5.1. Moreover, the integer h is bounded from below by $h \geq -1$. By neglecting all terms of the order $\mathcal{O}((y-x)^{-2})$, to the leading singularity $h = -1$ the sum over all Feynman diagrams (2.5.11) takes the form (cf. [**59**, eq. (2.2.57)])

$$\chi_c \, \tilde{s}(x,y) = \chi_c \, \text{Pexp}\left(-i \int_x^y (y-x)_j \, A^j_c(z) \, dz \right) s(x,y) \;+\; \mathcal{O}((y-x)^{-2})$$

for $c \in \{L, R\}$, where Pexp is defined as follows (see [**59**, Definition 2.2.5]).

Definition 2.5.3. *For a smooth one-parameter family of matrices $F(\alpha)$, $\alpha \in \mathbb{R}$, the* **ordered exponential** $\text{Pexp}(\int F(\alpha)\, d\alpha)$ *is given by the Dyson series*

$$\begin{aligned} \text{Pexp}\left(\int_a^b F(\alpha)\, d\alpha \right) &= \mathbb{1} + \int_a^b F(t_0)\, dt_0 \;+\; \int_a^b dt_0\, F(t_0) \int_{t_0}^b dt_1\, F(t_1) \\ &\quad + \int_a^b dt_0\, F(t_0) \int_{t_0}^b dt_1\, F(t_1) \int_{t_1}^b dt_2\, F(t_2) + \cdots . \end{aligned}$$

For ordered exponentials over the chiral potentials, we use the short notations

$$\begin{aligned} &\text{Pexp}\left(-i \int_x^y (y-x)_j \, A^j_c(z)\, dz \right) = \text{Pexp}\left(-i \int_x^y A^j_c \, (y-x)_j \right) = \text{Pe}^{-i \int_x^y A^j_c \, (y-x)_j} \\ &:= \text{Pexp}\left(-i \int_0^1 A^j_c\big|_{\alpha y + (1-\alpha) x} \, (y-x)_j \, d\alpha \right). \end{aligned}$$

PROPOSITION 2.5.4. *Every contribution* (2.5.36) *to the light-cone expansion of Theorem 2.5.2 can be written as a finite sum of expressions of the form*

$$\chi_c\, C\, (y-x)^K\, W^{(0)}(x) \int_x^y [l_1, r_1 \,|\, n_1]\, dz_1\; W^{(1)}(z_1) \int_{z_1}^y [l_2, r_2 \,|\, n_2]\, dz_2\; W^{(2)}(z_2) \\ \cdots \int_{z_{\alpha-1}}^y [l_\alpha, r_\alpha \,|\, n_\alpha]\, dz_\alpha\; W^{(\alpha)}(z_\alpha)\; \gamma^J\; S^{(h)}(x,y)$$

with $\alpha \leq k$, where the factors $W^{(\beta)}$ are composed of the potentials and their partial derivatives,

$$W^{(\beta)} = (\partial^{K_{a_\beta}} \Box^{p_{a_\beta}} V^{(a_\beta)}_{J_{a_\beta}, c_{a_\beta}}) \cdots (\partial^{K_{b_\beta}} \Box^{p_{b_\beta}} V^{(b_\beta)}_{J_{b_\beta}, c_{b_\beta}})$$

with $a_1 = 1$, $a_{\beta+1} = b_\beta + 1$, $b_\beta \geq a_\beta - 1$ (in the case $b_\beta = a_\beta - 1$, $W^{(\beta)}$ is identically one), and $b_\alpha = k$. The parameters l_a, r_a, and n_a are non-negative integers, C is a complex number, and $c, c_a \in \{L, R\}$ are chiral indices. The potentials $V^{(a)}$ are again given by (2.5.37)*; their chirality is determined by the rule (i) in Theorem 2.5.2. The tensor indices of the multi-indices J, K, J_a, and K_a are all contracted with each other, according to the rules (a),(b) of Theorem 2.5.2 and*

(c) The tensor indices of $(y-x)^K$ are all contracted with the tensor indices of the factors $V^{(a)}_{J_a}$ or γ^J (but not with the factors ∂^{K_a}).

We have the relation

$$2h = k - 1 - |K| + \sum_{a=1}^k \big(|K_a| + 2p_a\big)\,.$$

DEFINITION 2.5.5. *A contribution of the form* (2.5.36) *to the light-cone expansion of Proposition 2.5.4 is called* **phase-free** *(see* [**59**, Definition 2.2.7]*) if*

$$|K_a| + 2p_a > 0 \qquad \text{whenever} \qquad J_a \text{ is contracted with } (y-x)^K.$$

From every phase-free contribution the corresponding **phase-inserted** *contribution is obtained as follows: We insert ordered exponentials according to the replacement rule*

$$W^{(\beta)}(z_\beta) \;\longrightarrow\; W^{(\beta)}(z_\beta)\; \text{Pexp}\left(-i \int_{z_\beta}^{z_{\beta+1}} A^{j_\beta}_{c_\beta}\, (z_{\beta+1} - z_\beta)_{j_\beta}\right), \quad \beta = 0, \ldots, \alpha\,,$$

where we set $z_0 = x$ and $z_{\alpha+1} = y$. The chiralities c_β ($\beta = 0, \ldots, \alpha$) are determined by the relations $c_0 = c$ and

$$c_{\beta-1} \text{ and } c_\beta \left\{ \begin{array}{c} \text{coincide} \\ \text{are opposite} \end{array} \right\} \\ \text{if } W^{(\beta-1)} \text{ contains an } \left\{ \begin{array}{c} \text{even} \\ \text{odd} \end{array} \right\} \text{number of factors } Y_\cdot.$$

THEOREM 2.5.6. *To every order on the light cone, the number of phase-free contributions is finite. The light-cone expansion of the Green's function $\tilde{s}(x,y)$ is given by the sum of the corresponding phase-inserted contributions. See* [**59**, Theorem 2.2.8].

After these technical preliminaries, we are in the position to outline the structure of the solution P^{sea} of the Dirac equation (2.5.35),

$$(i\partial\!\!\!/ + \mathcal{B} - mY)\, \psi(x) = 0\,, \tag{2.5.38}$$

in a self-contained way. To this end, we consider the Green's functions

$$s^\pm(p) = \not{p}\, S^\pm_{a\,|\,a=0}(p) \qquad \text{with} \qquad S^\pm_a(p) = \lim_{\nu \searrow 0} \frac{1}{p^2 - a \mp i\nu}$$

and again use the notation (2.5.31),

$$S^{\pm\,(l)} = \left(\frac{d}{da}\right)^l S^\pm_{a\,|\,a=0} \qquad (l = 0, 1, 2, \ldots)\,.$$

The perturbation expansion of these Dirac Green's functions is, in analogy to (2.5.30) or (2.5.11), given by the formal series

$$\tilde{s}^+ := \sum_{n=0}^\infty (-s^+\, B)^n s^+ \qquad \text{and} \qquad \tilde{s}^- := \sum_{n=0}^\infty (-s^-\, B)^n s^-\,.$$

Next, we introduce the *residual fundamental solution* $\tilde{p}^{\text{res}}$ in analogy to (2.5.23) by

$$\tilde{p}^{\text{res}} := \frac{1}{2\pi i} \left(\tilde{s}^+ - \tilde{s}^-\right),$$

and define the *residual fermionic projector* $\tilde{P}^{\text{res}}(x,y)$ by

$$\tilde{P}^{\text{res}}(x,y) := \frac{1}{2} \left(\tilde{p}^{\text{res}} - \tilde{k}\right)(x,y)\,.$$

In analogy to the mass expansion of the Green's functions (2.5.31), we set

$$T^{(l)}_{\text{formal}} = \left(\frac{d}{da}\right)^l T_a\big|_{a=0} \qquad (l = 0, 1, 2, \ldots)\,. \tag{2.5.39}$$

Then due to [**59**, Proposition 2.2.11], the formal light-cone expansion of the residual fermionic projector $\tilde{P}^{\text{res}}(x,y)$ is obtained from that of the causal Green's functions by the replacement

$$S^{(l)} \to T^{(l)}_{\text{formal}}\,.$$

Unfortunately, the formal a-derivatives of T_a in (2.5.39) contain logarithmic poles of the form $\log|a|$ (cf. [**59**, eq. (2.2.116)]) which cause serious problems. These difficulties, arising from the logarithm in (2.5.24), are referred to as *logarithmic mass problem*. In order to resolve the logarithmic mass problem, the first step is to "regularize" the formal light-cone expansion by subtracting the problematic $\log|a|$ term,

$$T^{\text{reg}}_a(x,y) := T_a(x,y) - \frac{a}{32\pi^3}\, \log|a| \sum_{j=0}^\infty \frac{(-1)^j}{j!\,(j+1)!}\, \frac{(a\xi^2)^j}{4^j}$$

(where $\xi^2 \equiv \xi^j \xi_j$ denotes again the Minkowski inner product). We then introduce

$$T^{(l)} := \left(\frac{d}{da}\right)^l T^{\text{reg}}_{a\,|\,a=0} \qquad (l = 0, 1, 2, \ldots)\,. \tag{2.5.40}$$

By definition (see [**59**, Definition 2.2.12]), the so-called *causal contribution* $\tilde{P}^{\text{causal}}$ to the fermionic projector is obtained from the residual Dirac sea $\tilde{P}^{\text{res}}$ by replacing all factors $T^{(h)}_{\text{formal}}$ in the formal light-cone expansion of $\tilde{P}^{\text{res}}(x,y)$ by $T^{(h)}$. Next, the *non-causal low energy contribution* $\tilde{P}^{\text{le}}$ to the fermionic projector is given by

$$\tilde{P}^{\text{le}}(x,y) = \tilde{P}^{\text{res}}(x,y) - \tilde{P}^{\text{causal}}(x,y)\,.$$

Due to [**59**, Theorem 2.2.13], the light-cone expansion of the causal Green's functions also holds for the causal contribution $\tilde{P}^{\text{causal}}$ to the fermionic projector if one simply

replaces $S^{(l)} \to T^{(l)}$ with $T^{(l)}$ according to (2.5.40). Moreover, the *non-causal high energy contribution* $\tilde{P}^{\text{he}}(x,y)$ to the fermionic projector is given by

$$\tilde{P}^{\text{he}}(x,y) = P^{\text{sea}}(x,y) - \tilde{P}^{\text{res}}(x,y)\,.$$

Note that to every order in the external potential $\mathcal{B}$, both $\tilde{P}^{\text{le}}(x,y)$ and $\tilde{P}^{\text{he}}(x,y)$ are smooth functions in x and y (see [**59**, Theorem 2.2.14 and Theorem 2.2.16]).

Taken together, the above constructions show that in the presence of chiral and scalar/pseudoscalar potentials (see (2.5.32), (2.5.38), (2.5.34)) the fermionic projector has a representation of the form

$$\begin{aligned} P^{\text{sea}}(x,y) = \sum_{n=-1}^{\infty} &(\text{phase-inserted line integrals}) \times T^{(n)}(x,y) \\ &+ \tilde{P}^{\text{le}}(x,y) + \tilde{P}^{\text{he}}(x,y)\,, \end{aligned} \tag{2.5.41}$$

where $(y-x)_k\, T^{(-1)}(x,y)$ is defined by the distributional equation (cf. [**59**, eq. (3.4.13)])

$$\frac{\partial}{\partial x^k} T^{(0)}(x,y) = \frac{1}{2}\,(y-x)_k\, T^{(-1)}(x,y)\,.$$

The series in (2.5.41) is a light-cone expansion which describes the singular behavior of the fermionic projector on the light cone. It is obtained from the light-cone expansion of the Green's functions by the simple replacement rule

$$S^{(n)} \longrightarrow T^{(n)}$$

with $T^{(n)}$ as defined in (2.5.40). Note that the expression (2.5.41) is "causal" in the sense that all phase-inserted line integrals are *bounded*. The important point in what follows is that the higher a-derivatives of $T_a(x,y)$ contain more factors $(y-x)^2$ and are thus of higher order on the light cone. This makes it possible to make mathematical sense of the infinite series in (2.5.41) as a light-cone expansion.

In order to put the following subsection into the context, we recall that composite expressions of the regularized fermionic projector enter the causal principle in form of the closed chain. In order to give such expressions a mathematical meaning, it is useful to evaluate composite expressions in $T^{(n)}$ and $\overline{T^{(n)}}$ weakly on the light cone (where $\overline{T^{(n)}}$ denotes the complex conjugate of $T^{(n)}$), as we now explain.

2.5.6. Weak Evaluation on the Light Cone. The techniques presented in the previous subsection §2.5.5 provide a method for analyzing the *un*regularized kernel of the fermionic projector. The purpose of the formalism of the continuum limit is to extend these methods to the regularized setting. More precisely, the representation (2.5.41), which describes the singular behavior of the fermionic projector on the light cone, is the starting point for the formalism of the continuum limit. Introducing a "regularization," the formalism of the continuum limit allows to evaluate composite expressions in the fermionic projector in a well-defined way. For details concerning the formalism of the continuum limit we refer to [**59**, Section 2.4 and Section 3.5] as well as [**53**, Chapter 4].

In order to explain the basic ideas of introducing a regularization in more detail, let $P^{\text{sea}}(x,y)$ be a solution of the interacting Dirac equation

$$(i\partial\!\!\!/ + \mathcal{B} - mY)\, P^{\text{sea}}(x,y) = 0\,. \tag{2.5.42}$$

The solution $P^{\text{sea}}(x,y)$ can be written as (cf. [**59**, Section 3.4])

$$P^{\text{sea}}(x,y) = \sum_{n=-1}^{\infty} \sum_{k} m^{p_k} (\text{phase-inserted nested line integrals}) \times T^{(n)}(x,y) + \tilde{P}^{\text{le}}(x,y) + \tilde{P}^{\text{he}}(x,y)\,. \tag{2.5.43}$$

Here the n-summands describe the different orders of the singularities on the light cone, whereas the k-sum describes all contributions to a given order on the light cone. The phase-inserted nested line integrals involve $\mathcal{B}$ and its partial derivatives, possibly sandwiched between time-ordered exponentials of chiral potentials. Since these nested line integrals are smooth functions in x and y, the series in (2.5.43) is a light-cone expansion in the sense of Definition 2.5.1, provided that the k-sum is finite for every n. This is indeed the case if $\mathcal{B}$ is composed of scalar, pseudoscalar and chiral potentials [**52**], whereas for a more general perturbation operator $\mathcal{B}$ this condition still needs to be verified. This expansion is *causal* in the sense that it depends on $\mathcal{B}$ and its partial derivatives only along the line segment $\overline{xy}$.

Note that particles and anti-particles are introduced in analogy to (2.5.44) by

$$P^{\text{aux}}(x,y) = P^{\text{sea}}(x,y) - \frac{1}{2\pi} \sum_{k=1}^{n_{\text{p}}} \psi_k(x) \overline{\psi_k(y)} + \frac{1}{2\pi} \sum_{l=1}^{n_{\text{a}}} \phi_l(x) \overline{\phi_l(y)}\,. \tag{2.5.44}$$

The formalism of the continuum limit is essentially based on the physically reasonable *assumption of macroscopic potentials and wave functions* which states that both the bosonic potentials in (2.5.42) and the fermionic wave functions in (2.5.44) vary only on the "macroscopic" length scale $\ell_{macro} \gg \ell_P$. In order to outline the basic ideas of the formalism of the continuum limit, it is necessary to first introduce a regularization of regularization length $\varepsilon > 0$. As the regularized objects are thought of as describing the "microscopic structure of spacetime," it is natural to choose $\varepsilon \lesssim \ell_P$. We are then led to "evaluate weakly" for spacetime points $x, y \in \mathscr{M}$ in such a way that $\varepsilon \ll |\vec{\xi}| \ll \ell_{macro}$, where $\xi := y - x$ and $\xi = (\xi^0, \vec{\xi})$ (cf. [**59**, eq. (3.7.77)]).

In view of the causal action principle (2.1.1), (2.1.6) our main task is to evaluate composite expressions in the fermionic projector $P(x,y)$ on the light cone. As smooth functions can be evaluated for any spacetime point, it clearly suffices to restrict attention to singular expressions on the light cone. In other words, we introduce a regularization in such a way that the smooth contributions are all left unchanged. Motivated by the special case of an $i\varepsilon$-regularization (see [**59**, §2.4.1]), introducing a regularization amounts in view of (2.5.43) to regularizing the expressions $T^{(n)}$ by the formal replacements

$$m^p\, T^{(n)} \to m^p\, T^{(n)}_{[p]}\,,$$

where the subscript [.] is added to the factors $T^{(n)}$ in order to count the power in m, and the factors $T^{(n)}_{[p]}$ are smooth functions of ξ. More precisely, one first rewrites every summand of the light-cone expansion (2.5.43) in such a way that it involves at most one factor $\not{\xi}$ (which can always be arranged), and then associates to every factor $\not{\xi}$ the corresponding factor $T^{(n)}_{[p]}$. This gives rise to the replacements (cf. [**59**, eq. (2.4.43)])

$$m^p \not{\xi} T^{(n)} \to m^p \not{\xi}^{(n)}_{[p]}\, T^{(n)}_{[p]}\,.$$

Nevertheless, composite expressions diverge in the limit $\varepsilon \searrow 0$. Setting $r := |\vec{\xi}|$ for every $\xi = (t, \vec{\xi})$, the proper method in order to analyze this singular behavior is to "evaluate weakly" in t for fixed r. To this end, one considers integrals of the form

$$\int_{-\infty}^{\infty} \eta(t) \left(\cdots \right) dt$$

for a smooth test function η, where "$\cdots$" stands for a composite expression in the $T^{(n)}$ and $\overline{T^{(n)}}$. In cases where the lower index does not need to be specified we write $T_\circ^{(n)}$. Next, to any factor $T_\circ^{(n)}$ we associate the *degree* $\deg T_\circ^{(n)}$ by

$$\deg T_\circ^{(n)} = 1 - n \,.$$

The degree is additive in products, whereas the degree of a quotient is defined as the difference of the degrees of numerator and denominator. The degree of an expression can be thought of as describing the order of its singularity on the light cone, in the sense that a larger degree corresponds to a stronger singularity. Using formal Taylor series, we can expand in the degree. In all our applications, this will give rise to terms of the form

$$\eta(x,y)\, \frac{T_\circ^{(a_1)} \cdots T_\circ^{(a_\alpha)}\, \overline{T_\circ^{(b_1)} \cdots T_\circ^{(b_\beta)}}}{T_\circ^{(c_1)} \cdots T_\circ^{(c_\gamma)}\, \overline{T_\circ^{(d_1)} \cdots T_\circ^{(d_\delta)}}} \qquad \text{with } \eta(x,y) \text{ smooth} \,.$$

The quotient of the two monomials in this equation is referred to as a *simple fraction*. A simple fraction can be given a quantitative meaning by considering one-dimensional integrals along curves which cross the light cone transversely away from the origin $\xi = 0$. This procedure is called *weak evaluation on the light cone*. For our purpose, it suffices to integrate over the time coordinate $t = \xi^0$ for fixed $\vec{\xi} \neq 0$. light cone $t \approx |\vec{\xi}|$. The resulting integrals, which in the leading degree are expressions of the form

$$\int_{|\vec{\xi}|-\varepsilon}^{|\vec{\xi}|+\varepsilon} dt\; \eta(t, \vec{\xi})\, \frac{T_\circ^{(a_1)} \cdots T_\circ^{(a_\alpha)}\, \overline{T_\circ^{(b_1)} \cdots T_\circ^{(b_\beta)}}}{T_\circ^{(c_1)} \cdots T_\circ^{(c_\gamma)}\, \overline{T_\circ^{(d_1)} \cdots T_\circ^{(d_\delta)}}} \,, \tag{2.5.45}$$

diverge if the regularization is removed. In view of the scalings $\varepsilon \ll |\vec{\xi}| \ll \ell_{macro}$, it is natural to neglect in the resulting expressions error terms of the form

$$\cdots + (\text{higher orders in } \varepsilon/|\vec{\xi}|)$$

as well as

$$\cdots + (\text{higher orders in } \varepsilon/\ell_{\text{macro}}) \,,$$

where ℓ_{macro} denotes the "macroscopic" length scale on which η varies. Moreover, for large values of $|\vec{\xi}| \gg \ell_{macro}$, the resulting expansions are of "rapid decay" in the "$l$arge spacetime variable l" (for details see [**59**, Section 2.4]), whereas expressions for small values $|\vec{\xi}|$ describe pure "regularization effects" which should not be required in order to furnish the connection to conventional "macroscopic" physics. In this way, the weak evaluation on the light cone allows to extract information of composite expressions for values of $|\vec{\xi}|$ in the scaling $\varepsilon \ll |\vec{\xi}| \ll \ell_{macro}$ (see [**59**, eq. (3.7.77)]).

The method of weak evaluation will be useful in order to analyze the Euler-Lagrange equations corresponding to a minimizer of the causal action, which we present in the following subsection. A detailed analysis of these equations will eventually reveal the desired connection to physics.

2.5.7. The Euler-Lagrange Equations. The derivation of the Euler-Lagrange (EL) equations outlined in this subsection (for details see [**59**, Section 3.2]) is motivated by the variational principle (2.4.7)–(2.4.8) (cf. [**53**]). For this reason, variations of a minimizing fermionic projector are considered. For simplicity, in analogy to [**59**, eq. (2.1.3)] let us consider an operator $P : C_0^\infty(\mathscr{M}, S\mathscr{M}) \to C(\mathscr{M}, S\mathscr{M})$ of the form

$$(P\psi)(x) = \int_{\mathscr{M}} P(x,y)\, \psi(y)\, d^4y \tag{2.5.46}$$

with an integral kernel $P(x,y)$, where $\mathscr{M}$ denotes Minkowski space and $C_0^\infty(\mathscr{M}, S\mathscr{M})$ is the set of smooth wave functions with compact support. Furthermore, we shall assume that P is symmetric with respect to the Lorentz invariant inner product

$$<\psi \mid \phi> = \int_{\mathscr{M}} \overline{\psi(x)}\, \phi(x)\, d^4x \tag{2.5.47}$$

(where $\overline{\psi} \equiv \psi^\dagger \gamma^0$ is the usual adjoint spinor, and $\psi^\dagger$ the complex conjugate spinor). We then refer to P as the *fermionic projector*,[21] and the vectors in the image of P are called *physical wave functions*. Moreover, we introduce the *closed chain* $A_{xy} = A_{xy}[P]$ by

$$A_{xy} = P(x,y)\, P(y,x) \qquad \text{for all } x, y \in \mathscr{M}$$

and define the *spectral weight* $|A|$ by

$$|A| = \sum_{i=1}^{4} |\lambda_i| \,,$$

where $\lambda_1, \ldots, \lambda_4$ are the eigenvalues of A counted with algebraic multiplicities. Then for any $x, y \in \mathscr{M}$ the *Lagrangian* $\mathcal{L}$ is given by

$$\mathcal{L}_{xy}[P] = |A_{xy}^2| - \frac{1}{4}\, |A_{xy}|^2 \,,$$

and the functionals (2.1.1), (2.1.4) formally read (cf. [**59**, Section 3.2])

$$\begin{aligned} \mathcal{S}[P] &\overset{\text{formally}}{=} \iint_{\mathscr{M}\times\mathscr{M}} \mathcal{L}_{xy}[P]\, d^4x\, d^4y \,, \\ \mathcal{T}[P] &\overset{\text{formally}}{=} \iint_{\mathscr{M}\times\mathscr{M}} |A_{xy}|^2\, d^4x\, d^4y \,. \end{aligned} \tag{2.5.48}$$

These expressions are only formal because the integrands need not decay for large x or y, and thus the integrals may be infinite. The functional $\mathcal{S}$ is called the *causal action*.

In order to clarify the class in which to minimize the causal action, let us point out that in [**53**, Section 3.5] it is argued to only consider continuous variations of P of the form UPU^{-1} with unitary transformations U. In order to avoid divergences in (2.5.47) and (2.5.48), we shall only consider such unitary variations U with respect to the inner product (2.5.47) which outside of a compact set coincide with the identity. Variations of this kind are referred to as *unitary in a compact region* (see [**59**, Definition 3.2.1]). According to [**59**, Definition 3.2.2], a fermionic projector P of the form (2.5.46) is said to be a *minimizer* of the variational principle

$$\text{minimize } \mathcal{S} \text{ for fixed } \mathcal{T} \tag{2.5.49}$$

[21] Note that, with a slight abuse of language, sometimes the kernel $P(x,y)$ is also referred to as "fermionic projector."

if for any operator U which is unitary in a compact region and satisfies the constraint

$$\int_{\mathscr{M}} d^4x \int_{\mathscr{M}} d^4y \left(\big|A_{xy}[P]\big|^2 - \big|A_{xy}[UPU^{-1}]\big|^2 \right) = 0\,, \tag{2.5.50}$$

the functional $\mathcal{S}$ satisfies the inequality

$$\int_{\mathscr{M}} d^4x \int_{\mathscr{M}} d^4y \left(\mathcal{L}_{xy}[UPU^{-1}] - \mathcal{L}_{xy}[P] \right) \;\geq\; 0\,.$$

In [**59**], the connection to conventional physics is established essentially by means of the Euler-Lagrange (EL) equations corresponding to a minimizer of the causal action principle (2.5.49). To this end, the existence of a family $(P^\varepsilon)_{\varepsilon>0}$ of regularized vacuum minimizers is taken for granted (see [**59**, Assumption 3.3.1]). In order to derive the corresponding Euler-Lagrange (EL) equations, it is convenient to treat the side condition (2.5.50) with a Lagrange multiplier μ (for details and a justification of this procedure we refer to [**59**, §3.5.2] and [**12**]). More precisely, introducing the regularized closed chain $A^\varepsilon_{xy} = P^\varepsilon(x,y)P^\varepsilon(y,x)$ as well as the functional

$$\mathcal{S}_\mu[P^\varepsilon] \overset{\text{formally}}{=} \iint_{\mathscr{M}\times\mathscr{M}} \mathcal{L}_\mu[A^\varepsilon_{xy}]\, d^4x\, d^4y \qquad \text{with} \qquad \mathcal{L}_\mu[A^\varepsilon_{xy}] = |(A^\varepsilon_{xy})^2| - \mu |A^\varepsilon_{xy}|^2\,,$$

unitary variations of P in a compact region according to [**59**, Definition 3.2.1] give rise to the Euler-Lagrange (EL) equations

$$\boxed{[P^\varepsilon, Q^\varepsilon] = 0\,,} \tag{2.5.51}$$

where Q^ε is an operator with integral kernel

$$Q^\varepsilon(x,y) = \frac{1}{4}\big(R^\varepsilon(x,y) + R^\varepsilon(y,x)^*\big)\,, \tag{2.5.52}$$

and $R^\varepsilon(y,x)$ is given by $R^\varepsilon(y,x) := \nabla \mathcal{L}_\mu[A^\varepsilon_{xy}]$ with (cf. [**59**, eq. (3.5.15)])

$$(\nabla \mathcal{L}_\mu)^\alpha_\beta := \frac{\partial \mathcal{L}_\mu}{\partial \operatorname{Re} P(x,y)^\beta_\alpha} - i \frac{\partial \mathcal{L}_\mu}{\partial \operatorname{Im} P(x,y)^\beta_\alpha} \qquad \text{for } \alpha, \beta \in \{1, \ldots, 4\}\,.$$

Employing the method of *testing on null lines* (for details see [**59**, Appendix A]), the EL equations (2.5.51) reduce to (cf. [**59**, eq. (3.5.29)])

$$\boxed{Q^\varepsilon(x,y) = 0 \quad \text{if evaluated weakly on the light cone}\,.} \tag{2.5.53}$$

We refer to (2.5.53) as the *Euler-Lagrange equations in the continuum limit*. The crucial point in what follows is to observe that the EL equations (2.5.51) pose conditions on the regularized fermionic projector $P^\varepsilon(x,y)$ and in this way on the interaction $\mathcal{B}$ entering the fermionic projector. In a sufficiently general model, the connection to conventional physics will be obtained by means of EL equations of this form. For a generalization to several generations we refer to [**59**, Section 3.4].

In order to analyze the EL equations (2.5.53), it seems reasonable and sufficiently general to assume that the regularized fermionic projector of the vacuum $P^\varepsilon(x,y)$ is *homogeneous* in the sense that $P^\varepsilon(x,y) = P^\varepsilon(y-x)$ for all $x, y \in \mathscr{M}$ and that P^ε has a *vector-scalar structure*, meaning that it is of the form (cf. [**59**, eq. (2.6.2)])

$$P^\varepsilon(x,y) = g_j(x,y)\, \gamma^j + h(x,y)\, \mathbb{1}$$

with appropriate smooth functions g_j and h. We then introduce the closed chain by

$$A^\varepsilon_{xy} = P^\varepsilon(x,y)\, P^\varepsilon(y,x)$$

in view of (2.1.3). Omitting the superscript "ε" and considering g generations, in the formalism of the continuum limit the quantities (cf. [**59**, eqs. (2.6.15) and (3.6.5)])

$$\lambda_+^{xy} = g^2 T_{[0]}^{(0)}\, \overline{T_{[0]}^{(-1)}} + (\deg < 3)\,, \qquad \lambda_-^{xy} = g^2 T_{[0]}^{(-1)}\, \overline{T_{[0]}^{(0)}} + (\deg < 3)\,. \tag{2.5.54}$$

can be interpreted as the eigenvalues of the closed chain, where the bracket $(\deg < 3)$ stands for terms of degree at most two. The corresponding spectral projectors read (see [**59**, eqs. (2.6.16) and (3.6.6)] as well as [**53**, eq. (5.3.21)])

$$F_\pm^{xy} = \frac{1}{2}\Big(\mathbb{1} \pm \frac{[\not\xi, \overline{\not\xi}]}{z - \overline{z}}\Big) + \not\xi(\deg \leq 0) + (\deg < 0) \tag{2.5.55}$$

(where the functions z and $\overline{z}$ are given in terms of the so-called "contraction rules," see [**59**, eqs. (2.4.44)–(2.4.46)]). The corresponding regularized fermionic projector is given by (cf. [**59**, eq. (3.6.2)])

$$P(x,y) = \frac{ig}{2}\, \not\xi\, T_{[0]}^{(-1)} + (\deg < 2)\,,$$

where the indices $\overset{(-1)}{[0]}$ of the factor $\not\xi$ are omitted for notational convenience. In view of [**59**, Proposition 3.6.1], the operator Q as defined by (2.5.52) takes the form

$$Q(x,y) = i\not\xi\, g^3\, (1 - 4\mu)\, T_{[0]}^{(0)} T_{[0]}^{(-1)}\, \overline{T_{[0]}^{(-1)}} + (\deg < 5)\,. \tag{2.5.56}$$

Although only formal, the underlying calculations have a well-defined meaning in the formalism of the continuum limit, because to the resulting formulas one can apply the weak evaluation formula (2.5.45). In view of (2.5.56) we conclude that the degree of the leading singularity of $Q(x,y)$ is five. Analyzing the EL equations (2.5.53) degree by degree, the connection to conventional "macroscopic" physics can be accomplished as we now explain.

2.5.8. Derivation of Classical Field Equations. In a few words, the result of the detailed analysis carried out in [**59**, Chapter 3] shows that the EL equations (2.5.53) can be satisfied to degree five on the light cone in the case of a single charged sector. More precisely, it is useful to rewrite a chiral potential $\mathcal{B}$ in the form

$$\mathcal{B} = \not{A}_\mathrm{v} + \Gamma \not{A}_\mathrm{a}$$

with a *vector potential* A_v and an *axial potential* A_a defined by

$$A_\mathrm{v} = (A_L + A_R)/2 \qquad \text{and} \qquad A_\mathrm{a} = (A_L - A_R)/2\,.$$

It is found that the eigenvalues (λ_s^c) of the closed chain and the corresponding spectral projectors F_s^c with $c \in \{L, R\}$ and $s \in \{+, -\}$ are given by (cf. [**59**, eq. (3.6.21)])

$$\lambda_\pm^{L/R} = \nu_{L/R}\, \lambda_\pm \qquad \text{and} \qquad F_\pm^{L/R} = \chi_{L/R}\, F_\pm$$

with λ_s and F_s as in (2.5.54) and (2.5.55), where (cf. [**59**, eq. (3.6.20)])

$$\nu_L = \overline{\nu_R} = \exp\Big(- 2i \int_x^y A_\mathrm{a}^j\, \xi_j \Big)\,.$$

The fact that all eigenvalues have the same absolute value (cf. [**59**, eq. (3.6.22)]) implies that the EL equations in the continuum limit (2.5.53) are satisfied to degree five on the light cone (cf. [**59**, Section 3.6]).

Thus the task is to deal with the EL equations (2.5.53) to the next lower degree four on the light cone in the case of a single sector (cf. [**59**, Section 3.6]). Introducing the *axial (bosonic) current* by (cf. [**59**, eq. (3.7.4)])

$$j_{\mathrm{a}}^k = \partial_j^k A_{\mathrm{a}}^j - \Box A_{\mathrm{a}}^k$$

as well as the *axial Dirac current* by (cf. [**59**, eq. (3.7.8)])

$$J_{\mathrm{a}}^i = \sum_{k=1}^{n_{\mathrm{p}}} \overline{\psi_k} \Gamma \gamma^i \psi_k - \sum_{l=1}^{n_{\mathrm{a}}} \overline{\phi_l} \Gamma \gamma^i \phi_l \,,$$

it turns out that the Euler-Lagrange equations (2.5.53) give rise to the equation

$$\xi_k \left(j_{\mathrm{a}}^k \, N_1 - m^2 A_{\mathrm{a}}^k \, N_2 - J_{\mathrm{a}}^k \, N_3 \right) = 0 \,,$$

where N_1, N_2 and N_3 are simple fractions which are defined by [**59**, eqs. (3.7.5), (3.7.6) and (3.7.9)], respectively. Evaluating weakly on the light cone, by contrast to N_3 the simple fractions N_1 and N_2 give rise to *logarithmic poles on the light cone*. In order to obtain physically reasonable equations, one must get rid of these logarithmic poles. In short, this is achieved by means of a so-called *microlocal chiral transformation* (we refer the interested reader to [**59**, Section 3.7]). After this has been accomplished, the EL equations to degree four give rise to the desired field equations, including several "correction terms" (see [**59**, Theorem 3.8.2]). Dropping one of these correction terms, one obtains the Dirac-Maxwell equations in [**59**, eq. (3.8.23)].

Unfortunately, the situation is more complicated for a system involving neutrinos (see [**59**, Section 4.3]). Nevertheless, employing the ansatz of "massive neutrinos" [**59**, eq. (4.1.8)], in the setting of one charged and one neutrino sector one gets the following results: the chiral gauge potentials satisfy a classical Yang-Mills-type equation, coupled to the fermions (cf. [**59**, eq. (4.1.10) and Theorem 4.8.1]). The chiral potentials are of the form (cf. [**59**, eq. (4.8.2)])

$$\mathcal{B} = \chi_R \begin{pmatrix} \mathcal{A}_L^{11} & \mathcal{A}_L^{12}\, U_{\mathrm{MNS}}^* \\ \mathcal{A}_L^{21}\, U_{\mathrm{MNS}} & -\mathcal{A}_L^{11} \end{pmatrix} + \chi_L \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

with left-handed $\mathfrak{su}(2)$-valued gauge potentials and a unitary matrix $U_{\mathrm{MNS}} \in \mathrm{U}(3)$, which in analogy to the Standard Model is referred to as MNS matrix. Moreover, the setting allows for describing a gravitational field by the Einstein equations in [**59**, eq. (4.1.11)].

These results are obtained for a system of one charged and one neutrino sector (including right-handed neutrinos). In this case, the auxiliary fermionic projector is given in analogy to (2.5.14)–(2.5.15) with *one* charged sector instead of seven and comprises seven direct summands. In analogy to (2.5.16) and (2.5.17), the chiral asymmetry matrix X and the mass matrix Y are given with *one* charged sector instead of seven as well. The regularization is built in by the formal replacements (cf. [**59**, eqs. (4.2.49)–(4.2.50)])

$$m^p \, T^{(n)} \to m^p \, T_{[p]}^{(n)} \,,$$
$$\tau_{\mathrm{reg}} \, T^{(n)} \to \tau_{\mathrm{reg}} \sum_{k=0}^{\infty} \frac{1}{k!} \, \frac{1}{\delta^{2k}} T_{[R,2n]}^{(k+n)} \,,$$

where the factor δ^{-2n} is required in order to get the scaling dimensions right, while the subscript "R" indicates the regularization of the right-handed component (for details

see [**59**, Section 4.2]). The scaling of the length scale $\delta \gg \varepsilon$ will be specified in (2.5.60) below, whereas the parameter τ_{reg} is chosen in accordance with [**59**, eq. (4.3.34)],

$$\tau_{\text{reg}} = (m\varepsilon)^{p_{\text{reg}}} \qquad \text{with} \qquad 0 < p_{\text{reg}} < 2\,.$$

Next, particles and anti-particles are introduced by occupying additional states or by removing states from the sea in the fashion of (2.5.44). The regularized interacting auxiliary fermionic projector $\tilde{P}^{\text{aux}}$ is then obtained in the fashion of §2.5.4. Finally, the regularized fermionic projector $\tilde{P}$ is introduced by forming the *sectorial projection* (cf. [**59**, eq. (4.2.51)])

$$(\tilde{P})^i_j = \sum_{\alpha,\beta} (\tilde{P}^{\text{aux}})^{(i,\alpha)}_{(j,\beta)}\,,$$

where $i, j \in \{1, 2\}$ is the sector index, and the generation indices α and β take the values $\alpha, \beta \in \{1, \ldots 4\}$ if $i = 1$ and $\alpha, \beta \in \{1, 2, 3\}$ if $i = 2$.[22]

The basic procedure is to first prove that in the presence of a chiral potential $\mathcal{B}$ the EL equations in the continuum limit to degree five on the light cone can be satisfied. To this end, one needs to pose certain conditions on the regularization. More precisely, introducing

$$L^{(n)}_{[p]} = T^{(n)}_{[p]} + \frac{1}{3}\,\tau_{\text{reg}}\, T^{(n)}_{[R,p]}\,,$$

the regularization is required to satisfy the condition (cf. [**59**, eq. (4.3.35)])

$$T^{(n)}_{[p]} = L^{(n)}_{[p]}\,\big(1 + \mathcal{O}\big((m\varepsilon)^{p_{\text{reg}}}\big)\big) \qquad \text{pointwise}\,.$$

Following the arguments in [**59**, §4.3.2], one obtains the following two cases:

$$\textbf{(i)} \quad |\vec{\xi}| \gg \frac{(m\varepsilon)^{p_{\text{reg}}}}{\|A^{12}_L\| + \|A^{21}_L\|}\,, \qquad\qquad \textbf{(ii)} \quad |\vec{\xi}| \ll \frac{(m\varepsilon)^{p_{\text{reg}}}}{\|A^{12}_L\| + \|A^{21}_L\|}\,. \tag{2.5.57}$$

It is found that in both cases, the EL equations can be satisfied to degree five on the light cone (see [**59**, Section 4.3]). However, the analysis in [**59**, Chapter 5] indicates that Case **(ii)** in (2.5.57) is of physical relevance, which is therefore assumed in what follows.

In order to allow for left-handed gauge fields in the neutrino sector, the simplest method is to pose the following additional condition on the regularization functions,

$$\big|L^{(0)}_{[0]}\big| = \big|T^{(0)}_{[0]}\big|\,\big(1 + \mathcal{O}\big((m\varepsilon)^{2p_{\text{reg}}}\big)\big) \qquad \text{pointwise}\,. \tag{2.5.58}$$

Then according to [**59**, Proposition 4.6.8] the EL equations to degree four on the light cone can be satisfied only if (2.5.57) and (2.5.58) are satisfied. Furthermore, in order to compensate the logarithmic poles on the light cone, one again employs a microlocal chiral transformation. It turns out that the logarithmic poles can be compensated if the condition in [**59**, eq. (4.4.64)] holds (see [**59**, Proposition 4.4.6]). For the relation of this condition to the neutrino masses we refer to [**59**, Remark 4.4.9].

Finally, in order to derive the Einstein equations, in the case $n = -1$ and $p = 0$ one needs to pose the following condition on the regularization,

$$\big|L^{(-1)}_{[0]}\big| = \big|T^{(-1)}_{[0]}\big|\,\big(1 + \mathcal{O}\big((m\varepsilon)^{2p_{\text{reg}}}\big)\big) \qquad \text{pointwise}\,. \tag{2.5.59}$$

The connection to the Einstein field equations is then obtained in [**59**, Theorem 4.9.3] provided that (2.5.58) and (2.5.59) hold. Nevertheless, this connection becomes more

[22] Forming the sectorial projection is necessary in order to obtain the correct size of the effective gauge group; concerning further details see §2.5.9 below.

clear in [**59**, Chapter 5]. Introducing the Einstein-Hilbert Lagrangian $\mathcal{L}_{\text{EH}}$ according to [**59**, Theorem 5.4.4], then under the assumptions (2.5.57), (2.5.58) and

$$\varepsilon \ll \delta \ll \frac{1}{m}(m\varepsilon)^{p_{\text{reg}}/2}\,, \tag{2.5.60}$$

the EL equations to degree four can be expressed in terms of the effective action

$$\mathcal{S}_{\text{eff}} = \int_{\mathscr{M}} (\mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{EH}}) \sqrt{-\det g}\, d^4x \tag{2.5.61}$$

by suitably choosing the parameter τ in the Dirac Lagrangian $\mathcal{L}_{\text{Dirac}}$ in [**59**, eq. (5.4.3)] (and a Yang-Mills-type Lagrangian $\mathcal{L}_{\text{YM}}$ of the form [**59**, eqs. (5.4.8)–(5.4.10)]). Then variations of the effective action (2.5.61) with respect to the Lorentzian metric g yield the Einstein field equations (also cf. [**59**, eqs. (4.7.16) and (4.7.17)]), where the energy-momentum tensor is taken into account according to [**59**, §5.4.3]. On the other hand, varying the effective action (2.5.61) with respect to the gauge potential, one obtains the bosonic field equations (see [**59**, §5.4.1]). Note that all variations are carried out under the constraint that the Dirac equation [**59**, eq. (5.4.2)] is satisfied.

Following the arguments in [**59**, §3.9.4], it is expected that the EL equations to degree three or lower on the light cone do *not* give rise to dynamical field equations. For a discussion of *nonlocal* potentials we refer to [**59**, Section 3.10].

2.5.9. Connection to the Standard Model. In order to furnish the connection to the Standard Model, we are led to consider the auxiliary fermionic projector as outlined in §2.5.4. The interaction by chiral potentials is introduced by inserting the following operator into the Dirac equation (for details see [**59**, Section 5.2]),

$$\mathcal{B} = \chi_L \mathcal{A}_R + \chi_R \mathcal{A}_L\,,$$

where A_L^j and A_R^j are Hermitian 25×25-matrices acting on the sectors (implying that $\mathcal{B}$ is a symmetric operator). However, in order to obtain unitary phase transformations, one needs to assume that $\mathcal{B}$ is a unitary operator. Bearing this in mind, the chiral gauge potentials can a-priori be chosen according to the gauge group

$$\mathrm{U}(25)_L \times \mathrm{U}(25)_R\,.$$

This gauge group, however, is too large for mathematical and physical reasons. On the one hand, the causality compatibility condition (2.5.19) gives rise to the smaller gauge group

$$\mathrm{U}(24)_L \times \mathrm{U}(24)_R \times \mathrm{U}(1)_R\,,$$

where the group U(24) acts on the first three direct summands of P_{aux}^N and on the 21 direct summands in P_{aux}^C in (2.5.15). On the other hand, taking into consideration the corresponding regularized interacting auxiliary fermionic projector $\tilde{P}^{\text{aux}}$ (for details we refer to [**59**, Appendix F] and §2.5.4), in order to build in the sectorial projection

$$(\tilde{P})^i_j = \sum_{\alpha,\beta} (\tilde{P}^{\text{aux}})^{(i,\alpha)}_{(j,\beta)}$$

in a compatible way (where $i, j \in \{1, \ldots, 8\}$ is the sector index, and the generation indices α, β take values $\alpha, \beta \in \{1, \ldots 4\}$ if $i = 1$ and $\alpha, \beta \in \{1, 2, 3\}$ if $i = 2, \ldots, 8$), one is led to consider the ansatz [**59**, eq. (5.2.17)]. This gives rise to the gauge group

$$\mathrm{U}(8)_L \times \mathrm{U}(1)_R \times \mathrm{U}(7)_R\,. \tag{2.5.62}$$

As a consequence, whenever the assumptions in [**59**, Theorem 5.3.2] are satisfied, one obtains precisely the effective gauge group of the Standard Model,

$$\mathcal{G} = \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)\,. \tag{2.5.63}$$

For clarity, however, we remark that the assumptions of [**59**, Theorem 5.3.2] can*not* be derived from the causal action principle. Nevertheless, under these assumptions, one recovers the mixing matrix U_{CKM} (which is regarded as the CKM matrix), and in view of [**59**, Proposition 5.4.3] one obtains precise agreement with electroweak theory (for details see [**59**, Theorem 5.4.2]). Let us finally note that the reduction from the large gauge group (2.5.62) to the effective gauge group (2.5.63) and to gauge potentials of the specific form given in [**59**, Theorem 5.3.2] can be regarded as spontaneous breaking of the gauge symmetry; this effect is referred to as "spontaneous block formation."

CHAPTER 3

Causal Variational Principles in the σ-Locally Compact Setting

ABSTRACT. We prove the existence of minimizers of causal variational principles on second-countable, locally compact Hausdorff spaces. Moreover, the corresponding Euler-Lagrange equations are derived. The method is to first prove the existence of minimizers of the causal variational principle restricted to compact subsets for a lower semi-continuous Lagrangian. Exhausting the underlying topological space by compact subsets and rescaling the corresponding minimizers, we obtain a sequence which converges vaguely to a regular Borel measure of possibly infinite total volume. It is shown that, for continuous Lagrangians of compact range, this measure solves the Euler-Lagrange equations. Furthermore, we prove that the constructed measure is a minimizer under variations of compact support. Under additional assumptions, it is proven that this measure is a minimizer under variations of finite volume. We finally extend our results to continuous Lagrangians decaying in entropy.

3.1. Introduction

In the physical theory of causal fermion systems, spacetime and the structures therein are described by a minimizer of the so-called causal action principle (for an introduction to the physical background and the mathematical context, we refer the interested reader to §3.2.1, the textbook [**59**] and the survey articles [**64, 62**]). *Causal variational principles* evolved as a mathematical generalization of the causal action principle [**56, 65**]. The starting point in [**65**] is a smooth manifold $\mathcal{F}$ and a non-negative function

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+ := [0, \infty)$$

(the *Lagrangian*) which is assumed to be lower semi-continuous. The causal variational principle is to minimize the *action* $\mathcal{S}$ defined as the double integral over the Lagrangian

$$\mathcal{S}(\rho) = \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y)$$

under variations of the measure ρ within the class of regular Borel measures, keeping the total volume $\rho(\mathcal{F})$ fixed (*volume constraint*). The aim of this chapter is to extend the existence theory for minimizers of such variational principles to the case that $\mathcal{F}$ is non-compact and the total volume is infinite. Furthermore, we drop the manifold structure of the underlying space $\mathcal{F}$ and consider a σ-locally compact topological space instead. We also work out the corresponding Euler-Lagrange (EL) equations.

In order to put this chapter into the mathematical context, in [**53**] it was proposed to formulate physics by minimizing a new type of variational principle in spacetime. The suggestion in [**53**, Section 3.5] led to the causal action principle in discrete spacetime, which was first analyzed mathematically in [**54**]. A more general and systematic inquiry of causal variational principles on measure spaces was carried out in [**56**]. In this article,

the existence of minimizers is proven in the case that the total volume is finite. In [**65**], the setting is generalized to non-compact manifolds of possibly infinite volume and the corresponding EL equations are analyzed. However, the existence of minimizers is not proved. Here we fill this gap and develop the existence theory in the non-compact setting.

The main difficulty in dealing with measures of infinite total volume is to properly implement the volume constraint. Indeed, the naive prescription $\rho(\mathcal{F}) = \infty$ leaves the freedom to change the total volume by any finite amount, which is not sensible. The way out is to only allow for variations which leave the measure unchanged outside a set of finite volume (so-called *variations of finite volume*; see Definition 3.2.1). In order to prove existence of minimizers within this class, we exhaust $\mathcal{F}$ by compact sets K_n and show that minimizers for the variational principle restricted to each K_n exist. Making essential use of the corresponding EL equations, we rescale the minimizing measures in such a way that a subsequence converges vaguely to a measure ρ on $\mathcal{F}$. We proceed by proving that this measure satisfies the EL equations globally. Finally, we prove that, under suitable assumptions, this measure is even a minimizer under variations of finite volume. This minimizing property is proved in two steps: We first assume that the Lagrangian is of *compact range* (see Definition 3.3.3) and prove that ρ is a minimizer under *variations of compact support* (see Definition 3.4.8 and Theorem 3.4.9). In a second step we extend this result to variations of finite volume (see Definition 3.2.1 and Theorem 3.4.10) under the assumption that property (iv) in §3.2.2 holds, i.e.

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty\,.$$

Sufficient conditions for this assumption to hold are worked out (see Lemma 3.4.7). Finally, we generalize our results to Lagrangians which do not have compact range, but instead have suitable decay properties (see Definition 3.5.1 and Theorem 3.5.9).

The chapter is organized as follows. In Section 3.2 we give a short physical motivation (§3.2.1) and recall the main definitions and existence results as worked out in [**65**] (§3.2.2). In Section 3.3 causal variational principles in the σ-locally compact setting are introduced (§3.3.1), and the existence of minimizers is proved for the causal variational principle restricted to compact subsets, making use of the Banach-Alaoglu theorem and the Riesz representation theorem (§3.3.2). In Section 3.4 minimizers are constructed for continuous Lagrangians of compact range. To this end, in §3.4.1 we exhaust the underlying topological space by compact subsets and take a vague limit of suitably rescaled minimizers thereon to obtain a regular Borel measure on the whole topological space. In §3.4.2 it is shown that this measure satisfies the EL equations. Furthermore, we prove in §3.4.3 that this measure is a minimizer under variations of compact support (see Definition 3.4.8). Finally, in §3.4.4 it is shown that, under additional assumptions, this measure is also a minimizer under variations of finite volume (see Definition 3.3.2). In Section 3.5 we conclude this chapter by weakening the assumption that the Lagrangian is of compact range to Lagrangians which *decay in entropy* (see Definition 3.5.1). Then the EL equations are again satisfied, and under similar additional assumptions as before we prove that the constructed Borel measure is a minimizer of the causal action principle as intended in [**65**].

3.2. Physical Background and Mathematical Preliminaries

3.2.1. Physical Context and Motivation. The purpose of this subsection is to outline a few concepts of causal fermion systems and to explain how this chapter fits into the general physical context and the ongoing research program. The reader not interested in the physical background may skip this section.

The theory of causal fermion systems is a recent approach to fundamental physics. The original motivation was to resolve shortcomings of relativistic quantum field theory. Namely, due to ultraviolet divergences, perturbative quantum field theory is well-defined only after regularization, which is usually understood as a set of prescriptions for how to make divergent integrals finite (e.g. by introducing a suitable "cutoff" in momentum space). The regularization is then removed using the renormalization procedure. However, this concept is not convincing from neither the physical nor the mathematical point of view. More precisely, in view of Heisenberg's uncertainty principle, physicists infer a correspondence between large momenta and small distances. Because of that, the regularization length is often associated to the Planck length $\ell_P \approx 1.6 \cdot 10^{-35}$ m. Accordingly, by introducing an ultraviolet cutoff in momentum space, one disregards distances which are smaller than the Planck length. As a consequence, the microscopic structure of spacetime is completely unknown. Unfortunately, at present there is no consensus on what the correct mathematical model for "Planck scale physics" should be.

The simplest and maybe most natural approach is to assume that on the Planck scale, spacetime is no longer a continuum but becomes in some way "discrete." This is the starting point in the monograph [**53**], where the physical system is described by an ensemble of wave functions in a discrete spacetime. Motivated by the Lagrangian formulation of classical field theory, physical equations are formulated by a variational principle in discrete spacetime. In the meantime, this setting was generalized and developed to the theory of causal fermion systems. It is an essential feature of the approach that spacetime does not enter the variational principle a-priori, but instead it emerges when minimizing the action. Thus causal fermion systems allow for the description of both discrete and continuous spacetime structures.

In order to get the connection to this chapter, let us briefly outline the main structures of causal fermion systems. As initially introduced in [**61**], a *causal fermion system* consists of a triple $(\mathcal{H}, \mathcal{F}, \rho)$ together with an integer $n \in \mathbb{N}$, where $\mathcal{H}$ denotes a complex Hilbert space, $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ being the set of all self-adjoint operators on $\mathcal{H}$ of finite rank with at most n positive and at most n negative eigenvalues, and ρ, referred to as *universal measure*, being a measure on the Borel σ-algebra over $\mathcal{F}$. Then for any $x, y \in \mathcal{F}$, the product xy is an operator of rank at most $2n$. Denoting its non-trivial eigenvalues (counting algebraic multiplicities) by $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy} \in \mathbb{C}$, and introducing the spectral weight $|\,.\,|$ of an operator as the sum of the absolute values of its eigenvalues, the *Lagrangian* can be introduced as a mapping

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+ , \qquad \mathcal{L}(x,y) = \big|(xy)^2\big| - \frac{1}{2n} |xy|^2 \;.$$

As being of relevance for this article, we point out that the Lagrangian is a continuous function which is symmetric in the sense that

$$\mathcal{L}(x,y) = \mathcal{L}(y,x) \qquad \text{for all } x, y \in \mathcal{F} \,.$$

In analogy to classical field theory, one defines the *causal action* by

$$\mathcal{S}(\rho) = \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d\rho(x)\, d\rho(y)\,.$$

Finally, the corresponding *causal action principle* is introduced by varying the measure ρ in the class of a suitable class of Borel measures under additional constraints (which assert the existence of non-trivial minimizers). Given a minimizing measure ρ, *spacetime* M is defined as its support,

$$M := \operatorname{supp} \rho\,.$$

As being outlined in detail in [**59**], critical points of the causal action give rise to Euler-Lagrange (EL) equations, which describe the dynamics of the causal fermion system. In a certain limiting case, the so-called *continuum limit*, one gets a connection to the conventional formulation of physics in a spacetime continuum. In this limiting case, the EL equations give rise to classical field equations like the Maxwell and Einstein equations. Moreover, quantum mechanics is obtained in a limiting case, and close connections to relativistic quantum field theory have been established (see [**58**] and [**63**]).

In order for the causal action principle to be mathematically sensible, the existence theory is of crucial importance. In the case that the dimension of $\mathcal{H}$ is finite, the existence of minimizers was proven in [**56**, Section 2] (based on the existence theory in discrete spacetime [**54**]), giving rise to minimizing measures of finite total volume $\rho(\mathcal{F}) < \infty$. The remaining open problem is to extend the existence theory to the case that $\mathcal{H}$ is infinite-dimensional. Then the total volume $\rho(\mathcal{F})$ is necessarily infinite (for a counter example see [**59**, Exercise 1.3]). Proving existence of minimizers in the resulting *infinite-dimensional setting* (i.e. $\dim \mathcal{H} = \infty$ and $\rho(\mathcal{F}) = \infty$) is a difficult task. Therefore, our strategy is to approach the problem in two steps. The first step is to deal with infinite total volume; it is precisely the objective of this chapter to address this problem in sufficient generality. The second step, which involves the difficulty of dealing with non-locally compact spaces, is currently under investigation.

3.2.2. Causal Variational Principles in the Non-Compact Setting. Before introducing the σ-locally compact setting in Section 3.3, we now recall known results in the slightly less general situation of causal variational principles in the non-compact setting as studied in [**65**, Section 2]. The starting point in [**65**] is a (possibly non-compact) smooth manifold $\mathcal{F}$ of dimension $m \geq 1$. We let ρ be a (positive) measure on the Borel algebra of $\mathcal{F}$ (the *universal measure*). Moreover, let $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ be a non-negative function (the *Lagrangian*) with the following properties:

(i) $\mathcal{L}$ is symmetric, i.e. $\mathcal{L}(x,y) = \mathcal{L}(y,x)$ for all $x, y \in \mathcal{F}$.
(ii) $\mathcal{L}$ is lower semi-continuous, i.e. for all sequences $x_n \to x$ and $y_{n'} \to y$,

$$\mathcal{L}(x,y) \leq \liminf_{n,n'\to\infty} \mathcal{L}(x_n, y_{n'})\,.$$

The *causal variational principle* is to minimize the action

$$\mathcal{S}(\rho) = \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \tag{3.2.1}$$

under variations of the measure ρ, keeping the total volume $\rho(\mathcal{F})$ fixed (*volume constraint*). Here we are interested in the case that the total volume is infinite. In order to implement the volume constraint, we make the following additional assumptions:

(iii) The measure ρ is *locally finite* (meaning that any $x \in \mathcal{F}$ has an open neighborhood $U \subset \mathcal{F}$ with $\rho(U) < \infty$).

(iv) The function $\mathcal{L}(x,.)$ is ρ-integrable for all $x \in \mathcal{F}$ and

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty\,. \tag{3.2.2}$$

By Fatou's lemma, the integral in (3.2.2) is lower semi-continuous in the variable x. A measure on the Borel algebra which satisfies (iii) will be referred to as a *Borel measure*. A Borel measure is said to be *regular* if it is inner and outer regular.

In order to give the causal variational principle (3.2.1) a mathematical meaning, we first note that the difference of two measures $\rho, \tilde{\rho} : \mathcal{B}(\mathcal{F}) \to [0, +\infty]$ with $\rho(\mathcal{F}), \tilde{\rho}(\mathcal{F}) = \infty$ is not a signed measure (due to expressions of the kind "$\infty - \infty$"); as a consequence, the total variation for signed measures does not apply (see e.g. [**139**, §6.1]). For this reason, we introduce the total variation of the difference of two such measures by saying that $|\rho - \tilde{\rho}|(\mathcal{F}) < \infty$ if and only if there exists a Borel set $B \subset \mathcal{F}$ with $\rho(B), \tilde{\rho}(B) < \infty$ and $\rho|_{\mathcal{F}\setminus B} = \tilde{\rho}|_{\mathcal{F}\setminus B}$. In this case, the signed measure $\rho - \tilde{\rho}$ for any Borel set $\Omega \subset \mathcal{F}$ is defined by

$$(\rho - \tilde{\rho})(\Omega) := \rho(\Omega \cap B) - \tilde{\rho}(\Omega \cap B)\,.$$

We then vary in the following class of measures:

Definition 3.2.1. *Given a regular Borel measure ρ on $\mathcal{F}$, a regular Borel measure $\tilde{\rho}$ on $\mathcal{F}$ is said to be a* **variation of finite volume** *if*

$$\big|\tilde{\rho} - \rho\big|(\mathcal{F}) < \infty \qquad \text{and} \qquad \big(\tilde{\rho} - \rho\big)(\mathcal{F}) = 0\,. \tag{3.2.3}$$

Assuming that (i), (ii) and (iv) hold and that $\tilde{\rho}$ is a variation of finite volume, the difference of the actions as given by

$$\begin{aligned} \big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) &= \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) + \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \end{aligned} \tag{3.2.4}$$

is well-defined in view of [**65**, Lemma 2.1].

Definition 3.2.2. *A regular Borel measure ρ for which the above conditions* (i)–(iv) *hold is said to be a* **minimizer** *of the causal action if the difference* (3.2.4) *is non-negative for all regular Borel measures $\tilde{\rho}$ satisfying* (3.2.3)*, i.e.*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

We denote the support of the measure ρ by M,

$$M := \operatorname{supp} \rho = \mathcal{F} \setminus \bigcup \big\{ \Omega \subset \mathcal{F} \;\big|\; \Omega \text{ is open and } \rho(\Omega) = 0 \big\} \tag{3.2.5}$$

(thus the support is the set of all points for which every open neighborhood has a strictly positive measure; for details and generalizations see [**43**, Subsection 2.2.1]).

It is shown in [**65**, Lemma 2.3] (based on a similar result in the compact setting in [**68**, Lemma 3.4]) that a minimizer satisfies the following *Euler-Lagrange (EL) equations*, which state that for a suitable value of the parameter $\mathfrak{s} > 0$, the lower semi-continuous function $\ell : \mathcal{F} \to \mathbb{R}^+_0$ defined by

$$\ell(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) - \mathfrak{s}$$

is minimal and vanishes on the support of ρ,

$$\ell|_M \equiv \inf_{\mathcal{F}} \ell = 0\,. \tag{3.2.6}$$

The parameter $\mathfrak{s}$ can be interpreted as the Lagrange parameter corresponding to the volume constraint. For the derivation of the EL equations and further details we refer to [**65**, Section 2].

3.3. Causal Variational Principles on σ-Locally Compact Spaces

3.3.1. Basic Definitions. In the setup of causal variational principles in the non-compact setting (see §3.2.2) it is assumed that $\mathcal{F}$ is a smooth manifold. Since this manifold structure is not needed in what follows, we now slightly generalize the setting.

DEFINITION 3.3.1. *Let $\mathcal{F}$ be a second-countable, locally compact Hausdorff space, and let the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ be a symmetric and lower semi-continuous function (see conditions* (i) *and* (ii) *in §3.2.2). Moreover, we assume that $\mathcal{L}$ is strictly positive on the diagonal, i.e.*

$$\mathcal{L}(x,x) > 0 \qquad \text{for all } x \in \mathcal{F}\,. \tag{3.3.1}$$

The **causal variational principle on σ-locally compact spaces** *is to minimize the causal action* (3.2.1) *under variations of finite volume (see Definition 3.2.1).*

Note that we do not impose the conditions (iii) and (iv) in §3.2.2. For this reason, it is a-priori not clear whether the integrals in (3.2.4) exist. Therefore, we include this condition into our definition of a minimizer:

DEFINITION 3.3.2. *A regular Borel measure ρ is said to be a* **minimizer** *of the causal action under variations of finite volume if the difference* (3.2.4) *is well-defined and non-negative for all regular Borel measures $\tilde{\rho}$ satisfying* (3.2.3)*,*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

We point out that a minimizer again satisfies the EL equations (3.2.6) (as is proved exactly as in [**65**, Lemma 2.3]). The condition in (3.3.1) is needed in order to avoid trivial minimizers supported at a point where $\mathcal{L}(x,x) = 0$ (see [**68**, Section 1.2]). Moreover, condition (3.3.1) is a reasonable assumption in view of [**59**, Exercise 1.2].

For clarity, we note that, following the conventions in [**76**], by a Borel measure we mean a measure $\rho : \mathcal{B}(\mathcal{F}) \to [0, +\infty]$ on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$ which is *locally finite* (meaning that every point has an open neighborhood of finite volume).[1]

In view of [**9**, Theorem 29.12], every Borel measure on $\mathcal{F}$ is regular (in the sense that the measure of a set can be recovered by approximation with compact sets from inside and with open sets from outside). In particular, it is inner regular and therefore a Radon measure [**143**]. More generally, every Borel measure on a Souslin space is regular by Meyer's theorem (see [**40**, Satz VIII.1.17]).

A topological space which is locally compact and σ-compact is also referred to as being *σ-locally compact* (see for example [**150**]). We note that every second-countable, locally compact Hausdorff space is σ-compact (cf. [**9**, §29]). Therefore, $\mathcal{F}$ is a σ-locally compact space. Moreover, in view of [**72**, Proposition 4.31] and [**161**, Theorem 14.3], the space $\mathcal{F}$ is regular, and hence separable and metrizable by Urysohn's theorem (see for instance [**161**, Theorem 23.1]), where the resulting metric is complete (see [**9**, p. 185]). Thus we can arrange that $\mathcal{F}$ is a Polish space. Since each Polish space is Souslin, any Borel measure on $\mathcal{F}$ is regular, and therefore its support is given by (3.2.5).

[1] We remark that a Borel measure is not usually taken to be locally finite by those working outside of *topological* measure theory (cf. [**76**]).

A metric space X is said to have the *Heine-Borel property* if every closed bounded subset is compact [**162**].[2] In this case, the corresponding metric is referred to as *Heine-Borel metric*. Clearly, every Heine-Borel metric is complete. According to [**162**, Theorem 2'], every σ-locally compact Polish space is metrizable by a Heine-Borel metric. Since the topological space $\mathcal{F}$ is σ-locally compact and Polish we can arrange that bounded sets in $\mathcal{F}$ (with respect to the Heine-Borel metric) are relatively compact, i.e. have compact closure.

Moreover, in order to construct solutions of the EL equations, we first impose the following assumption (see Section 3.4).

DEFINITION 3.3.3. *The Lagrangian has* **compact range** *if for every $K \subset \mathcal{F}$ compact there is a compact set $K' \subset \mathcal{F}$ such that*

$$\mathcal{L}(x,y) = 0 \qquad \text{for all } x \in K \text{ and } y \notin K' .$$

Later on we will show that this assumption can be weakened (see Section 3.5).

3.3.2. Existence of Minimizers on Compact Subsets. Our strategy is to exhaust $\mathcal{F}$ by compact sets, to minimize on each compact set, and to analyze the limit of the resulting measures. In preparation, we now consider the variational principle on a compact subset $K \subset \mathcal{F}$. Since the restriction of a Borel measure (according to [**76**, Definition 2.1]) to K has finite volume, by rescaling we may arrange that the total volume equals one. This leads us to the variational principle

$$\text{minimize} \qquad \mathcal{S}_K(\rho) := \int_K d\rho(x) \int_K d\rho(y)\, \mathcal{L}(x,y)$$

in the class

$$\rho \in \mathfrak{M}_K := \{\text{normalized Borel measures on } K\} ,$$

where *normalized* means that $\rho(K) = 1$ (since we are not concerned with probability theory, in our context it seems preferable to avoid the notion of a probability measure).

Existence of minimizers follows from abstract compactness arguments in the spirit of [**56**, Section 1.2]. We give the proof in detail because the generalization to the lower semi-continuous setting is not quite obvious.

THEOREM 3.3.4. *Let $K \subset \mathcal{F}$ be compact. Moreover, let $(\rho_k)_{k\in\mathbb{N}}$ be a minimizing sequence in $\mathfrak{M}_K$ for the action $\mathcal{S}_K$, i.e.*

$$\lim_{k\to\infty} \mathcal{S}_K(\rho_k) = \inf_{\rho \in \mathfrak{M}_K} \mathcal{S}_K(\rho) .$$

Then the sequence $(\rho_k)_{k\in\mathbb{N}}$ contains a subsequence which converges weakly to a minimizer $\rho_K \in \mathfrak{M}_K$.

PROOF. Let $(\rho_k)_{k\in\mathbb{N}}$ be a minimizing sequence. For clarity, note that the compact subset $K \subset \mathcal{F}$ is a locally compact Hausdorff space. Moreover, the continuous, real-valued functions on K, denoted by $C(K)$, form a normed vector space (with respect to the sup norm $\|\cdot\|_\infty$), and the functions in $C(K)$ are all bounded and have compact support, i.e. $C(K) = C_b(K) = C_c(K)$. For each $k \in \mathbb{N}$, the mapping

$$I_k : C(K) \to \mathbb{R}, \qquad I_k(f) := \int_K f(x)\, d\rho_k(x)$$

[2] In coarse geometry, such metric spaces are also called *proper* (cf. [**136**, Definition 1.4]). For instance, every connected complete Riemannian manifold is a proper metric space (see [**135**, Chapter 2]).

defines a continuous positive linear functional. Since

$$\|I_k\| := \sup_{\substack{f \in C(K) \\ \|f\| \leq 1}} \left| \int_K f(x)\, d\rho_k(x) \right| \leq \|\rho_k\|(K)$$

and $\|\rho_k\|(K) = \rho_k(K) = 1$ for all $k \in \mathbb{N}$ (where $\|\cdot\|(K)$ denotes the total variation, and $\|\cdot\|$ the operator norm on $C(K)^*$), the sequence $(I_k)_{k \in \mathbb{N}}$ is bounded in $C(K)^*$. In view of the Banach-Alaoglu theorem, a subsequence $(I_{k_j})_{j \in \mathbb{N}}$ converges to a linear functional $I \in C(K)^*$ in the weak*-topology,

$$I_{k_j} \rightharpoonup^* I \in C(K)^* .$$

Applying the Riesz representation theorem, we obtain a regular Borel measure ρ_K such that

$$I(f) = \int_K f(x)\, d\rho_K(x) \qquad \text{for all } f \in C(K) .$$

Since $\rho_K(K) = I(1_K) = \lim_{j \to \infty} I_{k_j}(1_K) = 1$ (where 1_K is the function which is identically equal to one), one sees that ρ_K is again normalized.

It remains to show that ρ_K is a minimizer. Since K is compact, σ-compactness of K implies that the measure space $(K, \mathcal{B}(K))$ is σ-finite (according to [**87**, §7]; this also results from the fact that any Borel measure is locally finite and K is second-countable). Due to [**87**, §35, Theorem B], for all $j \in \mathbb{N}$ there is a uniquely determined product measure

$$\eta_{k_j} := \rho_{k_j} \times \rho_{k_j} \colon \mathcal{B}(K) \otimes \mathcal{B}(K) \to \mathbb{R}$$

(see also [**72**, Theorem 7.20]) such that

$$\eta_{k_j}(A \times B) := \rho_{k_j}(A) \cdot \rho_{k_j}(B) ,$$

where $A \times B \in \mathcal{B}(K) \otimes \mathcal{B}(K)$. Since $K \subset \mathcal{F}$ is a second-countable Hausdorff space, it is separable according to [**161**, §5F], and the Cartesian product $K \times K$ is compact (see e.g. [**41**, Theorem 3.2.3]). Moreover, any countable product of second-countable topological spaces is again second-countable and thus separable. By [**13**, Theorem 2.8] we obtain weak convergence

$$\eta_{k_j} = \rho_{k_j} \times \rho_{k_j} \rightharpoonup \rho_K \times \rho_K =: \eta_K .$$

In particular, $(\eta_{k_j})_{j \in \mathbb{N}}$ is a sequence of normalized Borel measures, and η_K is a normalized Borel measure on $K \times K$. Since $K \times K$ is metrizable due to [**122**, §34], and the Lagrangian $\mathcal{L}|_{K \times K} : K \times K \to \mathbb{R}_0^+$ is a measurable non-negative real valued function on $K \times K$, Fatou's lemma for sequences of measures [**44**, eq. (1.5)] yields

$$\begin{aligned}
\mathcal{S}_K(\rho_K) &= \int_K \int_K \mathcal{L}(x,y)\, d\rho_K(x)\, d\rho_K(y) = \iint_{K \times K} \mathcal{L}(x,y)\, d\eta_K(x,y) \\
&\leq \liminf_{j \to \infty} \iint_{K \times K} \mathcal{L}(x,y)\, d\eta_{k_j}(x,y) = \liminf_{j \to \infty} \int_K \int_K \mathcal{L}(x,y)\, d\rho_{k_j}(x)\, d\rho_{k_j}(y) \\
&= \liminf_{j \to \infty} \mathcal{S}_K(\rho_{k_j}) \leq \lim_{j \to \infty} \mathcal{S}_K(\rho_{k_j}) = \inf_{\rho \in \mathfrak{M}_K} \mathcal{S}_K(\rho) .
\end{aligned}$$

Hence ρ_K is a minimizer of the action $\mathcal{S}_K$. □

A minimizing measure $\rho_K \in \mathfrak{M}_K$ satisfies the corresponding Euler-Lagrange equations, which in analogy to (3.2.6) read

$$\ell_K|_{\operatorname{supp} \rho_K} \equiv \inf_K \ell_K = 0 , \tag{3.3.2}$$

where $\ell_K : \mathcal{F} \to \mathbb{R}$ is the function

$$\ell_K(x) := \int_K \mathcal{L}(x,y)\, d\rho_K(y) - \mathfrak{s}\,, \tag{3.3.3}$$

and $\mathfrak{s} > 0$ is a suitably chosen parameter. For clarity, we point out that the integral in (3.3.3) is strictly positive for the following reason: Each normalized measure $\rho \in \mathfrak{M}_K$ has non-empty support; thus for any $x \in \operatorname{supp}\rho \neq \varnothing$, the fact that the Lagrangian is lower semi-continuous as well as strictly positive on the diagonal (3.3.1) implies that there is an open neighborhood U of x such that $\mathcal{L}(x,y) > \mathcal{L}(x,x)/2 > 0$ for all $y \in U$ and $\rho(U) > 0$ in view of (3.2.5). As a consequence,

$$\int_K \mathcal{L}(x,y)\, d\rho(y) \geq \int_U \mathcal{L}(x,y)\, d\rho(y) > 0\,.$$

Comparing this inequality for $x \in K$ with (3.3.2) and (3.3.3), we conclude that the last integral is strictly positive and constant on K.

3.4. Minimizers for Lagrangians of Compact Range

3.4.1. Construction of a Global Borel Measure. Let $(K_n)_{n\in\mathbb{N}}$ be an exhaustion of the σ-locally compact space $\mathcal{F}$ by compact sets such that each compact set is contained in the interior of its successor (see e.g. [**9**, Lemma 29.8]). For every $n \in \mathbb{N}$, we let ρ_{K_n} in $\mathfrak{M}_{K_n}$ be a corresponding minimizer on K_n as constructed in Theorem 3.3.4. We extend these measures by zero to $\mathcal{B}(\mathcal{F})$,

$$\rho^{[n]}(A) := \lambda_n\, \rho_{K_n}(A \cap K_n)\,,$$

where λ_n are positive parameters which will be chosen such that the parameter $\mathfrak{s}$ in the EL equations (3.3.2) and (3.3.3) is equal to one. Thus

$$\ell^{[n]}\big|_{\operatorname{supp}\rho^{[n]}} \equiv \inf_{K_n} \ell^{[n]} = 0\,, \tag{3.4.1}$$

where

$$\ell^{[n]}(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{[n]}(y) - 1\,. \tag{3.4.2}$$

For clarity, we point out that the measures $\rho^{[n]}$ are *not normalized*. More precisely,

$$\rho^{[n]}(\mathcal{F}) = \lambda_n\,,$$

and the sequence $(\lambda_n)_{n\in\mathbb{N}}$ will typically be unbounded.

LEMMA 3.4.1. *For every compact subset $K \subset \mathcal{F}$ there is a constant $C_K > 0$ such that*

$$\rho^{[n]}(K) \leq C_K \qquad \textit{for all } n \in \mathbb{N}\,.$$

PROOF. Since $\mathcal{L}(x,.)$ is lower semi-continuous and strictly positive at x (see (3.3.1)), there is an open neighborhood $U(x)$ of x with

$$\mathcal{L}(y,z) \geq \frac{\mathcal{L}(x,x)}{2} > 0 \qquad \text{for all } y,z \in U(x)\,.$$

Covering K by a finite number of such neighborhoods $U(x_1),\ldots,U(x_L)$, it suffices to show the inequality for the sets $K \cap U(x_\ell)$ for any $\ell \in \{1,\ldots,L\}$. Moreover, we choose N so large that $K_N \supset K$ and fix $n \geq N$. If $K \cap \operatorname{supp}\rho^{[n]} = \varnothing$, there is nothing to prove.

Otherwise, there is a point $z \in K \cap \operatorname{supp} \rho^{[n]}$. Using the EL equations (3.4.1) at z, it follows that

$$1 = \int_{\mathcal{F}} \mathcal{L}(z,y)\, d\rho^{[n]}(y) \geq \int_{U(x_\ell)} \mathcal{L}(z,y)\, d\rho^{[n]}(y) \geq \frac{\mathcal{L}(x_\ell, x_\ell)}{2}\, \rho^{[n]}(U(x_\ell))\,.$$

Hence for any $n \geq N$,

$$\rho^{[n]}(U(x_\ell)) \leq \frac{2}{\mathcal{L}(x_\ell, x_\ell)}\,. \tag{3.4.3}$$

Next, let $c(x_\ell)$ be the maximum of $\rho^{[1]}(U(x_\ell)), \ldots, \rho^{[N-1]}(U(x_\ell))$ and $2/\mathcal{L}(x_\ell, x_\ell)$. Since the open sets $U(x_1), \ldots, U(x_L)$ cover K, we finally introduce the constant C_K as the sum of the constants $c(x_1), \ldots, c(x_L)$. □

Now we proceed as follows. Denoting by $(K_n)_{n\in\mathbb{N}}$ the above exhaustion of $\mathcal{F}$ by compact sets, we first restrict the measures $\rho^{[n]}$ to the compact set K_1. According to Lemma 3.4.1, the resulting sequence of measures is bounded. Therefore, a subsequence converges as a measure on K_1 (using again the Banach-Alaoglu theorem and the Riesz representation theorem). Out of the resulting subsequence $(\rho^{[1,n_k]})_{k\in\mathbb{N}}$, we then choose a subsequence of measures $(\rho^{[2,n_k]})_{k\in\mathbb{N}}$ which converges weakly on K_2. We proceed iteratively and denote the resulting diagonal sequence by

$$\rho^{(k)} := \rho^{[k,n_k]} \qquad \text{for all } k \in \mathbb{N}\,. \tag{3.4.4}$$

In the following, we restrict attention to the compact exhaustion $(K_m)_{m\in\mathbb{N}}$, where for convenience by K_m we denote the sets K_{n_m} for $m \in \mathbb{N}$ (thus $\rho^{(m)}$ is a minimizer on K_m for each $m \in \mathbb{N}$). By construction, the sequence $(\rho^{(k)}|_{K_m})_{k\in\mathbb{N}}$ converges weakly to some measure ρ_{K_m} for every $m \in \mathbb{N}$, i.e.

$$\rho^{(k)}|_{K_m} \rightharpoonup \rho_{K_m} \qquad \text{for all } m \in \mathbb{N}\,. \tag{3.4.5}$$

In particular,

$$\lim_{k\to\infty} \rho^{(k)}(K_m) = \rho_{K_m}(K_m) \qquad \text{for all } m \in \mathbb{N}\,.$$

Denoting the interior of K_m by K_m°, we claim that for all $n \geq m$,

$$\int_{K_m} f\, d\rho_{K_m} = \int_{K_n} f\, d\rho_{K_n} \qquad \text{for all } f \in C_c(K_m^\circ)\,. \tag{3.4.6}$$

Namely, in view of $C_c(K_m^\circ) \subset C_b(K_m) \cap C_b(K_n)$, weak convergence (3.4.5) yields

$$\int_{K_m} f\, d\rho_{K_m} = \lim_{k\to\infty} \int_{K_m} f\, d\rho^{(k)}|_{K_m} \overset{(\star)}{=} \lim_{k\to\infty} \int_{K_n} f\, d\rho^{(k)}|_{K_n} = \int_{K_n} f\, d\rho_{K_n}$$

for all $f \in C_c(K_m^\circ)$, where in $(\star)$ we made use of the fact that $\operatorname{supp} f \subset K_m^\circ \subset K_n$.

In order to construct a global measure ρ on $\mathcal{F}$, we proceed by introducing the functional

$$I : C_c(\mathcal{F}) \to \mathbb{R}\,, \qquad f \mapsto I(f) := \lim_{n\to\infty} \int_{\mathcal{F}} f(x)\, d\rho_{K_n}(x)\,.$$

Note that, in view of (3.4.6), the last integral is independent of $n \in \mathbb{N}$ for sufficiently large integers $n \in \mathbb{N}$, and thus well-defined. Thus $I : C_c(\mathcal{F}) \to \mathbb{R}$ defines a positive functional. Making use of the fact that $\mathcal{F}$ is locally compact, the Riesz representation theorem [**40**, Darstellungssatz VIII.2.5] or [**9**, Theorem 29.1 and Theorem 29.6] yields the existence of

a uniquely determined Radon measure $\rho : \mathcal{B}(\mathcal{F}) \to [0,\infty]$ (i.e. an inner regular, locally finite measure [**143**]) such that

$$I(f) = \int_{\mathcal{F}} f \, d\rho \qquad \text{for all } f \in C_c(\mathcal{F}) \tag{3.4.7}$$

(also see [**36**, Definition 4 and Theorem 5]). From the equality

$$\lim_{n\to\infty} \int_{\mathcal{F}} f(x)\, d\rho_{K_n}(x) = I(f) = \int_{\mathcal{F}} f \, d\rho \qquad \text{for all } f \in C_c(\mathcal{F}) \tag{3.4.8}$$

we conclude that the sequence $(\rho_{K_n})_{n\in\mathbb{N}}$ converges vaguely to the measure ρ (for details see [**9**, Definition 30.1]). Moreover, given $f \in C_c(\mathcal{F})$, we know that $\operatorname{supp} f \subset K_m^\circ$ for sufficiently large $m \in \mathbb{N}$. Hence from vague convergence (3.4.8) and weak convergence (3.4.5) we conclude that for any $\varepsilon > 0$ there exists $m' \in \mathbb{N}$ such that for all $m \geq m'$,

$$\left| \int_{\mathcal{F}} f \, d\rho - \int_{\mathcal{F}} f \, d\rho^{(k)} \right| \leq \left| \int_{\mathcal{F}} f \, d\rho - \int_{\mathcal{F}} f \, d\rho_{K_m} \right| + \left| \int_{\mathcal{F}} f \, d\rho_{K_m} - \int_{\mathcal{F}} f \, d\rho^{(k)}|_{K_m} \right| < \varepsilon \,.$$

Since $f \in C_c(\mathcal{F})$ was arbitrary, we also conclude that

$$\rho^{(k)} \xrightarrow{v} \rho \qquad \text{vaguely} \,. \tag{3.4.9}$$

Regularity follows from Ulam's theorem [**40**, Satz VIII.1.16] and the fact that $\mathcal{F}$ is Polish. In Appendix A.1 (see Lemma A.1.1) it is shown that the measure ρ given by (3.4.7) is non-zero and possibly has infinite total volume.

Similar to (3.4.2), we introduce the notation

$$\ell^{(n)}(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) - 1 \,. \tag{3.4.10}$$

In particular, the following EL equations hold,

$$\ell^{(n)}\big|_{\operatorname{supp}\rho^{(n)}} \equiv \inf_{K_n} \ell^{(n)} = 0 \,. \tag{3.4.11}$$

3.4.2. Derivation of the Euler-Lagrange Equations. In this section, we assume that the Lagrangian $\mathcal{L}$ is continuous and of compact range (see Definition 3.3.3). Our goal is to prove the following result.

THEOREM 3.4.2 (**Euler-Lagrange equations**). *Assume that $\mathcal{L}$ is continuous and of compact range. Then the measure ρ constructed in* (3.4.7) *satisfies the Euler-Lagrange equations*

$$\ell|_{\operatorname{supp}\rho} \equiv \inf_{x\,\in\,\mathcal{F}} \ell(x) = 0 \,, \tag{3.4.12}$$

where $\ell \in C(\mathcal{F})$ is defined by

$$\ell(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) - 1 \,. \tag{3.4.13}$$

In order to prove Theorem 3.4.2 we proceed in several steps. The proof will be completed at the end of this subsection.

LEMMA 3.4.3. *Assume that the sequence of measures $(\rho^{(n)})_{n\in\mathbb{N}}$ converges vaguely to a regular Borel measure $\rho \neq 0$. Then for every $x \in \operatorname{supp}\rho$ there is a sequence $(x_k)_{k\in\mathbb{N}}$ and a subsequence $\rho^{(n_k)}$ such that $x_k \in \operatorname{supp}\rho^{(n_k)}$ for all $k \in \mathbb{N}$ and $x_k \to x$.*

PROOF. Assume conversely that for some $x \in \operatorname{supp} \rho$ there is no such subsequence. Then there is an open neighborhood U of x which does not intersect the support of the measures $\rho^{(n)}$ for almost all $n \in \mathbb{N}$. As a consequence, for every compact neighborhood V of x with $V \subset U$ (which exists by [**72**, Proposition 4.30]) we have

$$\rho^{(n)}(V) = 0 \qquad \text{for almost all } n \in \mathbb{N}\,,$$

and

$$\rho(V) > 0$$

in view of (3.2.5). Since U is an open cover of the compact Hausdorff space V, there exists $f \in C_c(U; [0,1])$ with $f|_V \equiv 1$ (see for instance [**4**, Lemma 2.92]). Thus vague convergence $\rho^{(n)} \overset{v}{\to} \rho$ yields the contradiction

$$0 < \int_{\mathcal{F}} f \, d\rho = \lim_{n\to\infty} \int_{\mathcal{F}} f \, d\rho^{(n)} = 0\,,$$

which proves the claim. □

For notational simplicity, we denote the subsequence $\rho^{(n_k)}$ again by $\rho^{(k)}$.

LEMMA 3.4.4. *Let $\mathcal{L}$ be continuous and of compact range. Then the function ℓ defined by* (3.4.13) *is continuous.*

PROOF. For any $x \in \mathcal{F}$, let $(x_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in $\mathcal{F}$ converging to x, and let U be an open, relatively compact neighborhood of x (which exists by local compactness of $\mathcal{F}$). Since $\mathcal{L}$ is of compact range, there is a compact set $K' \subset \mathcal{F}$ such that $\mathcal{L}(\tilde{x}, y) = 0$ for all $\tilde{x} \in K := \overline{U}$ and $y \notin K'$. Since the sequence $(x_n)_{n\in\mathbb{N}}$ converges to x, there is an integer $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. By continuity of $\mathcal{L}$, the mapping $\mathcal{L} : K \times K' \to \mathbb{R}$ is bounded. Therefore, the functions $\mathcal{L}(x_n, \cdot) : K' \to \mathbb{R}$ are uniformly bounded for all $n \geq N$. Thus Lebesgue's dominated convergence theorem yields

$$\begin{aligned}\ell(x) &= \int_{K'} \mathcal{L}(x,y)\, d\rho(y) - 1 = \int_{K'} \lim_{n\to\infty} \mathcal{L}(x_n, y)\, d\rho(y) - 1 \\ &= \lim_{n\to\infty} \int_{K'} \mathcal{L}(x_n, y)\, d\rho(y) - 1 = \lim_{n\to\infty} \ell(x_n)\,,\end{aligned}$$

proving continuity of ℓ. □

In the next proposition, we show that the sequence $(\ell^{(n)})_{n\in\mathbb{N}}$ converges pointwise to ℓ. Choosing $K = \{x\}$ in Definition 3.3.3, we denote the corresponding compact set K' by K_x, i.e.

$$\mathcal{L}(x,y) = 0 \qquad \text{for all } y \notin K_x\,. \tag{3.4.14}$$

PROPOSITION 3.4.5. *Let $\mathcal{L}$ be continuous and of compact range, and let $(\ell^{(n)})_{n\in\mathbb{N}}$ and ℓ be the functions defined in* (3.4.10) *and* (3.4.13)*, respectively. Then $(\ell^{(n)})_{n\in\mathbb{N}}$ converges pointwise to ℓ, i.e.*

$$\lim_{n\to\infty} \ell^{(n)}(x) = \ell(x) \qquad \textit{for all } x \in \mathcal{F}\,. \tag{3.4.15}$$

PROOF. Let $x \in \mathcal{F}$. Since $\mathcal{L}$ is assumed to be of compact range, using the notation (3.4.14) implies that $\mathcal{L}(x,y) = 0$ for all $y \notin K_x$. We conclude that $\mathcal{L}(x, \cdot) \in C_c(\mathcal{F})$, and thus by vague convergence (3.4.9) we obtain

$$\ell(x) = \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) - 1 = \lim_{n\to\infty} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) - 1 = \lim_{n\to\infty} \ell^{(n)}(x)\,.$$

Since $x \in \mathcal{F}$ is arbitrary, the sequence $(\ell^{(n)})_{n\in\mathbb{N}}$ converges pointwise to ℓ. □

Our proof of Theorem 3.4.2 will be based on equicontinuity of the family $(\ell^{(n)}|_K)_{n\in\mathbb{N}}$ for arbitrary compact subsets $K \subset \mathcal{F}$. We know that the functions $\ell^{(n)}$ are continuous and uniformly bounded on compact sets. However, as can be seen from the example $(f_n)_{n\in\mathbb{N}}$ with

$$f_n : [0,1] \to \mathbb{R}, \qquad f_n(x) = \sin nx \qquad \text{for all } n \in \mathbb{N}\,,$$

these conditions are in general not sufficient to ensure equicontinuity. Nonetheless, the additional assumption that the Lagrangian $\mathcal{L}\colon \mathcal{F}\times\mathcal{F} \to \mathbb{R}_0^+$ is of compact range (see Definition 3.3.3) gives rise to equicontinuity of the family $(\ell^{(n)}|_K)_{n\in\mathbb{N}}$, as the following proposition shows.

PROPOSITION 3.4.6. *Let $\mathcal{L}$ be continuous and of compact range. Then for any compact subset $K \subset \mathcal{F}$, the family $F_K := \{\ell^{(n)}|_K \colon n \in \mathbb{N}\}$ is equicontinuous.*

PROOF. Consider an arbitrary compact set $K \subset \mathcal{F}$. In order to prove equicontinuity of F_K, we have to show that for every $\varepsilon > 0$ and every $x \in K$ there is a corresponding neighborhood $V = V(x)$ of x with

$$\sup_{f\in F_K}\sup_{z\in V} |f(x) - f(z)| < \varepsilon\,.$$

Let $x \in K$ and consider an arbitrary $\varepsilon > 0$. Since $\mathcal{L}$ is of compact range, there is a compact set $K' \subset \mathcal{F}$ such that

$$\mathcal{L}(\tilde{x}, y) = 0 \qquad \text{for all } \tilde{x} \in K \text{ and } y \notin K'\,. \tag{3.4.16}$$

In view of Lemma 3.4.1 there is a positive constant $C_{K'} > 0$ such that $\rho^{(n)}(K') \le C_{K'}$ for all $n \in \mathbb{N}$.

Since $\mathcal{L}$ is continuous and $K \times K'$ is compact, the mapping

$$\mathcal{L}|_{K\times K'} \colon K \times K' \to \mathbb{R}$$

is uniformly continuous. Moreover, in view of (3.4.16), the same is true for $\mathcal{L}|_{K\times\mathcal{F}}$. Hence for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\Big|\mathcal{L}|_{K\times\mathcal{F}}(x_1,y_1) - \mathcal{L}|_{K\times\mathcal{F}}(x_2,y_2)\Big| < \varepsilon \qquad \text{for all } (x_2,y_2) \in B_\delta((x_1,y_1))\,.$$

Choosing $\delta > 0$ such that $|\mathcal{L}|_{K\times\mathcal{F}}(x,\cdot) - \mathcal{L}|_{K\times\mathcal{F}}(z,\cdot)| < \varepsilon/C_{K'}$ for all $z \in B_\delta(x) \cap K$, we obtain

$$\begin{aligned}
&\sup_{n\in\mathbb{N}}\sup_{z\in B_\delta(x)\cap K} \Big|\ell^{(n)}|_K(x) - \ell^{(n)}|_K(z)\Big| \\
&\quad= \sup_{n\in\mathbb{N}}\sup_{z\in B_\delta(x)\cap K} \left|\int_{\mathcal{F}} \Big(\mathcal{L}|_{K\times\mathcal{F}}(x,y) - \mathcal{L}|_{K\times\mathcal{F}}(z,y)\Big)\, d\rho^{(n)}(y)\right| \\
&\quad\le \sup_{n\in\mathbb{N}}\sup_{z\in B_\delta(x)\cap K} \int_{K'} \Big|\mathcal{L}|_{K\times\mathcal{F}}(x,y) - \mathcal{L}|_{K\times\mathcal{F}}(z,y)\Big|\, d\rho^{(n)}(y) \\
&\quad< \sup_{n\in\mathbb{N}} \rho^{(n)}(K')\,\frac{\varepsilon}{C_{K'}} \le \varepsilon\,.
\end{aligned}$$

This yields equicontinuity of F_K as desired. □

After these preparations, we are able to prove Theorem 3.4.2.

PROOF OF THEOREM 3.4.2. Let $(K_n)_{n\in\mathbb{N}}$ be a compact exhaustion of $\mathcal{F}$. Moreover, let $(\rho^{(n)})_{n\in\mathbb{N}}$ be the corresponding sequence of vaguely converging measures according to (3.4.4) such that (3.4.10) and (3.4.11) hold. The main idea of the proof is to make use of pointwise convergence (3.4.15) and equicontinuity of the sequence $(\ell^{(n)}|_K)_{n\in\mathbb{N}}$ for arbitrary compact sets $K \subset \mathcal{F}$ as established in Proposition 3.4.5 and Proposition 3.4.6, respectively.

First of all, application of Proposition 3.4.5 shows that $\ell(x) \geq 0$ for every $x \in \mathcal{F}$. Namely, since $\rho^{(n)}$ is a minimizer of the action $\mathcal{S}_{K_n}$ for every $n \in \mathbb{N}$, and x is contained in all compact sets $(K_n)_{n\geq N}$ for some integer $N = N(x)$, we have

$$\ell(x) \overset{(3.4.15)}{=} \lim_{n\to\infty} \ell^{(n)}(x) = \lim_{n\to\infty} \ell^{(n)}|_{K_n}(x) \overset{(3.4.11)}{\geq} 0 \qquad \text{for all } x \in \mathcal{F}\,. \tag{3.4.17}$$

In order to derive the EL equations (3.4.12), it remains to prove that $\ell(x)$ vanishes for every $x \in \operatorname{supp}\rho$. Since $\mathcal{F}$ is locally compact, each $x \in \operatorname{supp}\rho$ is contained in a compact neighborhood K_x (cf. (3.4.14)) such that $\mathcal{L}(x,.)$ vanishes outside K_x. Due to vague convergence $\rho^{(n)} \to \rho$ as $n \to \infty$ (cf. (3.4.9)), by virtue of Lemma 3.4.3 there exists a sequence $x^{(n)} \to x$ as $n \to \infty$ such that $x^{(n)} \in \operatorname{supp}\rho^{(n)}$ for every $n \in \mathbb{N}$. We choose $N' \in \mathbb{N}$ such that $x^{(n)} \in K_x$ for all $n \geq N'$. For this reason, it suffices to focus on the restriction $\ell|_{K_x}$. Equicontinuity of the family $\{\ell^{(n)}|_{K_x} : n \in \mathbb{N}\}$ (see Proposition 3.4.6) yields

$$\lim_{n\to\infty} \sup_{k\in\mathbb{N}} \left| \ell^{(k)}|_{K_x}\big(x^{(n)}\big) - \ell^{(k)}|_{K_x}(x) \right| = 0\,.$$

Moreover, the expression

$$\lim_{n\to\infty} \left| \ell^{(n)}|_{K_x}(x) - \ell|_{K_x}(x) \right| = 0 \qquad \text{for all } x \in K_x$$

holds in view of pointwise convergence (3.4.15). Taken together, for every $x \in \operatorname{supp}\rho$ we finally obtain

$$\begin{aligned} \lim_{n\to\infty} \left| \ell^{(n)}\big(x^{(n)}\big) - \ell(x) \right| &= \lim_{n\to\infty} \left| \ell^{(n)}|_{K_x}\big(x^{(n)}\big) - \ell|_{K_x}(x) \right| \\ &\leq \lim_{n\to\infty} \left| \ell^{(n)}|_{K_x}\big(x^{(n)}\big) - \ell^{(n)}|_{K_x}(x) \right| + \lim_{n\to\infty} \left| \ell^{(n)}|_{K_x}(x) - \ell|_{K_x}(x) \right| = 0\,. \end{aligned}$$

In view of (3.4.17), the Euler-Lagrange equations (3.4.12) hold due to

$$\ell(x) = \lim_{n\to\infty} \ell^{(n)}\big(x^{(n)}\big) \overset{(3.4.11)}{=} 0 \qquad \text{for all } x \in \operatorname{supp}\rho\,,$$

which completes the proof. □

In the remainder of this subsection, we discuss the properties (iii) and (iv) in §3.2.2. Condition (iii) holds by construction because we are working with locally finite measures (see §3.3.1). Condition (iv) does not hold in general, but it can be checked a-posteriori for a constructed measure ρ. Under suitable assumptions on $\mathcal{L}$, however, this condition can even be verified a-priori, i.e. without knowing ρ. This is exemplified in the following lemma.

LEMMA 3.4.7. *Let $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ be continuous and of compact range. Moreover, assume that the following conditions hold:*

(a) $c := \inf_{x\in\mathcal{F}} \mathcal{L}(x,x) > 0$.
(b) $\sup_{x,y\in\mathcal{F}} \mathcal{L}(x,y) \leq \mathcal{C} < \infty$.

(c) *There is an integer $N > 0$ such that every K_x (as defined in* (3.4.14)*) can be covered by open sets $U_1, \ldots, U_N$ with the property that for all $i \in \{1, \ldots, N\}$,*

$$\mathcal{L}(x,y) > \frac{c}{2} \qquad \text{for all } y \in U_i \,.$$

Then the measure ρ constructed in (3.4.7) *satisfies condition* (iv) *in §3.2.2.*

PROOF. Since $\mathcal{L}$ is continuous and of compact range, we have

$$\int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) = \int_{K_x} \mathcal{L}(x,y)\, d\rho(y) \leq \sup_{y \in K_x} \mathcal{L}(x,y)\, \rho(K_x) < \infty \,,$$

showing that $\mathcal{L}(x,\cdot)$ is ρ-integrable for every $x \in \mathcal{F}$. It remains to prove that

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty \,.$$

Since $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ satisfies (a)–(c), inequality (3.4.3) yields

$$\rho(U_i) \leq \sup_{x \in \mathcal{F}} \frac{2}{\mathcal{L}(x,x)} \leq \frac{2}{c} \qquad \text{for all } i \in \{1, \ldots, N\} \,.$$

Thus we obtain

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) \leq \sup_{x,y \in \mathcal{F}} \mathcal{L}(x,y)\, \rho(K_x) \leq \mathcal{C} \sum_{i=1}^{N} \rho(U_i) < \frac{2\mathcal{C}N}{c} < \infty$$

as desired. □

3.4.3. Existence of Minimizers under Variations of Compact Support. This subsection is intended to prove the existence of minimizers under variations of compact support. Apart from technical convenience, it is natural and sufficient for many applications to restrict attention to variations of compact support (in contrast to the more general variations of finite volume as introduced in Definition 3.2.1 and Definition 3.3.2).

DEFINITION 3.4.8. *A measure $\rho \in \mathfrak{B}_{\mathcal{F}}$ is said to be a* **minimizer under variations of compact support** *of the causal action if for any $\tilde{\rho} \in \mathfrak{B}_{\mathcal{F}}$ which satisfies* (3.2.3) *such that the signed measure $(\rho - \tilde{\rho})$ is compactly supported, the inequality*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0$$

holds.

The goal of this subsection is to prove that the measure ρ constructed in (3.4.7) is a minimizer under variations of compact support. Before stating our result (see Theorem 3.4.9 below), we show that the difference (3.2.4) is well-defined. Indeed, considering variations of compact support, the signed measure $\mu := \tilde{\rho} - \rho$ is compactly supported. Considering its Jordan decomposition $\mu = \mu^+ - \mu^-$ (see e.g. [**87**, §29]), the measures μ^+ and μ^- have compact support. Hence, using that the Lagrangian is continuous,

$$\int_{\mathcal{F}} d\mu^+(x)\, \ell(x) \leq \left(\sup_{x \in \operatorname{supp} \mu^+} \ell(x) \right) \mu^+(\operatorname{supp} \mu^+) < \infty \,,$$

and similarly for μ^-. Now we can proceed as in the proof of [**65**, Lemma 2.1] to conclude that all the integrals in (3.2.4) are well-defined and finite.

THEOREM 3.4.9. *Assume that the Lagrangian $\mathcal{L}$ is continuous and of compact range. Then the measure ρ constructed in* (3.4.7) *is a minimizer under variations of compact support.*

PROOF. Let $\tilde{\rho}$ be a regular Borel measure on $\mathcal{F}$ with $K := \text{supp}(\tilde{\rho} - \rho) \subset \mathcal{F}$ compact such that

$$0 < \tilde{\rho}(K) = \rho(K) < \infty \,.$$

Then $(\tilde{\rho} - \rho)(\mathcal{F}) = 0$, i.e. (3.2.3) is satisfied. Since the Lagrangian is supposed to be of compact range, there exists a compact set $K' \subset \mathcal{F}$ such that $\mathcal{L}(x, y) = 0$ for all $x \in K$ and $y \in \mathcal{F} \setminus K'$.

By regularity of ρ and $\tilde{\rho}$, for arbitrary $\tilde{\varepsilon} > 0$ there exist $U \supset K$ and $U' \supset K'$ open with $U' \supset U$ such that

$$\begin{aligned} \rho(U \setminus K) &< \tilde{\varepsilon} \,, \qquad & \rho(U' \setminus K') &< \tilde{\varepsilon} \,, \\ \tilde{\rho}(U \setminus K) &< \tilde{\varepsilon} \,, \qquad & \tilde{\rho}(U' \setminus K') &< \tilde{\varepsilon} \,. \end{aligned}$$

In particular, $\rho|_{U'}$ is a non-negative finite Borel measure on the topological space U'. Since $\mathcal{F}$ is metrizable, it is completely regular. As a consequence, the class $\Gamma_{\rho|_{U'}}$ of all Borel sets $E \subset U'$ with boundaries of ρ-measure zero contains a base (consisting of open sets) of the topology of U' (see [**16**, Proposition 8.2.8]). Since K and K' are compact, they can be covered by finitely many relatively compact, open sets $V_1, \ldots, V_N \subset U$ and $V'_1, \ldots, V'_{N'} \subset U'$ in $\Gamma_{\rho|_{U'}}$ (cf. [**41**, Section 1.1]). By construction, the closure of the sets $V := \bigcup_{i=1}^N V_i \subset U$ and $V' := \bigcup_{i=1}^{N'} V'_i \subset U'$, denoted by $\overline{V}$ and $\overline{V'}$, respectively, is compact. By choosing $V \subset U$ suitably, we can arrange that $\mathcal{L}(x, y) = 0$ for all $x \in V$ and $y \notin V'$. Moreover, considering the restriction $\rho^{(n)}|_V$, for any $f \in C_c(V)$ we obtain

$$\lim_{n\to\infty} \int_V f \, d\rho^{(n)}|_V = \lim_{n\to\infty} \int_{\mathcal{F}} f \, d\rho^{(n)} \overset{(3.4.9)}{=} \int_{\mathcal{F}} f \, d\rho = \int_V f \, d\rho|_V \,,$$

and similarly for V'. This proves vague convergence

$$\rho^{(n)}|_V \overset{v}{\to} \rho|_V \qquad \text{and} \qquad \rho^{(n)}|_{V'} \overset{v}{\to} \rho|_{V'} \,.$$

Since $\Gamma_{\rho|_{U'}}$ is a subalgebra in $\mathcal{B}(\mathcal{F})$ (see [**16**, Proposition 8.2.8]), the sets V and V' are also contained in $\Gamma_{\rho|_{U'}}$, implying that

$$\rho(\partial V) = 0 = \rho(\partial V') \,.$$

Making use of vague convergence (3.4.9) and applying [**9**, Theorem 30.2], we obtain

$$\rho(V) = \rho(\overline{V}) \geq \limsup_{n\to\infty} \rho^{(n)}(\overline{V}) \geq \limsup_{n\to\infty} \rho^{(n)}(V) \geq \liminf_{n\to\infty} \rho^{(n)}(V) \geq \rho(V) \,,$$

proving that

$$\rho(V) = \lim_{n\to\infty} \rho^{(n)}(V) \tag{3.4.18}$$

(and similarly for V'). We point out that, for each $n \in \mathbb{N}$, the measure $\rho^{(n)}|_V / \rho^{(n)}(V)$ is normalized in the sense of §3.3.2. Moreover, for any $f \in C_c(V)$ we are given

$$\lim_{n\to\infty} \int_V f \, d\rho^{(n)}|_V / \rho^{(n)}(V) = \int_V f \, d\rho|_V / \rho(V) \,.$$

That is, the sequence of normalized measures $(\rho^{(n)}|_V / \rho^{(n)}(V))_{n\in\mathbb{N}}$ converges vaguely to the normalized measure $\rho|_V / \rho(V)$. From [**9**, Corollary 30.9] we conclude that the sequence $(\rho^{(n)}|_V / \rho^{(n)}(V))_{n\in\mathbb{N}}$ converges *weakly* to the normalized measure $\rho|_V / \rho(V)$.

Since V is separable (see e.g. [**4**, Corollary 3.5]), we may apply [**13**, Theorem 2.8] to obtain *weak* convergence of the corresponding product measures,

$$\begin{aligned}\big(\rho^{(n)}|_V/\rho^{(n)}(V)\big)\otimes\big(\rho^{(n)}|_V/\rho^{(n)}(V)\big) &\rightharpoonup \big(\rho|_V/\rho(V)\big)\otimes\big(\rho|_V/\rho(V)\big)\,,\\ \big(\rho^{(n)}|_V/\rho^{(n)}(V)\big)\otimes\big(\rho^{(n)}|_{V'}/\rho^{(n)}(V')\big) &\rightharpoonup \big(\rho|_V/\rho(V)\big)\otimes\big(\rho|_{V'}/\rho(V')\big)\,.\end{aligned} \tag{3.4.19}$$

We now proceed as follows. First of all, in accordance with (3.2.4) we have

$$\begin{aligned}\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho) = \int_{\mathcal{F}} d(\tilde\rho-\rho)(x)\int_{\mathcal{F}} d\rho(y)\,\mathcal{L}(x,y) + \int_{\mathcal{F}} d\rho(x)\int_{\mathcal{F}} d(\tilde\rho-\rho)(y)\,\mathcal{L}(x,y)\\ + \int_{\mathcal{F}} d(\tilde\rho-\rho)(x)\int_{\mathcal{F}} d(\tilde\rho-\rho)(y)\,\mathcal{L}(x,y)\,.\end{aligned}$$

Making use of the symmetry of the Lagrangian and applying Fubini's theorem, we can write this expression as

$$\begin{aligned}\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho) = 2\int_{\mathcal{F}} d(\tilde\rho-\rho)(x)\int_{\mathcal{F}} d\rho(y)\,\mathcal{L}(x,y)\\ + \int_{\mathcal{F}} d(\tilde\rho-\rho)(x)\int_{\mathcal{F}} d(\tilde\rho-\rho)(y)\,\mathcal{L}(x,y)\,,\end{aligned}$$

and the fact that $\mathcal{L}$ is of compact range yields

$$\begin{aligned}\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho) = 2\int_{K} d(\tilde\rho-\rho)(x)\int_{K'} d\rho(y)\,\mathcal{L}(x,y)\\ + \int_{K} d(\tilde\rho-\rho)(x)\int_{K} d(\tilde\rho-\rho)(y)\,\mathcal{L}(x,y)\,.\end{aligned}$$

Making use of the fact that $\rho(V\setminus K)<\tilde\varepsilon$, $\tilde\rho(V\setminus K)<\tilde\varepsilon$ and $\rho(V'\setminus K')<\tilde\varepsilon$, $\tilde\rho(V'\setminus K')<\tilde\varepsilon$, we can arrange that, up to an arbitrarily small error term $\varepsilon>0$,

$$\begin{aligned}\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho) \geq 2\int_V d\tilde\rho(x)\int_{V'} d\rho(y)\,\mathcal{L}(x,y) - 2\int_V d\rho(x)\int_{V'} d\rho(y)\,\mathcal{L}(x,y)\\ + \int_V d\tilde\rho(x)\int_V d\tilde\rho(y)\,\mathcal{L}(x,y) - \int_V d\tilde\rho(x)\int_V d\rho(y)\,\mathcal{L}(x,y)\\ - \int_V d\rho(x)\int_V d\tilde\rho(y)\,\mathcal{L}(x,y) + \int_V d\rho(x)\int_V d\rho(y)\,\mathcal{L}(x,y) - \varepsilon\,.\end{aligned}$$

Since $\mathcal{L}\in C_b(\mathcal{F}\times\mathcal{F})$, by applying weak convergence (3.4.19) we obtain

$$\begin{aligned}&\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho)\\ &\geq \lim_{n\to\infty}\Bigg[2\left(\int_V d\tilde\rho(x)\int_{V'} d\rho^{(n)}|_{V'}(y) - \int_V d\rho^{(n)}|_V(x)\int_{V'} d\rho^{(n)}|_{V'}(y)\right)\,\mathcal{L}(x,y)\\ &\quad + \int_V d\tilde\rho(x)\int_V d\tilde\rho(y)\,\mathcal{L}(x,y) - \int_V d\tilde\rho(x)\int_V d\rho^{(n)}|_V(y)\,\mathcal{L}(x,y)\\ &\quad - \int_V d\rho^{(n)}|_V(x)\int_V d\tilde\rho(y)\,\mathcal{L}(x,y) + \int_V d\rho^{(n)}|_V(x)\int_V d\rho^{(n)}|_V(y)\,\mathcal{L}(x,y)\Bigg] - \varepsilon\,,\end{aligned}$$

or equivalently,

$$\begin{aligned}\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho) \geq \lim_{n\to\infty}\Bigg[2\int_V d\big(\tilde\rho-\rho^{(n)}\big)(x)\int_{V'} d\rho^{(n)}(y)\,\mathcal{L}(x,y)\\ + \int_V d\big(\tilde\rho-\rho^{(n)}\big)(x)\int_V d\big(\tilde\rho-\rho^{(n)}\big)(y)\,\mathcal{L}(x,y)\Bigg] - \varepsilon\,.\end{aligned}$$

Next, for any $n \in \mathbb{N}$ we introduce the measures $\tilde{\rho}_n : \mathcal{B}(\mathcal{F}) \to [0, +\infty]$ by

$$\tilde{\rho}_n := \begin{cases} c_n\, \tilde{\rho} & \text{on } V \\ \rho^{(n)} & \text{on } \mathcal{F} \setminus V \end{cases} \qquad \text{with} \qquad c_n := \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)} \quad \text{for all } n \in \mathbb{N}\,.$$

Considering the compact exhaustion $(K_n)_{n\in\mathbb{N}}$, we thus have $\rho^{(n)}(K_n) = \tilde{\rho}_n(K_n)$ for each $n \in \mathbb{N}$. Furthermore,

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)} \overset{(3.4.18)}{=} \frac{\rho(V)}{\tilde{\rho}(V)} = \frac{\rho(V\setminus K) + \rho(K)}{\tilde{\rho}(V\setminus K) + \tilde{\rho}(K)} = 1 \tag{3.4.20}$$

according to $\rho|_{\mathcal{F}\setminus K} = \tilde{\rho}|_{\mathcal{F}\setminus K}$ and $\tilde{\rho}(K) = \rho(K)$. In view of (3.4.20) we may write

$$\begin{aligned} \mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \lim_{n\to\infty} \Big[& 2 \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\rho^{(n)}(y)\, \mathcal{L}(x,y) \\ & + \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(y)\, \mathcal{L}(x,y) \Big] - \varepsilon\,. \end{aligned}$$

Since $\tilde{\rho}_n$ and $\rho^{(n)}$ coincide on $K_n \setminus V$ for all sufficiently large $n \in \mathbb{N}$, and $\mathcal{L}(x,y) = 0$ for all $x \in V$ and $y \notin V'$, the difference $\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)$ can finally be written as

$$\begin{aligned} \mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) = \lim_{n\to\infty} \Big[& 2 \int_{K_n} d\big(\tilde{\rho}_n - \rho^{(n)}\big)(x) \int_{K_n} d\rho^{(n)}(y)\, \mathcal{L}(x,y) \\ & + \int_{K_n} d\big(\tilde{\rho}_n - \rho^{(n)}\big)(x) \int_{K_n} d\big(\tilde{\rho}_n - \rho^{(n)}\big)(y)\, \mathcal{L}(x,y) \Big] - \varepsilon\,. \end{aligned}$$

Since $\rho^{(n)}$ is a minimizer on K_n for every $n \in \mathbb{N}$, we are given

$$\big(\mathcal{S}_{K_n}(\tilde{\rho}_n) - \mathcal{S}_{K_n}(\rho^{(n)})\big) \geq 0 \qquad \text{for all } n \in \mathbb{N}\,. \tag{3.4.21}$$

Taking the limit $n \to \infty$ on the left hand side of (3.4.21), one obtains exactly the above expression in square brackets, which therefore is bigger than or equal to zero. Since $\varepsilon > 0$ is arbitrary, we thus arrive at

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

Hence ρ is a minimizer under variations of compact support. □

3.4.4. Existence of Minimizers under Variations of Finite Volume. In order to prove the existence of minimizers in the sense of Definition 3.3.2, we additionally assume that property (iv) in §3.2.2 is satisfied, i.e.

$$\sup_{x\in\mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty\,.$$

Under this additional assumption, the difference (3.2.4) is well-defined. Moreover, we obtain the following existence result.

THEOREM 3.4.10. *Let $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ be continuous, bounded, and of compact range, and assume that condition* (iv) *in §3.2.2 is satisfied. Then ρ is a minimizer under variations of finite volume (see Definition 3.3.2).*

PROOF. Let $\tilde{\rho} \in \mathfrak{B}_{\mathcal{F}}$ be a positive Borel measure on $\mathcal{F}$ satisfying (3.2.3), i.e.

$$|\tilde{\rho} - \rho|(\mathcal{F}) < \infty \qquad \text{and} \qquad (\tilde{\rho} - \rho)(\mathcal{F}) = 0\,.$$

Introducing $B := \operatorname{supp}(\tilde{\rho} - \rho)$, we thus are given $0 < \rho(B) = \tilde{\rho}(B) < \infty$. Without loss of generality, we may assume that $\rho(B) = \tilde{\rho}(B) > 1$ (otherwise we multiply the measures ρ

and $\tilde{\rho}$ by a suitable constant). By assuming that condition (iv) in §3.2.2 holds we know that the difference (3.2.4) is well-defined, thus giving rise to

$$\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) = 2 \int_B d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_B d(\tilde{\rho} - \rho)(x) \int_B d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\,.$$

Let $\tilde{\varepsilon} > 0$ be arbitrary. In analogy to the proof of Theorem 3.4.10, by regularity of ρ and $\tilde{\rho}$ we approximate B by open sets $U \supset B$ from outside such that

$$\rho(U \setminus B) < \tilde{\varepsilon}/4\,, \qquad \tilde{\rho}(U \setminus B) < \tilde{\varepsilon}/4\,.$$

By adding and subtracting the terms

$$2 \int_{U\setminus B} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_{U\setminus B} d(\tilde{\rho} - \rho)(x) \int_B d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)$$

as well as

$$\int_U d(\tilde{\rho} - \rho)(x) \int_{U\setminus B} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\,,$$

one can show that

$$\begin{aligned}\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) &= 2 \int_U d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ &+ \int_U d(\tilde{\rho} - \rho)(x) \int_U d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) - \bigg\{ \int_U d(\tilde{\rho} - \rho)(x) \int_{U\setminus B} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \\ &+ 2 \int_{U\setminus B} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_{U\setminus B} d(\tilde{\rho} - \rho)(x) \int_B d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \bigg\}\,.\end{aligned}$$

Choosing $U \supset B$ suitably, property (iv) implies (along with (3.4.13)) that the expression

$$\left| \int_{U\setminus B} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \right| \leq \underbrace{\Big(\sup_{x\in\mathcal{F}} \ell(x) + 1 \Big)}_{<\infty} \underbrace{\big(|\tilde{\rho}(U \setminus B)| + |\rho(U \setminus B)| \big)}_{<\tilde{\varepsilon}/2}$$

can be arranged to be arbitrarily small. Similarly, one can show that the other summands in the expression in curly brackets are arbitrarily small for a suitable choice of $U \supset B$. We thus can arrange that

$$\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq 2 \int_U d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_U d(\tilde{\rho} - \rho)(x) \int_U d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) - \varepsilon$$

for any given $\varepsilon > 0$. Next, by regularity of ρ and $\tilde{\rho}$ we may approximate U by compact sets $V \subset U$ from inside such that

$$\rho(U \setminus V) < \tilde{\varepsilon}/4\,, \qquad \tilde{\rho}(U \setminus V) < \tilde{\varepsilon}/4\,.$$

Proceeding in analogy to the proof of Theorem 3.4.9, we may assume that $V \subset U$ is an open, relatively compact set such that $\rho^{(n)}(V) \to \rho(V)$. Moreover, since $\mathcal{L}$ is supposed to be of compact range, there is some relatively compact, open subset $V' \subset \mathcal{F}$ such

that $\mathcal{L}(x,y) = 0$ for all $x \in V$ and $y \notin V'$ (and vice versa, for details see the proof of Theorem 3.4.9). Applying similar arguments as before to $V \subset U$, we arrive at

$$\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \Big[2 \int_V d(\tilde{\rho}-\rho)(x) \int_{V'} d\rho(y)\, \mathcal{L}(x,y) + \int_V d(\tilde{\rho}-\rho)(x) \int_V d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y) \Big] - 2\varepsilon\,.$$

Proceeding in analogy to the proof of Theorem 3.4.9, one can show that the term in square brackets is greater than or equal to zero, up to an arbitrarily small error term. Indeed, introducing the measures

$$\tilde{\rho}_n := \begin{cases} c_n\, \tilde{\rho} & \text{on } V \\ \rho^{(n)} & \text{on } \mathcal{F} \setminus V \end{cases} \qquad \text{with} \qquad c_n := \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)} \quad \text{for all } n \in \mathbb{N}\,,$$

we are given

$$\tilde{\rho}_n(V) = c_n\, \tilde{\rho}(V) = \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)}\, \tilde{\rho}(V) = \rho^{(n)}(V)\,.$$

Note that $\rho(V), \tilde{\rho}(V) \in (\rho(B) - \tilde{\varepsilon}/2, \rho(B) + \tilde{\varepsilon}/2) \subset (1, \infty)$ by choosing $\tilde{\varepsilon} > 0$ suitably, implying that $1 < \rho(V) < \tilde{\rho}(V) + \tilde{\varepsilon}$. In view of $\rho^{(n)}(V) \to \rho(V)$, for sufficiently large $n \in \mathbb{N}$ we thus obtain

$$c_n = \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)} < \frac{\tilde{\rho}(V) + \tilde{\varepsilon}}{\tilde{\rho}(V)} = 1 + \frac{\tilde{\varepsilon}}{\tilde{\rho}(V)} \leq 1 + \tilde{\varepsilon}$$

(and similarly, $c_n > 1 - \tilde{\varepsilon}$ for sufficiently large $n \in \mathbb{N}$). As a consequence,

$$c_n^2 < (1+\tilde{\varepsilon})^2 < 1 + 3\tilde{\varepsilon}\,,$$

and henceforth the term in square brackets can be estimated by

$$\begin{aligned} &\bigg(2 \int_V d\big(\tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\rho^{(n)}(y) + \int_V d\big(\tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\big(\tilde{\rho} - \rho^{(n)}\big)(y) \bigg)\, \mathcal{L}(x,y) \\ &\quad \geq \Big\{ 2 \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\rho^{(n)}(y)\, \mathcal{L}(x,y) \\ &\qquad + \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(y)\, \mathcal{L}(x,y) \Big\} \\ &\qquad - 4\tilde{\varepsilon} \int_V d\tilde{\rho}(x) \int_{V'} d\rho^{(n)}(y)\, \mathcal{L}(x,y) - 3\tilde{\varepsilon} \int_V d\tilde{\rho}(x) \int_{V'} d\tilde{\rho}(y)\, \mathcal{L}(x,y)\,. \end{aligned}$$

Choosing $\tilde{\varepsilon} > 0$ suitably, one can arrange that

$$\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \lim_{n\to\infty} \Big\{ 2 \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\rho^{(n)}(y)\, \mathcal{L}(x,y) + \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(y)\, \mathcal{L}(x,y) \Big\} - 9\varepsilon\,.$$

Note that $V' \subset K_n$ for sufficiently large $n \in \mathbb{N}$. Making use of the fact that $\rho^{(n)}$ is a minimizer on K_n and arguing as in the proof of Theorem 3.4.9, we conclude that the term in curly brackets in the last inequality is greater than or equal to zero.

Finally, since $\varepsilon > 0$ was chosen arbitrarily, we arrive at

$$\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq 0\,,$$

which proves the claim. □

3.5. Minimizers for Lagrangians Decaying in Entropy

3.5.1. Preliminaries. The goal of this section is to deal with the question if it is possible to weaken the assumption that the Lagrangian $\mathcal{L}$ is of compact range. To this end, we specialize the above setting as follows. As before, we let $\mathcal{F}$ be a second-countable, locally compact Hausdorff space. Then $\mathcal{F}$ is completely metrizable, and hence can be endowed with a Heine-Borel metric as mentioned in §3.3.1 such that $\mathcal{F}$ is proper, i.e. closed, bounded subsets in $\mathcal{F}$ are compact. As every relatively compact set is precompact, any bounded subset of $\mathcal{F}$ can be covered by a finite number of sets of diameter less than $\delta > 0$ (cf. [**30**, §3.16, §3.17]). Thus for any $r > 0$ and $x \in \mathcal{F}$, the closed ball $\overline{B_r(x)}$ is compact, and hence can be covered by finitely many balls of radius $\delta > 0$. We denote the smallest such number by $E_x(r, \delta)$.[3] In particular, for all $r' < r$ the annuli $\overline{B_r(x) \setminus B_{r'}(x)}$ can be covered by at most $E_x(r, \delta)$ balls of radius δ. If ρ is a uniform measure on $\mathcal{F}$, the number $E_x(r, \delta)$ can be determined more specifically (see [**136**, Example 3.13]).

In the following, we additionally assume that the Lagrangian decays in entropy, which is defined as follows.

Definition 3.5.1. *Assume that $\mathcal{F}$ is endowed with an unbounded Heine-Borel metric d. The Lagrangian $\mathcal{L}\colon \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is said to* **decay in entropy** *if the following conditions are satisfied:*

(a) $c := \inf_{x \in \mathcal{F}} \mathcal{L}(x, x) > 0$.

(b) *There is a compact set $K \subset \mathcal{F}$ such that*

$$\delta := \inf_{x \in \mathcal{F} \setminus K} \sup \Big\{ s \in \mathbb{R} \,:\, \mathcal{L}(x, y) \geq \frac{c}{2} \quad \text{for all } y \in B_s(x) \Big\} > 0 \,.$$

(c) *The Lagrangian has the following decay property: There is a monotonically decreasing, integrable function $f \in L^1(\mathbb{R}^+, \mathbb{R}_0^+)$ such that*

$$\mathcal{L}(x, y) \leq \frac{f\big(d(x, y)\big)}{C_x\big(d(x, y), \delta\big)} \qquad \text{for all } x, y \in \mathcal{F} \text{ with } x \neq y \,,$$

where

$$C_x(r, \delta) := C\, E_x(r + 2, \delta) \qquad \text{for all } r > 0 \,,$$

and the constant C is given by

$$C := 1 + \frac{2}{c} < \infty \,.$$

In Definition 3.5.1 (b) we may assume without loss of generality that $\delta = 1$ (otherwise we rescale the metric suitably). Then

$$\mathcal{L}(x, y) \geq \frac{\mathcal{L}(x, x)}{2} \qquad \text{for all } y \in B_1(x) \,.$$

Now let $(\rho^{(n)})_{n \in \mathbb{N}}$ be the sequence of measures given by (3.4.4), and let ρ be its vague limit constructed in (3.4.7). Then by (3.4.3), for every $x \in \mathcal{F}$ we have $\rho(B_1(x)) \leq C$ as well as $\rho^{(n)}(B_1(x)) \leq C$ for sufficiently large $n \in \mathbb{N}$.

Condition (b) determines the behavior of the Lagrangian locally (more precisely, it gives a uniform bound for the size of balls in which the Lagrangian is bounded from below).

[3] In coarse geometry, this number is called *entropy* of a set (cf. [**136**, Definition 3.1]). In the literature, however, also the logarithm of this number is referred to as *δ-entropy* (see [**30**, §3.16, Problem 4]).

In condition (c), on the other hand, the function f characterizes the decay properties of the Lagrangian at infinity. In particular, condition (c) implies that for any $\varepsilon > 0$ there is an integer $N_0 = N_0(\varepsilon) > 1$ such that

$$\sum_{k=n}^{\infty} f(k) \leq \int_{n-1}^{\infty} f(x)\, dx < \varepsilon/3 \qquad \text{for all } n \geq N_0\,. \tag{3.5.1}$$

Considering arbitrary $\varepsilon > 0$ and $x \in \mathcal{F}$, the Heine-Borel property of $\mathcal{F}$ ensures that the closed ball

$$K_{x,\varepsilon} := \overline{B_{N_0}(x)} \tag{3.5.2}$$

is compact. Since $\mathcal{L}$ decays in entropy, we thus obtain

$$\begin{aligned}\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\rho(y) &= \sum_{k=N_0}^{\infty} \int_{B_{k+1}(x)\setminus\overline{B_k(x)}} \mathcal{L}(x,y)\, d\rho(y) \\ &\leq \sum_{k=N_0}^{\infty} \sup_{y\in B_{k+1}(x)\setminus\overline{B_k(x)}} \mathcal{L}(x,y)\, \underbrace{\rho\Big(B_{k+1}(x)\setminus\overline{B_k(x)}\Big)}_{\leq C_x(k,1)} \leq \sum_{k=N_0}^{\infty} \frac{f(k)}{C_x(k,1)}\, C_x(k,1) < \varepsilon/3\,,\end{aligned}$$

where in the last step we made use of (3.5.1). Analogously,

$$\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(\tilde{x},y)\, d\rho(y) < \varepsilon/3 \tag{3.5.3}$$

for all $\tilde{x} \in B_\delta(x)$ and sufficiently small $\delta > 0$. Under the assumption that the Lagrangian is continuous, the same is true for the measures $\rho^{(n)}$: Given $x \in \mathcal{F}$, we introduce the compact sets $A_k(x) \subset \mathcal{F}$ by

$$A_k(x) := \overline{B_{k+1}(x)\setminus B_k(x)} \qquad \text{for all } k \geq N_0$$

and choose open sets $U_k(x) \supset A_k(x)$ with $U_k(x) \subset B_{k+2}(x)\setminus B_{k-1}(x)$ such that

$$\rho\Big(U_k \setminus \big(B_{k+1}(x)\setminus\overline{B_k(x)}\big)\Big) < 2^{-k-1}\,\varepsilon/3 \qquad \text{for all } k \geq N_0\,.$$

Then in view of [**4**, Lemma 2.92], for every $k \geq N_0$ there exists $\eta_k \in C_c(U_k(x);[0,1])$ such that $\eta_k|_{A_k(x)} \equiv 1$, implying that $\mathcal{L}(x,.)\,\eta_k \in C_c(U_k(x))$ for all $k \geq N_0$. Hence for arbitrary $m > 0$, by vague convergence (3.4.9) there exists some integer $N_k = N_k(m)$ such that

$$\left| \int_{U_k(x)} \mathcal{L}(x,y)\, \eta_k(y)\, d\rho^{(n)}(y) - \int_{U_k(x)} \mathcal{L}(x,y)\, \eta_k(y)\, d\rho(y) \right| \leq 2^{-k}/m \tag{3.5.4}$$

for all $n \geq N_k$. For $k = N_0$, we choose a subsequence of $(\rho^{(n)})_{n\in\mathbb{N}}$ which we denote by $(\rho^{(1,n_\ell)})_{\ell\in\mathbb{N}}$ such that (3.5.4) holds for every $m \in \mathbb{N}$ and all $\ell \geq N(m) := N_1(m)$. Next, for $k = N_0 + 1$ we choose a subsequence of $(\rho^{(1,n_\ell)})_{\ell\in\mathbb{N}}$, denoted by $(\rho^{(2,n_\ell)})_{\ell\in\mathbb{N}}$, such that (3.5.4) is satisfied for every $m \in \mathbb{N}$ and all $\ell \geq N(m)$. Proceeding iteratively, we obtain a corresponding diagonal sequence $(\rho^{(\ell,n_\ell)})_{\ell\in\mathbb{N}}$, which for convenience we shall again denote by $(\rho^{(n)})_{n\in\mathbb{N}}$. This sequence has the property that, for each $k \geq N_0$ and for every $m \in \mathbb{N}$, there exists $N = N(m)$ such that (3.5.4) is satisfied for all $n \geq N$. As a

consequence, for arbitrary $\varepsilon > 0$ there exists $m \in \mathbb{N}$ such that $1/m < \varepsilon/6$, giving rise to

$$\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) = \sum_{k=N_0}^{\infty} \int_{B_{k+1}(x)\setminus\overline{B_k(x)}} \mathcal{L}(x,y)\, d\rho^{(n)}(y)$$
$$\leq \sum_{k=N_0}^{\infty} \int_{U_k(x)} \mathcal{L}(x,y)\, \eta_k(y)\, d\rho^{(n)}(y) \leq \sum_{k=N_0}^{\infty} \left(\int_{U_k} \mathcal{L}(x,y)\, \eta_k(y)\, d\rho(y) + 2^{-k}/m \right) < \varepsilon$$

for all $n \geq N(m)$, where we applied the geometric series.

Next, we employ the fact that $\mathcal{F}$ is separable. Consequently, there exists a countable dense subset $\{x_i : i \in \mathbb{N}\}$. Then the above arguments show that

$$\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) < \varepsilon \qquad \text{for all } n \geq N(m) \tag{3.5.5}$$

holds for $x = x_1$. By choosing a suitable subsequence $(\rho^{(2,n_\ell)})_{\ell\in\mathbb{N}}$, we can arrange that (3.5.5) holds for $x = x_2$. Proceeding iteratively, we end up with a corresponding diagonal sequence $(\rho^{(\ell,n_\ell)})_{\ell\in\mathbb{N}}$, which for simplicity we again denote by $(\rho^{(n)})_{n\in\mathbb{N}}$, such that (3.5.5) holds for $x = x_i$ for every $i \in \mathbb{N}$. Combining the above inequalities, we obtain

$$\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(\tilde{x},y)\, d\rho^{(n)}(y) \leq \sum_{k=N_0}^{\infty} \left(\int_{U_k} \mathcal{L}(\tilde{x},y)\, \eta_k(y)\, d\rho(y) + 2^{-k}/m \right).$$

Using (3.5.3), we conclude that for all $\tilde{x}$ in a small neighborhood of x and for sufficiently large $n \in \mathbb{N}$,

$$\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(\tilde{x},y)\, d\rho^{(n)}(y) < \varepsilon \qquad \text{and} \qquad \int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(\tilde{x},y)\, d\rho(y) < \varepsilon\,. \tag{3.5.6}$$

In particular, these inequalities imply that the measure ρ constructed in (3.4.7) is non-zero (for details see Lemma A.1.2).

3.5.2. Preparatory Results. Based on (3.5.6), the goal of this subsection is to derive results similar to Lemma 3.4.4, Proposition 3.4.5 and Proposition 3.4.6. We first prove continuity of the function ℓ (as defined in (3.4.13)).

PROPOSITION 3.5.2. *Assume that the Lagrangian $\mathcal{L} : \mathcal{F}\times\mathcal{F} \to \mathbb{R}_0^+$ is continuous and decays in entropy. Then the function $\ell : \mathcal{F} \to \mathbb{R}$ is continuous.*

PROOF. Let $x \in \mathcal{F}$ and $\varepsilon > 0$ be arbitrary, and let $(x_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in $\mathcal{F}$ converging to x. Introducing the associated compact set $K_{x,\varepsilon}$ (as defined in (3.5.2)), by continuity of $\mathcal{L}$ and $\rho(K_{x,\varepsilon}) < \infty$ we obtain

$$\big|\ell(x) - \ell(x_n)\big| \leq \left| \int_{K_{x,\varepsilon}} \big(\mathcal{L}(x,y) - \mathcal{L}(x_n,y)\big)\, d\rho(y) \right| + \left| \int_{\mathcal{F}\setminus K_{x,\varepsilon}} \big(\mathcal{L}(x,y) - \mathcal{L}(x_n,y)\big)\, d\rho(y) \right|$$
$$\leq \big|\mathcal{L}(x,y) - \mathcal{L}(x_n,y)\big|\, \rho(K_{x,\varepsilon}) + \left| \int_{\mathcal{F}\setminus K_{x,\varepsilon}} \big(\mathcal{L}(x,y) - \mathcal{L}(x_n,y)\big)\, d\rho(y) \right| \overset{(3.5.6)}{<} 3\varepsilon$$

for sufficiently large $n \in \mathbb{N}$. This proves continuity of ℓ. □

PROPOSITION 3.5.3. *Let $(\ell^{(n)})_{n\in\mathbb{N}}$ and ℓ be the functions defined in* (3.4.10) *and* (3.4.13). *Then $(\ell^{(n)})_{n\in\mathbb{N}}$ converges pointwise to ℓ, i.e.*

$$\lim_{n\to\infty} \ell^{(n)}(x) = \ell(x) \qquad \textit{for all } x \in \mathcal{F}\,.$$

PROOF. Let $x \in \mathcal{F}$ and $\varepsilon > 0$ be arbitrary, and let $K_{x,\varepsilon} \subset \mathcal{F}$ compact according to (3.5.2). Given $U \supset K_{x,\varepsilon}$ open, there exists $\eta \in C_c(U; [0,1])$ such that $\eta|_{K_{x,\varepsilon}} \equiv 1$ (see e.g. [**4**, Lemma 2.92]). Thus in view of (3.5.6), for sufficiently large $n \in \mathbb{N}$ we obtain

$$\left| \int_{U \setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, \eta(y)\, d\rho^{(n)}(y) \right| \leq \left| \int_{\mathcal{F} \setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\; d\rho^{(n)}(y) \right| \overset{(3.5.6)}{<} \varepsilon\,.$$

Making use of $\mathcal{L}(x,.)\, \eta \in C_c(\mathcal{F})$ and vague convergence (3.4.9), this yields

$$\begin{aligned}
\left| \ell(x) - \ell^{(n)}(x) \right| &= \left| \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) - \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) \right| \\
&\leq \left| \int_{K_{x,\varepsilon}} \mathcal{L}(x,y)\, \eta(y)\, d\big(\rho - \rho^{(n)}\big)(y) \right| + \underbrace{\left| \int_{\mathcal{F} \setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\big(\rho - \rho^{(n)}\big)(y) \right|}_{< 2\varepsilon} \\
&< \left| \int_{U} \mathcal{L}(x,y)\, \eta(y)\, d\big(\rho - \rho^{(n)}\big)(y) \right| + \underbrace{\left| \int_{U \setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, \eta(y)\, d\big(\rho - \rho^{(n)}\big)(y) \right|}_{< 2\varepsilon} + 2\varepsilon \\
&< \varepsilon + 2\varepsilon + 2\varepsilon < 5\varepsilon
\end{aligned}$$

for sufficiently large $n \in \mathbb{N}$, which gives the claim. □

PROPOSITION 3.5.4. *Assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ is continuous and decays in entropy, and let $K \subset \mathcal{F}$ be compact. Then for every $x \in K$ and every sequence $\big(x^{(n)}\big)_{n \in \mathbb{N}}$ in K with $x^{(n)} \to x$ we have*

$$\lim_{n \to \infty} \left| \ell^{(n)}|_K\big(x^{(n)}\big) - \ell^{(n)}|_K(x) \right| = 0\,.$$

PROOF. Let $K \subset \mathcal{F}$ be a compact subset. For any $x \in K$ and $\varepsilon > 0$, there is a compact subset $K_{x,\varepsilon} \subset \mathcal{F}$ (defined by (3.5.2)) such that (3.5.6) is satisfied. Let $C(x,\varepsilon) > 0$ be the positive constant according to Lemma 3.4.1 such that $\rho^{(n)}(K_{x,\varepsilon}) \leq C(x,\varepsilon)$ for all $n \in \mathbb{N}$. Since $\mathcal{L}$ is continuous and $K \times K_{x,\varepsilon}$ is compact, the mapping

$$\mathcal{L}|_{K \times K_{x,\varepsilon}} : K \times K_{x,\varepsilon} \to \mathbb{R}$$

is uniformly continuous. Hence we may choose $\delta > 0$ such that

$$\big| \mathcal{L}|_{K \times K_{x,\varepsilon}}(x,\cdot) - \mathcal{L}|_{K \times K_{x,\varepsilon}}(z,\cdot) \big| < \frac{\varepsilon}{2C(x,\varepsilon)} \qquad \text{for all } z \in B_\delta(x) \cap K\,.$$

In view of (3.5.6), for sufficiently large $n \in \mathbb{N}$ we thus obtain

$$\begin{aligned}
\sup_{z \in B_\delta(x) \cap K} \left| \ell^{(n)}|_K(x) - \ell^{(n)}|_K(z) \right| &= \sup_{z \in B_\delta(x) \cap K} \left| \int_{\mathcal{F}} \big(\mathcal{L}|_{K \times \mathcal{F}}(x,y) - \mathcal{L}|_{K \times \mathcal{F}}(z,y) \big)\, d\rho^{(n)}(y) \right| \\
&\leq \sup_{z \in B_\delta(x) \cap K} \left| \int_{\mathcal{F} \setminus K_{x,\varepsilon}} \big(\mathcal{L}|_{K \times \mathcal{F}}(x,y) - \mathcal{L}|_{K \times \mathcal{F}}(z,y) \big)\, d\rho^{(n)}(y) \right| \\
&\quad + \sup_{z \in B_\delta(x) \cap K} \left| \int_{K_{x,\varepsilon}} \big(\mathcal{L}|_{K \times \mathcal{F}}(x,y) - \mathcal{L}|_{K \times \mathcal{F}}(z,y) \big)\, d\rho^{(n)}(y) \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.
\end{aligned}$$

Considering a sequence $\big(x^{(n)}\big)_{n \in \mathbb{N}}$ in K with $x^{(n)} \to x$, we have

$$\lim_{n \to \infty} \left| \ell^{(n)}|_K\big(x^{(n)}\big) - \ell^{(n)}|_K(x) \right| = 0\,,$$

which completes the proof. □

3.5.3. The Euler-Lagrange Equations. Now we are able to prove the EL equations in the case that $\mathcal{L}$ decays in entropy (see Definition 3.5.1).

THEOREM 3.5.5. *Assume that $\mathcal{L}$ is continuous and decays in entropy. Then the measure ρ constructed in* (3.4.7) *satisfies the Euler-Lagrange equations*

$$\ell|_{\operatorname{supp}\rho} \equiv \inf_{x\in\mathcal{F}} \ell(x) = 0\,,$$

where $\ell \in C(\mathcal{F})$ is defined by (3.4.13).

PROOF. Proceed in analogy to the proof of Theorem 3.4.2, and make use of Proposition 3.5.2, Proposition 3.5.3, and Proposition 3.5.4. □

We now generalize Lemma 3.4.7.

LEMMA 3.5.6. *Assume that $\mathcal{L} : \mathcal{F}\times\mathcal{F} \to \mathbb{R}$ is continuous and decays in entropy. Under the additional assumptions* (a)–(c) *in Lemma 3.4.7 (with K_x replaced by $K_{x,\varepsilon}$), the measure ρ constructed in* (3.4.7) *satisfies property* (iv) *in §3.2.2.*

PROOF. Since continuity of $\mathcal{L}$ implies that

$$\int_{\mathcal{F}} \mathcal{L}(x,y)\,d\rho(y) = \int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\,d\rho(y) + \int_{K_{x,\varepsilon}} \mathcal{L}(x,y)\,d\rho(y) < \infty\,,$$

the function $\mathcal{L}(x,\cdot) : \mathcal{F} \to \mathbb{R}$ is ρ-integrable for every $x \in \mathcal{F}$. It remains to show that

$$\sup_{x\in\mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\,d\rho(y) < \infty\,.$$

For all $x \in M := \operatorname{supp}\rho$ this result follows from Theorem 3.5.5. Whenever $x \in \mathcal{F}\setminus M$, similar as in the proof of Lemma 3.4.7 we obtain

$$\begin{aligned}\sup_{x\in\mathcal{F}\setminus M} \int_{\mathcal{F}} \mathcal{L}(x,y)\,d\rho(y) &\le \sup_{x\in\mathcal{F}\setminus M} \left(\int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\,d\rho(y) + \int_{K_{x,\varepsilon}} \mathcal{L}(x,y)\,d\rho(y)\right)\\ &\le \varepsilon + \sup_{x\in\mathcal{F}\setminus M} \sup_{y\in K_{x,\varepsilon}} \mathcal{L}(x,y)\,\rho(K_{x,\varepsilon}) < \infty\end{aligned}$$

for some $\varepsilon > 0$ and the corresponding compact subset $K_{x,\varepsilon} \subset \mathcal{F}$ (see (3.5.2)). □

Moreover, under the additional assumptions (b) and (c) in Lemma 3.4.7 the following statement is true:

COROLLARY 3.5.7. *Assume that $\mathcal{L}$ is continuous and decays in entropy. Under the additional assumptions* (b) and (c) *in Lemma 3.4.7 (again with K_x replaced by $K_{x,\varepsilon}$), for all $\varepsilon \in (0,1)$ there is $\gamma > 0$ such that $\rho(K_{x,\varepsilon}) \ge \gamma$ for all $x \in \mathcal{F}$ (where $K_{x,\varepsilon}$ is given by* (3.5.2)*).*

PROOF. Consider an arbitrary $x \in \mathcal{F}$. In view of Theorem 3.5.5 we have

$$1 \le \int_{\mathcal{F}\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\,d\rho(y) + \int_{K_{x,\varepsilon}} \mathcal{L}(x,y)\,d\rho(y) < \varepsilon + \sup_{y\in K_{x,\varepsilon}} \mathcal{L}(x,y)\,\rho(K_{x,\varepsilon})$$

for some $\varepsilon > 0$ and the corresponding compact set $K_{x,\varepsilon} \subset \mathcal{F}$. Choosing $\varepsilon \in (0,1)$, we obtain

$$\rho(K_{x,\varepsilon}) \ge \frac{1-\varepsilon}{\mathcal{C}} =: \gamma\,,$$

which completes the proof. □

3.5.4. Existence of Minimizers under Variations of Compact Support. In the last two subsections we finally return to the question if the measure ρ is a minimizer of the causal variational principle. In preparation, we deal with the case of minimizers under variations of compact support (see Definition 3.4.8).

THEOREM 3.5.8. *Assume that $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is continuous and decays in entropy (see Definition 3.5.1). Then ρ is a minimizer under variations of compact support.*

PROOF. Let $\tilde{\rho} \in \mathfrak{B}_{\mathcal{F}}$ be a variation of compact support. Then $K := \text{supp}(\tilde{\rho} - \rho)$ is compact, and $\rho(K) = \tilde{\rho}(K) < \infty$. Since the Lagrangian is continuous and decays in entropy, the function $\ell(x)$ (see (3.4.13)) is locally bounded. As a consequence, the difference of the actions (3.2.4) is well-defined. Thus it remains to show that

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0$$

for all variations $\tilde{\rho}$ of compact support. Given $\tilde{\varepsilon} > 0$ and $x \in K$, we know that

$$\int_{\mathcal{F} \setminus K_{x,\tilde{\varepsilon}/2}} \mathcal{L}(x,y)\, d\rho(y) < \frac{\tilde{\varepsilon}}{2}$$

(where $K_{x,\tilde{\varepsilon}/2} \subset \mathcal{F}$ is given according to (3.5.2)). By continuity of the Lagrangian, there is an open neighborhood U_x of x such that

$$\int_{\mathcal{F} \setminus K_{x,\tilde{\varepsilon}/2}} \mathcal{L}(z,y)\, d\rho(y) < \tilde{\varepsilon} \qquad \text{for all } z \in U_x\,.$$

Covering the compact set K by a finite number of such neighborhoods $U_{x_1}, \ldots, U_{x_L}$ and introducing $K_{\tilde{\varepsilon}} := K_{x_1,\tilde{\varepsilon}/2} \cup \cdots \cup K_{x_L,\tilde{\varepsilon}/2}$, we conclude that

$$\int_{\mathcal{F} \setminus K_{\tilde{\varepsilon}}} \mathcal{L}(x,y)\, d\rho(y) < \tilde{\varepsilon} \qquad \text{for all } x \in K\,. \tag{3.5.7}$$

Similarly, for all $x \in K$ we have

$$\int_{\mathcal{F} \setminus K_{\tilde{\varepsilon}}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) < \tilde{\varepsilon} \qquad \text{for sufficiently large } n \in \mathbb{N}\,.$$

According to (3.2.4), we obtain

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) &= 2 \int_K d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F} \setminus K_{\tilde{\varepsilon}}} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + 2 \int_K d(\tilde{\rho} - \rho)(x) \int_{K_{\tilde{\varepsilon}}} d\rho(y)\, \mathcal{L}(x,y) + \int_K d(\tilde{\rho} - \rho)(x) \int_K d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\,.\end{aligned}$$

Choosing $\tilde{\varepsilon} > 0$ suitably and making use of (3.5.7), the expression

$$\int_K d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F} \setminus K_{\tilde{\varepsilon}}} d\rho(y)\, \mathcal{L}(x,y) \leq 2\tilde{\varepsilon}\, \rho(K)$$

can be arranged to be arbitrarily small, giving rise to

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \Big[&2 \int_K d(\tilde{\rho} - \rho)(x) \int_{K_{\tilde{\varepsilon}}} d\rho(y)\, \mathcal{L}(x,y) \\ &+ \int_K d(\tilde{\rho} - \rho)(x) \int_K d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \Big] - \varepsilon\end{aligned}$$

for any $\varepsilon > 0$. Proceeding similarly as in the proof of Theorem 3.4.9 and Theorem 3.4.10, one can show that the term in square brackets is bigger than or equal to zero, up to an arbitrarily small error term. Since $\varepsilon > 0$ was chosen arbitrarily, we finally arrive at

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,,$$

which proves the claim. □

3.5.5. Existence of Minimizers under Variations of Finite Volume. We finally can prove the existence of minimizers in the sense of Definition 3.3.2.

THEOREM 3.5.9. *Assume that $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is continuous, bounded, and decays in entropy (see Definition 3.5.1). Moreover, assume that condition* (iv) *in §3.2.2 holds. Then ρ is a minimizer under variations of finite volume (see Definition 3.3.2).*

PROOF. The idea is to proceed similarly to the proof of Theorem 3.4.10. Considering variations of finite volume $\tilde{\rho} \in \mathfrak{B}_{\mathcal{F}}$ satisfying the conditions in (3.2.3) and introducing the set $B := \operatorname{supp}(\tilde{\rho} - \rho)$, we know that $(\tilde{\rho} - \rho)(B) = 0$ and thus $\rho(B) = \tilde{\rho}(B) < \infty$. Under the assumption that condition (iv) in §3.2.2 holds, the difference (3.2.4) is well-defined, giving rise to

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) = 2 \int_B d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_B d(\tilde{\rho} - \rho)(x) \int_B d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\,.\end{aligned}$$

For arbitrary $\tilde{\varepsilon} > 0$, we first approximate B by open sets $U \supset B$ due to regularity of ρ and $\tilde{\rho}$ such that

$$\rho(U \setminus B) < \tilde{\varepsilon}/4 \qquad \text{and} \qquad \tilde{\rho}(U \setminus B) < \tilde{\varepsilon}/4\,,$$

and then we approximate U from inside by compact sets K such that

$$\rho(U \setminus K) < \tilde{\varepsilon}/4 \qquad \text{and} \qquad \tilde{\rho}(U \setminus K) < \tilde{\varepsilon}/4\,.$$

This gives rise to

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) = 2 \int_{B\setminus K} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + 2 \int_K d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_{B\setminus K} d(\tilde{\rho} - \rho)(x) \int_B d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) + \int_K d(\tilde{\rho} - \rho)(x) \int_B d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\,.\end{aligned}$$

Proceeding similarly to the proof of Theorem 3.4.10 and Theorem 3.5.8, one arrives at

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \Bigg[2 \int_K d(\tilde{\rho} - \rho)(x) \int_{K_{\tilde{\varepsilon}}} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_K d(\tilde{\rho} - \rho)(x) \int_K d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\Bigg] - \varepsilon\end{aligned}$$

for any $\varepsilon > 0$ by suitably choosing $U \supset B$ and $K \subset U$. Applying the same arguments as in the proof of Theorem 3.5.8, one can show that the term in square brackets is bigger than or equal to zero, up to an arbitrarily small error term. Since $\varepsilon > 0$ was chosen arbitrarily, we finally arrive at

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,,$$

which proves the claim. □

Theorem 3.5.9 concludes the existence theory in the σ-locally compact setting.

CHAPTER 4

Causal Variational Principles in the Infinite-Dimensional Setting

ABSTRACT. We provide a method for constructing (possibly non-trivial) measures on non-locally compact Polish subspaces of infinite-dimensional separable Banach spaces which, under suitable assumptions, are minimizers of causal variational principles in the non-locally compact setting. Moreover, for non-trivial minimizers the corresponding Euler-Lagrange equations are derived. The method is to exhaust the underlying Banach space by finite-dimensional subspaces and to prove existence of minimizers of the causal variational principle restricted to these finite-dimensional subsets of the Polish space under suitable assumptions on the Lagrangian. This gives rise to a corresponding sequence of minimizers. Restricting the resulting sequence to countably many compact subsets of the Polish space, by considering the resulting diagonal sequence we are able to construct a regular measure on the Borel algebra over the whole topological space. For continuous Lagrangians of bounded range it can be shown that, under suitable assumptions, the obtained measure is a (possibly non-trivial) minimizer under variations of compact support. Under additional assumptions, we prove that the constructed measure is a minimizer under variations of finite volume and solves the corresponding Euler-Lagrange equations. Afterwards, we extend our results to continuous Lagrangians vanishing in entropy. Finally, assuming that the obtained measure is locally finite, topological properties of spacetime are worked out and a connection to dimension theory is established.

4.1. Introduction

In the theory of causal fermion systems, spacetime and the structures therein are described by a minimizer of the so-called causal action principle (for an introduction to the physical background and the mathematical context, we refer the interested reader to §4.2.1, the textbook [**59**] and the survey articles [**62, 64**]). *Causal variational principles* evolved as a mathematical generalization of the causal action principle [**56, 65**], and were studied in more detail in [**66**] (see Chapter 3). The starting point in [**66**] is a second-countable, locally compact Hausdorff space $\mathcal{F}$ together with a non-negative function

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+ := [0, \infty)$$

(the *Lagrangian*) which is assumed to be lower semi-continuous, symmetric and positive on the diagonal. The causal variational principle is to minimize the *action* $\mathcal{S}$ defined as the double integral over the Lagrangian

$$\mathcal{S}(\rho) = \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y)$$

under variations of the measure ρ within the class of regular Borel measures on $\mathcal{F}$, keeping the (possibly infinite) total volume $\rho(\mathcal{F})$ fixed (*volume constraint*). The aim of this chapter is to extend the existence theory for minimizers of such variational principles to the case that $\mathcal{F}$ is non-locally compact and the total volume is infinite. We also work out the corresponding Euler-Lagrange (EL) equations.

In order to put the chapter into the mathematical context, in [**53**] it was proposed to formulate physics by minimizing a new type of variational principle in spacetime. The suggestion in [**53**, Section 3.5] led to the causal action principle in discrete spacetime, which was first analyzed mathematically in [**54**]. A more general and systematic inquiry of causal variational principles on measure spaces was carried out in [**56**]. In this article, the existence of minimizers is proven in the case that the total volume is finite. In [**65**], the setting is generalized to non-compact manifolds of possibly infinite volume and the corresponding EL equations are analyzed. However, the existence of minimizers is not proved. This is done in [**66**] in the slightly more general setting of second-countable, locally compact Hausdorff spaces. In this chapter, we extend the results of [**66**] by developing the existence theory in the non-locally compact setting.

The main difficulty in dealing with non-locally compact spaces is that it is no longer possible to restrict attention to compact neighborhoods. Moreover, it turns out that we can no longer assume that the underlying topological space $\mathcal{F}$ is σ-compact. As a consequence, at first sight it is not clear how to construct global measures on the whole topological space at all. The way out is to introduce a countable collection of suitable compact subsets, which indeed allows us to construct a global measure ρ on $\mathcal{F}$. For simplicity, we first assume that the Lagrangian is of *bounded range* (see Definition 4.3.7). In this case, the minimizing property of the measure ρ is proved in two steps: We first show that ρ is a minimizer under *variations of compact support*. In a second step, we extend this result to variations of finite volume under the assumption that property (iv) in §4.2.2 holds, i.e.

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty\,.$$

Afterwards, we generalize our results to Lagrangians which do not have bounded range, but instead have suitable decay properties. To this end, we consider Lagrangians *vanishing in entropy* (see Definition 4.6.2). Introducing *spacetime* as the support of the measure ρ, we finally analyze topological properties of spacetime; moreover, a connection to dimension theory is established.

The chapter is organized as follows. In Section 4.2 we give a short physical motivation (§4.2.1) and recall the main definitions and existence results as obtained in [**66**] (§4.2.2). In Section 4.3 we begin by working out important topological properties of infinite-dimensional causal fermion systems (§4.3.1); afterwards, we introduce causal variational principles in the non-locally compact (or infinite-dimensional) setting by considering non-locally compact Polish subspaces of (infinite-dimensional) separable Banach spaces (§4.3.2). Exhausting the underlying Banach space by finite-dimensional subspaces and making use of the results in [**66**], the existence of minimizers is proven for the causal variational principle restricted to finite-dimensional subspaces (§4.3.3). In Section 4.4 we provide a method for constructing a regular global measure on the Borel algebra over the whole non-locally compact Polish space. More precisely, we first introduce a countable collection of compact subsets of the underlying topological space (§4.4.1). Next, making use of Prohorov's theorem and applying Cantor's diagonal argument, we are able to construct a (possibly non-trivial) regular measure on the whole topological space (§4.4.2).

Finally, we derive useful properties of the constructed measure (§4.4.3). In Section 4.5 we prove that, under suitable assumptions, this measure is a minimizer for continuous Lagrangians of bounded range (see Definition 4.3.7). More precisely, we first introduce an appropriate assumption on the obtained measure (see condition (B) in §4.5.1). Next, we prove that the obtained measure is a minimizer under variations of finite-dimensional compact support (§4.5.2) as well as a minimizer under variations of compact support (§4.5.3). Under additional assumptions we show that the constructed measure is a minimizer under variations of finite volume (§4.5.4). Assuming that the measure under consideration is non-zero, we prove that the corresponding Euler-Lagrange (EL) equations are satisfied (§4.5.5). The goal in Section 4.6 is to weaken the assumption that the Lagrangian is of bounded range. To this end, we introduce Lagrangians *vanishing in entropy* (see Definition 4.6.1) which generalize the notion of Lagrangians decaying in entropy (see Definition 4.2.8). The concept of Lagrangians vanishing in entropy (§4.6.1) can be extended to non-locally compact topological spaces (see Definition 4.6.2). For such Lagrangians, we repeat the above construction steps, thus giving rise to a regular measure on the underlying topological space (§4.6.2). It is shown that, under suitable assumptions, the considered measure is minimizer of the causal action under variations of compact support as well as under variations of finite volume (§4.6.3). We finally derive the corresponding EL equations (§4.6.4). Introducing spacetime as the support of the minimizing measure under consideration, in Section 4.7 we conclude this chapter by analyzing topological properties of spacetime and establishing a connection to dimension theory. To this end, we first recall some concepts from dimension theory (§4.7.1), and afterwards apply them to causal fermion systems (§4.7.2). In the appendix we summarize useful results which will be referred to frequently: Appendix A.2 is dedicated to the proof that causal fermion systems are Polish (see Theorem A.2.1); the main result in Appendix A.3 states that the support of locally finite measures on Polish spaces is σ-compact (see Lemma A.3.2).

4.2. Physical Background and Mathematical Preliminaries

4.2.1. Physical Context and Motivation. The purpose of this subsection is to outline a few concepts of causal fermion systems and to explain how this chapter fits into the general physical context and the ongoing research program. The reader not interested in the physical background may skip this section.

The theory of causal fermion systems is a recent approach to fundamental physics motivated originally in order to resolve shortcomings of relativistic quantum field theory (QFT). Namely, due to ultraviolet divergences, perturbative quantum field theory is well-defined only after regularization, which is usually understood as a set of prescriptions for how to make divergent integrals finite (e.g. by introducing a suitable "cutoff" in momentum space). The regularization is then removed using the renormalization procedure. However, this concept is not convincing from neither the physical nor the mathematical point of view. More precisely, in view of Heisenberg's uncertainty principle, physicists infer a correspondence between large momenta and small distances. Because of that, the regularization length is often associated to the Planck length $\ell_P \approx 1.6 \cdot 10^{-35}$ m. Accordingly, by introducing an ultraviolet cutoff in momentum space, one disregards distances which are smaller than the Planck length. As a consequence, the microscopic structure of spacetime is completely unknown. Unfortunately, at present there is no consensus on what the correct mathematical model for "Planck scale physics" should be.

The simplest and maybe most natural approach is to assume that on the Planck scale, spacetime is no longer a continuum but becomes in some way "discrete." This is

the starting point in the monograph [**53**], where the physical system is described by an ensemble of wave functions in a discrete spacetime. Motivated by the Lagrangian formulation of classical field theory, physical equations are formulated by a variational principle in discrete spacetime. In the meantime, this setting was generalized and developed to the theory of causal fermion systems. It is an essential feature of the approach that spacetime does not enter the variational principle a priori, but instead it emerges when minimizing the action. Thus causal fermion systems allow for the description of both discrete and continuous spacetime structures.

In order to get the connection to this chapter, let us briefly outline the main structures of causal fermion systems. As initially introduced in [**61**], a *causal fermion system* consists of a triple $(\mathcal{H}, \mathcal{F}, \rho)$ together with an integer $n \in \mathbb{N}$, where $\mathcal{H}$ denotes a complex Hilbert space, $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ being the set of all self-adjoint operators on $\mathcal{H}$ of finite rank with at most n positive and at most n negative eigenvalues, and ρ being a positive measure on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$ (referred to as *universal measure*). Then for any $x, y \in \mathcal{F}$, the product xy is an operator of rank at most $2n$. Denoting its non-trivial eigenvalues (counting algebraic multiplicities) by $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy} \in \mathbb{C}$, and introducing the spectral weight $|.|$ of an operator as the sum of the absolute values of its eigenvalues, the *Lagrangian* can be introduced as a mapping

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+ , \qquad \mathcal{L}(x,y) = \big|(xy)^2\big| - \frac{1}{2n} |xy|^2 \;.$$

As being of relevance for this article, we point out that the Lagrangian is a continuous function which is symmetric in the sense that

$$\mathcal{L}(x,y) = \mathcal{L}(y,x) \qquad \text{for all } x, y \in \mathcal{F} .$$

In analogy to classical field theory, one defines the *causal action* by

$$\mathcal{S}(\rho) = \iint_{\mathcal{F} \times \mathcal{F}} \mathcal{L}(x,y)\, d\rho(x)\, d\rho(y) .$$

Finally, the corresponding *causal action principle* is introduced by varying the measure ρ in the class of regular measures on $\mathcal{B}(\mathcal{F})$ under additional constraints (which assert the existence of non-trivial minimizers). Given a minimizing measure ρ, *spacetime* M is defined as its support,

$$M := \operatorname{supp} \rho .$$

As being outlined in detail in [**59**], critical points of the causal action give rise to Euler-Lagrange (EL) equations, which describe the dynamics of the causal fermion system. In a certain limiting case, the so-called *continuum limit*, one can establish a connection to the conventional formulation of physics in a spacetime continuum. In this limiting case, the EL equations give rise to classical field equations like the Maxwell and Einstein equations. Moreover, quantum mechanics is obtained in a limiting case, and close connections to relativistic quantum field theory have been established (for details see [**58**] and [**63**]).

In order for the causal action principle to be mathematically sensible, the existence theory is of crucial importance. If the dimension of the Hilbert space $\mathcal{H}$ is finite, the existence of minimizers was proven in [**56**, Section 2] (based on existence results in discrete spacetime [**54**]), giving rise to minimizing measures ρ on $\mathcal{F}$ of finite total volume $\rho(\mathcal{F}) < \infty$. For this reason, it remains to extend these existence results by developing the existence theory in the case that $\mathcal{H}$ is infinite-dimensional. Then the total volume $\rho(\mathcal{F})$ is necessarily infinite (for a counterexample see [**59**, Exercise 1.3]). In the resulting *infinite-dimensional setting* (i.e. $\dim \mathcal{H} = \infty$ and $\rho(\mathcal{F}) = \infty$), the task is to deal with minimizers of infinite

total volume on non-locally compact spaces. In preparation, the existence theory of minimizers of possibly infinite total volume $\rho(\mathcal{F})$ on locally compact spaces is developed in [**66**] in sufficient generality (cf. Chapter 3). The remaining second step, which involves the difficulty of dealing with non-locally compact spaces, is precisely the objective of this chapter.

4.2.2. Causal Variational Principles in the σ-Locally Compact Setting. Before introducing causal variational principles on non-locally compact spaces in Section 4.3 below, we now recall the main results in the less general situation of causal variational principles in the σ-locally compact setting [**66**] (see Chapter 3) which are based on results concerning causal variational principles in the non-compact setting as studied in [**65**, Section 2].

The starting point in [**66**] is a second-countable, locally compact topological Hausdorff space $\mathcal{F}$. We let ρ be a (positive) measure on the Borel algebra over $\mathcal{F}$ (referred to as *universal measure*). Moreover, let $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ be a non-negative function (the *Lagrangian*) with the following properties:

(i) $\mathcal{L}$ is symmetric, i.e. $\mathcal{L}(x,y) = \mathcal{L}(y,x)$ for all $x, y \in \mathcal{F}$.
(ii) $\mathcal{L}$ is lower semi-continuous, i.e. for all sequences $x_n \to x$ and $y_{n'} \to y$,

$$\mathcal{L}(x,y) \leq \liminf_{n,n' \to \infty} \mathcal{L}(x_n, y_{n'}) \,.$$

The *causal variational principle* is to minimize the action

$$\mathcal{S}(\rho) = \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \tag{4.2.1}$$

under variations of the measure ρ, keeping the total volume $\rho(\mathcal{F})$ fixed (*volume constraint*). The papers [**65, 66**] mainly focus on the case that the total volume $\rho(\mathcal{F})$ is infinite. In order to implement the volume constraint and to derive the corresponding Euler-Lagrange equations, in [**65**] one makes the following additional assumptions:

(iii) The measure ρ is *locally finite* (meaning that any $x \in \mathcal{F}$ has an open neighborhood $U \subset \mathcal{F}$ with $\rho(U) < \infty$).
(iv) The function $\mathcal{L}(x,.)$ is ρ-integrable for all $x \in \mathcal{F}$ and

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty \,. \tag{4.2.2}$$

By Fatou's lemma, the integral in (4.2.2) is lower semi-continuous in the variable x.

A measure on the Borel algebra which satisfies (iii) will be referred to as a *Borel measure* (in the sense of [**76**]), and the set of Borel measures on $\mathcal{F}$ shall be denoted by $\mathfrak{B}_{\mathcal{F}}$. Moreover, the Borel σ-algebra over $\mathcal{F}$ is denoted by $\mathcal{B}(\mathcal{F})$. A Borel measure is said to be *regular* if it is inner and outer regular (cf. [**40**, Definition VIII.1.1]). An inner regular Borel measure is also called a *Radon measure* [**143**].

In [**65, 66**] one varies in the following class of measures:

DEFINITION 4.2.1. *Given a regular Borel measure ρ on $\mathcal{F}$, a regular Borel measure $\tilde{\rho}$ on $\mathcal{F}$ is said to be a* **variation of finite volume** *if*

$$\big|\tilde{\rho} - \rho\big|(\mathcal{F}) < \infty \qquad \text{and} \qquad \big(\tilde{\rho} - \rho\big)(\mathcal{F}) = 0 \,, \tag{4.2.3}$$

where the total variation $|\tilde{\rho}-\rho|$ of two possibly infinite measures ρ and $\tilde{\rho}$ on $\mathcal{B}(\mathcal{F})$ is defined in [**66**, §2.2] *(see §3.2.2) as follows: We say that $|\tilde{\rho} - \rho| < \infty$ if there exists $B \in \mathcal{B}(\mathcal{F})$*

with $\rho(B), \tilde{\rho}(B) < \infty$ *such that* $\rho|_{\mathcal{F}\setminus B} = \tilde{\rho}|_{\mathcal{F}\setminus B}$. *In this case,*

$$(\tilde{\rho} - \rho)(\Omega) := \tilde{\rho}(B \cap \Omega) - \rho(B \cap \Omega)$$

for any Borel set $\Omega \subset \mathcal{F}$.

Given a regular Borel measure $\rho \in \mathfrak{B}_{\mathcal{F}}$ and assuming that (i), (ii) and (iv) hold, for every variation of finite volume $\tilde{\rho} \in \mathfrak{B}_{\mathcal{F}}$ the difference of the actions as given by

$$\begin{aligned}\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) &= \int_{\mathcal{F}} d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y) + \int_{\mathcal{F}} d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y)\end{aligned} \tag{4.2.4}$$

is well-defined in view of [**65**, Lemma 2.1]. For clarity, we point out that condition (iii) is not required in order for (4.2.4) to hold.

Note that assumptions (iii) and (iv) are dropped in [**66**]. The causal variational principle in the σ-locally compact setting [**66**] is then defined as follows.

DEFINITION 4.2.2. *Let* $\mathcal{F}$ *be a second-countable, locally compact Hausdorff space, and let the Lagrangian* $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ *be a symmetric and lower semi-continuous function (see conditions* (i) *and* (ii) *above). Moreover, we assume that* $\mathcal{L}$ *is strictly positive on the diagonal, i.e.*

$$\mathcal{L}(x,x) > 0 \qquad \text{for all } x \in \mathcal{F}\,.$$

The **causal variational principle on σ-locally compact spaces** *is to minimize the causal action* (4.2.1) *under variations of finite volume (see Definition 4.2.1).*

We point out that (iv) is a sufficient condition for (4.2.4) to hold. However, since conditions (iii) and (iv) are not imposed in [**66**], it is a priori not clear whether the integrals in (4.2.4) exist. For this reason, condition (4.2.4) is included into the definition of a minimizer:

DEFINITION 4.2.3. *A regular Borel measure* ρ *on* $\mathcal{F}$ *is said to be a* **minimizer** *of the causal action* ***under variations of finite volume*** [**66**] *if the difference* (4.2.4) *is well-defined and non-negative for all regular Borel measures* $\tilde{\rho}$ *on* $\mathcal{F}$ *satisfying* (4.2.3)*,*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

We denote the support of the measure ρ by M,

$$M := \operatorname{supp} \rho = \mathcal{F} \setminus \bigcup \big\{\Omega \subset \mathcal{F} \;\;:\;\; \Omega \text{ is open and } \rho(\Omega) = 0\big\}$$

(thus the support is the set of all points for which every open neighborhood has a strictly positive measure; for details and generalizations see [**43**, Subsection 2.2.1]). According to Definition 4.2.1, the condition $|\tilde{\rho} - \rho| < \infty$ implies that there exists some Borel set $B \subset \mathcal{F}$ with $\rho(B), \tilde{\rho}(B) < \infty$ and $\rho|_{\mathcal{F}\setminus B} = \tilde{\rho}|_{\mathcal{F}\setminus B}$. In particular, $\rho|_B$ and $\tilde{\rho}|_B$ are finite Borel measures on $\mathcal{B}(B)$, and thus have support (see [**16**, Proposition 7.2.9]). Furthermore, the signed measure $\tilde{\rho} - \rho$ has support.

We now recall some results from [**66**] which will be referred to frequently (we refer to Chapter 3 for details). The first existence results in [**66**] are based on the assumption that the Lagrangian is of compact range. For convenience, let us state the definition (see Definition 3.3.3).

DEFINITION 4.2.4. *The Lagrangian has* **compact range** *if for every $K \subset \mathcal{F}$ compact there is a compact set $K' \subset \mathcal{F}$ such that*

$$\mathcal{L}(x,y) = 0 \qquad \text{for all } x \in K \text{ and } y \notin K' .$$

Moreover, the definition of minimizers under variations of compact support plays an important role in [**66**] (see Definition 3.4.8).

DEFINITION 4.2.5. *A regular Borel measure ρ on $\mathcal{F}$ is said to be a* **minimizer under variations of compact support** [**66**] *of the causal action if for any regular Borel measure $\tilde{\rho}$ on $\mathcal{F}$ which satisfies* (4.2.3) *such that the signed measure $\tilde{\rho} - \rho$ is compactly supported, the inequality*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0$$

holds.

Assuming that the Lagrangian is of compact range, the main results in [**66**] can be summarized as follows (see Theorem 3.4.2, Theorem 3.4.9 and Theorem 3.4.10).

THEOREM 4.2.6 (**Euler-Lagrange equations**). *Let $\mathcal{F}$ be a second-countable, locally compact Hausdorff space, and assume that $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is continuous and of compact range. Then there exists a regular Borel measure ρ on $\mathcal{F}$ which satisfies the Euler-Lagrange equations*

$$\ell|_{\text{supp}\,\rho} \equiv \inf_{x \in \mathcal{F}} \ell(x) = 0 ,$$

where $\ell \in C(\mathcal{F})$ is defined by

$$\ell(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) - 1 . \tag{4.2.5}$$

Combining [**66**, Theorem 4.9 and Theorem 4.10], we obtain the following result.

THEOREM 4.2.7. *Assume that $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is continuous and of compact range. Then there is a regular Borel measure ρ on $\mathcal{F}$ which is a minimizer under variations of compact support* [**66**] *(see Definition 4.2.5). Under the additional assumptions that the Lagrangian $\mathcal{L}$ is bounded and condition* (iv) *is satisfied (see* (4.2.2)*), the measure ρ is a minimizer under variations of finite volume* [**66**] *(see Definition 4.2.3).*

In [**66**, Section 5] (see Section 3.5) it was shown that the assumption that the Lagrangian $\mathcal{L}$ is of compact range can be weakened. To this end, we recall that every second-countable, locally compact Hausdorff space can be endowed with a Heine-Borel metric (for details we refer to the explanations in [**66**, §3.1 and §5.1]). Given an unbounded Heine-Borel metric on the second-countable, locally compact space $\mathcal{F}$, for any $r > 0$ and $x \in \mathcal{F}$ the closed ball $\overline{B_r(x)}$ is compact, and hence can be covered by finitely many balls of radius $\delta > 0$. The smallest such number is denoted by $E_x(r, \delta)$ and is called *entropy*. This gives rise to Lagrangians decaying in entropy, being defined as follows (see Definition 3.5.1 or [**66**, Definition 5.1]).

DEFINITION 4.2.8. *Assume that $\mathcal{F}$ is endowed with an unbounded Heine-Borel metric d. The Lagrangian $\mathcal{L}\colon \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is said to* **decay in entropy** *if the following conditions are satisfied:*

(a) $c := \inf_{x \in \mathcal{F}} \mathcal{L}(x,x) > 0$.

(b) *There is a compact set $K \subset \mathcal{F}$ such that*

$$\delta := \inf_{x \in \mathcal{F} \setminus K} \sup \Big\{ s \in \mathbb{R} \;:\; \mathcal{L}(x,y) \geq \frac{c}{2} \quad \text{for all } y \in B_s(x) \Big\} > 0 .$$

(c) *The Lagrangian has the following decay property: There is a monotonically decreasing, integrable function* $f \in L^1(\mathbb{R}^+, \mathbb{R}_0^+)$ *such that*

$$\mathcal{L}(x,y) \leq \frac{f\big(d(x,y)\big)}{C_x\big(d(x,y),\delta\big)} \qquad \text{for all } x,y \in \mathcal{F} \text{ with } x \neq y\,,$$

where

$$C_x(r,\delta) := C\, E_x(r+2,\delta) \qquad \text{for all } r > 0\,,$$

and the constant C *is given by*

$$C := 1 + \frac{2}{c} < \infty\,.$$

We point out that the above definition of Lagrangians decaying in entropy as introduced in [**66**, Section 5] (see Section 3.5) requires an *unbounded* Heine-Borel metric. For a more general definition we refer to §4.6.1 (see Definition 4.6.1).

For clarity we note that, if $(\mathcal{H}, \mathcal{F}, \rho)$ is a causal fermion system with $\dim(\mathcal{H}) < \infty$, the space $\mathrm{L}(\mathcal{H})$ of bounded linear operators on $\mathcal{H}$ is finite-dimensional. Combining the fact that all norms on finite-dimensional vector spaces are equivalent with the Heine-Borel theorem yields that the Fréchet metric induced by the operator norm on the vector space $\mathrm{L}(\mathcal{H})$ is an unbounded Heine-Borel metric on $\mathcal{F}$.

Let us now recall the main results in [**66**] under the assumption that the Lagrangian decays in entropy (see Theorem 3.5.5, Theorem 3.5.8 and Theorem 3.5.9).

THEOREM 4.2.9. *Let* $\mathcal{F}$ *be a second-countable, locally compact Hausdorff space, and assume that* $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ *is continuous and decays in entropy. Then there exists a regular Borel measure* ρ *on* $\mathcal{F}$ *which satisfies the Euler-Lagrange equations*

$$\ell|_{\operatorname{supp}\rho} \equiv \inf_{x \in \mathcal{F}} \ell(x) = 0\,,$$

where $\ell \in C(\mathcal{F})$ *is defined by* (4.2.5).

The following theorem ensures the existence of minimizing Borel measures.

THEOREM 4.2.10. *Assume that* $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ *is continuous and decays in entropy. Then there is a regular Borel measure* ρ *on* $\mathcal{F}$ *which is a minimizer under variations of compact support* [**66**]. *Under the additional assumptions that the Lagrangian* $\mathcal{L}$ *is bounded and condition* (iv) *is satisfied (see* (4.2.2)*), the measure* ρ *is a minimizer under variations of finite volume* [**66**].

The goal of this chapter is to extend the above results to the infinite-dimensional setting.

4.3. Causal Variational Principles in the Non-Locally Compact Setting

4.3.1. Motivation: Infinite-Dimensional Causal Fermion Systems. As explained in Section 4.1, causal variational principles evolved as a mathematical generalization of the causal action principle in order to develop the existence theory for causal fermion systems. In order to point out the connection to causal variational principles in the non-locally compact setting, let us briefly recall the basic structures of causal fermion systems (for details cf. [**59**, §1.1.1]). By definition, causal fermion systems are characterized by a separable complex Hilbert space $\mathcal{H}$, an integer $s \in \mathbb{N}$ (the so-called *spin dimension*) and a measure ρ defined on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$, where $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ consists of all self-adjoint operators on $\mathcal{H}$ which have at most s positive and at most s

negative eigenvalues. This gives rise to a triple $(\mathcal{H}, \mathcal{F}, \rho)$. The set $\mathcal{F}$ can be endowed with the topology induced by the operator norm on $\mathrm{L}(\mathcal{H})$, thus becoming a topological space. More precisely, denoting the Fréchet metric induced by the operator norm on $\mathrm{L}(\mathcal{H})$ by d, the space $(\mathcal{F}, d)$ is a separable complete metric space (Theorem A.2.1).

Whenever $\dim(\mathcal{H}) < \infty$, the topological space $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ is locally compact. On the contrary, whenever $\mathcal{H}$ is an infinite-dimensional Hilbert space, the corresponding set $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ is non-locally compact (see Lemma 4.3.3 below). In preparation, let us first state the following results.

PROPOSITION 4.3.1. *Any locally compact Banach space X is finite-dimensional.*

PROOF. See for instance [**140**, Theorem 1.22]. □

COROLLARY 4.3.2. *Any infinite-dimensional Banach space X is non-locally compact. The same holds true for open subsets of X.*

PROOF. This is an immediate consequence of Proposition 4.3.1. □

LEMMA 4.3.3. *Let $\mathcal{H}$ be an infinite-dimensional, separable complex Hilbert space, and let $\mathcal{F}^{\text{reg}} \subset \mathrm{L}(\mathcal{H})$ be the set of self-adjoint operators which have exactly s positive and exactly s negative eigenvalues for some $s \in \mathbb{N}$. Then $\mathcal{F}^{\text{reg}}$ is non-locally compact.*

PROOF. Since $\mathcal{F}^{\text{reg}}$ is a Banach manifold (for details see [**67**]), it can be covered by an atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ for some index set A (cf. [**169**, Chapter 73]). In particular, every point $x \in U^{\text{reg}}$ is contained in some open set U_α, whose image $V_\alpha := \phi_\alpha(U_\alpha)$ is open in some infinite-dimensional Banach space X_α. Due to Corollary 4.3.2, the set V_α is non-locally compact. As the mapping ϕ_α is a homeomorphism, we deduce that $U_\alpha \subset \mathcal{F}^{\text{reg}}$ is non-locally compact for any $\alpha \in A$, which proves the claim. □

Considering an *infinite-dimensional*, separable complex Hilbert space $\mathcal{H}$, then the set $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ as introduced in [**59**] is non-locally compact and Polish (see Lemma 4.3.3 and Theorem A.2.1). Our goal in the following is to prove the existence of a regular (possibly non-locally finite) measure ρ on the Borel algebra $\mathcal{B}(\mathcal{F})$ such that ρ is a minimizer of the corresponding causal action principle, giving rise to a causal fermion system $(\mathcal{H}, \mathcal{F}, \rho)$. Instead of immediately delving into the corresponding causal action principle (see [**59**, §1.1.1]), we deal with causal variational principles on $\mathcal{H}$, which can be viewed as generalizations of the causal action principle (as introduced in [**56**], [**65**] and considered in more detail in [**66**]). Corresponding results concerning the causal action principle are then obtained as a special case. With this in mind, it suffices to prove the existence of minimizers of the causal variational principle (4.3.2) under the constraints (4.2.3) in the non-locally compact setting as introduced in Definition 4.3.4.

In order to motivate the basic definitions in §4.3.2 below, we note that $\mathcal{F} \subset \mathcal{K}(\mathcal{H})$, where by $\mathcal{K}(\mathcal{H}) \subset \mathrm{L}(\mathcal{H})$ we denote the set of compact operators on $\mathcal{H}$. Since $\mathcal{H}$ is a separable, infinite-dimensional complex Hilbert space, let us point out that $\mathcal{K}(\mathcal{H})$ is a separable Banach space in view of [**99**, §12.E]. Making use of Proposition 4.3.1, Corollary 4.3.2 and Lemma 4.3.3, we infer that $\mathcal{K}(\mathcal{H}) \supset \mathcal{F}$ is infinite-dimensional. This allows us to approximate $\mathcal{K}(\mathcal{H})$ by finite-dimensional subspaces: There is a sequence of finite-dimensional subspaces $(L_n)_{n \in \mathbb{N}}$ in $\mathcal{K}(\mathcal{H})$ with $L_n \subset L_{n+1}$ for all $n \in \mathbb{N}$ such that $\bigcup_{n \in \mathbb{N}} L_n$ is dense in $\mathcal{K}(\mathcal{H})$ (cf. [**5**, Lemma 7.1] or [**21**]). Thus for every $n \in \mathbb{N}$, the space $\mathcal{F}^{(n)} := \mathcal{F} \cap L_n$ is locally compact. Denoting by d the (Fréchet) metric induced by the operator norm on $\mathrm{L}(\mathcal{H})$, the space $(\mathcal{F}, d)$ is Polish (cf. Theorem A.2.1) and thus a separable metric space. Due to [**4**, Corollary 3.5], the subsets $\mathcal{F}^{(n)}$ are separable for every $n \in \mathbb{N}$. Since separable

metric spaces are second-countable, for each $n \in \mathbb{N}$ the set $\mathcal{F}^{(n)}$ is a second-countable, locally compact Hausdorff space. Moreover, from Lemma 4.3.3 we conclude that $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ is non-locally compact.

In order to treat the corresponding causal variational principle in sufficient generality, it seems reasonable to vary in the class of regular, not necessarily locally finite measures on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$ (as intended in the textbook [**59**, §1.1.1]). As mentioned in [**59**], the causal action principle is ill-posed if the total volume $\rho(\mathcal{F})$ is finite and the Hilbert space $\mathcal{H}$ is infinite-dimensional. However, the causal action principle does make mathematical sense in the so-called *infinite-dimensional setting* where $\mathcal{H}$ is infinite-dimensional and the total volume is infinite, i.e. $\rho(\mathcal{F}) = \infty$. These considerations motivate causal variational principles in the infinite-dimensional (or non-locally compact) setting as defined in the next subsection.

4.3.2. Basic Definitions. Let us first state the causal variational principle in the non-locally compact setting and discuss its difficulties afterwards.

DEFINITION 4.3.4. *Assume that X is a separable, infinite-dimensional Banach space, and let $\mathcal{F} \subset X$ be a* non-locally compact *Polish space. Moreover, assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is a symmetric and lower semi-continuous function (see conditions* (i) *and* (ii) *in §4.2.2) which is strictly positive on the diagonal, i.e.*

$$\mathcal{L}(x,x) > 0 \qquad \text{for all } x \in \mathcal{F}\,. \tag{4.3.1}$$

The **causal variational principle in the non-locally compact setting**[1] *is to*

$$\text{minimize} \qquad \mathcal{S}(\rho) := \int_{\mathcal{F}} \int_{\mathcal{F}} d\rho(x)\, d\rho(y)\, \mathcal{L}(x,y) \tag{4.3.2}$$

under variations of finite volume (see Definition 4.3.5 below) in the class of all regular measures on $\mathcal{B}(\mathcal{F})$ (in the sense of [**76**]*, cf.* [**66**]*) with $\rho(\mathcal{F}) = \infty$.*

The condition (4.3.1) is needed in order to avoid trivial minimizers supported at $x \in \mathcal{F}$ with $\mathcal{L}(x,x) = 0$ (see [**68**, Section 1.2]). Furthermore, condition (4.3.1) is a plausible assumption in view of [**59**, Exercise 1.2]. Namely, given a minimizing measure ρ of the causal action principle (4.3.2), there exists a real constant c such that $\mathrm{tr}(x) = c$ for all $x \in \operatorname{supp}\rho$ according to [**59**, Proposition 1.4.1]. Under the reasonable assumption that $c \neq 0$ (cf. [**59**, §1.4.1]), we may conclude that $\mathcal{L}(x,x) > 0$ for all $x \in \mathcal{F}$ in view of [**59**, Exercise 1.2]. This motivates as well as justifies the assumption that the Lagrangian is strictly positive on the diagonal.

Dropping the assumption that the measures under consideration are locally finite, we slightly adapt the definition of a minimizer of the causal action as follows.

DEFINITION 4.3.5. *A regular measure ρ on $\mathcal{B}(\mathcal{F})$ is said to be a* **minimizer** *of the causal action* ***under variations of finite volume*** *if the difference* (4.2.4) *is well-defined and non-negative for all regular measures $\tilde{\rho}$ on $\mathcal{B}(\mathcal{F})$ satisfying* (4.2.3)*,*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

Given a measure ρ on the Borel algebra $\mathcal{B}(\mathcal{F})$ and assuming that $\tilde{\rho}$ is a variation of finite volume, in view of Definition 4.3.5 there exists $B \in \mathcal{B}(\mathcal{F})$ such that $\rho|_{\mathcal{F}\setminus B} = \tilde{\rho}|_{\mathcal{F}\setminus B}$. In particular, the measures $\rho|_B$ and $\tilde{\rho}|_B$ are finite. Henceforth, whenever ρ is locally finite, then the same holds true for the measure $\tilde{\rho}|_{\mathcal{F}\setminus B}$. From the fact that $\tilde{\rho}|_B$ is a finite measure

[1]For clarity we point out that "causal variational principles in the non-locally compact setting" and "causal variational principles in the infinite-dimensional setting" are used synonymously.

we conclude that $\tilde{\rho}|_B$ is locally finite. Consequently, the measure $\tilde{\rho}$ is locally finite if ρ is so. For this reason, Definition 4.3.5 can be viewed as a generalization of Definition 4.2.3 (cf. [**66**, Definition 2.1]). The same holds for Definition 4.3.6 below.

DEFINITION 4.3.6. *A regular measure ρ on $\mathcal{B}(\mathcal{F})$ is said to be a* **minimizer under variations of compact support** *of the causal action if for any regular measure $\tilde{\rho}$ on $\mathcal{B}(\mathcal{F})$ which satisfies* (4.2.3) *such that the signed measure $\tilde{\rho}-\rho$ is compactly supported, the inequality*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0$$

holds.

Let us now point out some difficulties regarding causal variational principles on non-locally compact spaces. First of all, let us recall that a topological space is called hemicompact if there is a sequence $(K_n)_{n\in\mathbb{N}}$ of compact subsets of X such that any compact set $K \subset X$ is contained in K_n for some $n \in \mathbb{N}$ (see [**161**, §17I]). Since $\mathcal{F}$ is first-countable and non-locally compact, by virtue of [**41**, Exercise 3.4.E] we conclude that $\mathcal{F}$ cannot be hemicompact.

Next, by contrast to the σ-locally compact setting as worked out in [**66**] (for details see Chapter 3), it is in general not even possible to assume that $\mathcal{F}$ is σ-compact, as the following argument shows: Every Polish space (as well as every locally compact Hausdorff space) is Baire according to [**99**, Theorem (8.4)].[2] In view of [**161**, §25B], a σ-compact topological space X is Baire if and only if the set of points at which X is locally compact is dense in X. Given an infinite-dimensional Hilbert space $\mathcal{H}$, and defining $\mathcal{F} \subset \mathcal{K}(\mathcal{H})$ in analogy to [**59**] (see §4.3.1), then $\mathcal{F}$ is a Polish space (see Appendix A.2). Consequently, the assumption that $\mathcal{F}$ is σ-compact implies that there exists $x \in \mathcal{F}$ being contained in a compact neighborhood $K \subset \mathcal{F}$ with $K^\circ \neq \varnothing$. From this we conclude that the intersection $K^{\text{reg}} := K \cap \mathcal{F}^{\text{reg}}$ is a compact set with non-empty interior, where the Banach manifold $\mathcal{F}^{\text{reg}} \subset \mathcal{F}$ is defined in Lemma 4.3.3 (for details see [**67**]). Given an atlas $(U_\alpha, \phi_\alpha)_{\alpha\in A}$ of $\mathcal{F}^{\text{reg}}$ (cf. [**169**]) and making use of the fact that each ϕ_α is a homeomorphism mapping to some infinite-dimensional Banach space X_α, we deduce that the image of K^{reg} is a compact subset with non-empty interior in contradiction to [**100**, Exercise 14.3]. For this reason, it is not possible to assume that the space $\mathcal{F}$ is σ-compact (by contrast to the setting in [**66**]).

Next, it is no longer possible to assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is simultaneously lower semi-continuous and of compact range (see Definition 4.2.4) as introduced in [**66**, Definition 3.3]. Namely, due to lower semi-continuity of the Lagrangian, the latter assumption already implies that $\mathcal{F}$ is locally compact.

Finally, it is not possible to assume that the Lagrangian decays in entropy in the sense of [**66**, Definition 5.1] (see Definition 4.2.8 or Definition 3.5.1); indeed, this assumption requires a Heine-Borel metric on $\mathcal{F}$, which clearly does not exist in non-locally compact spaces (otherwise each $x \in \mathcal{F}$ is contained in a corresponding ball with compact closure).

In view of these difficulties in the non-locally compact setting, let us begin by generalizing the assumption that $\mathcal{L}$ is of compact range in the following way.

DEFINITION 4.3.7. *Let $(\mathcal{F}, d)$ be a metric space. The Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is said to be* ***of bounded range*** *if every bounded set $B \subset \mathcal{F}$ is contained in a bounded*

[2]For clarity, we recall that a topological space X is said to be *Baire* if the intersection of each countable family of dense open sets in X is dense (see e.g. [**161**, Definition 25.1]).

neighborhood $B' \subset \mathcal{F}$ such that

$$\mathcal{L}(x,y) = 0 \qquad \text{for all } x \in B \text{ and } y \notin B' .$$

On proper metric spaces (that is, on spaces with the Heine-Borel property), this definition clearly implies that the Lagrangian $\mathcal{L}$ is of compact range (see Definition 4.2.4 or Definition 3.3.3) as defined in [**66**]. For this reason, Definition 4.3.7 provides a good starting point for dealing with causal variational principles on non-locally compact spaces. As we shall see below, the assumption that the Lagrangian is of bounded range can be weakened (see §4.6.1).

4.3.3. Finite-Dimensional Approximation. In the infinite-dimensional setting (see Definition 4.3.4), the space X is assumed to be an infinite-dimensional separable Banach space. Hence we may apply [**5**, Lemma 7.1] to deduce that there exists a sequence of finite-dimensional subspaces $(X_n)_{n\in\mathbb{N}}$ in X with $X_n \subset X_{n+1}$ for all $n \in \mathbb{N}$ such that $\bigcup_{n\in\mathbb{N}} X_n$ is dense in X (also see [**21**]). This allows us to introduce the topological spaces

$$\mathcal{F}^{(n)} := \mathcal{F} \cap X_n \qquad \text{for every } n \in \mathbb{N} .$$

Since finite-dimensional topological vector spaces are locally compact (see e.g. [**101**, §15.7]), we conclude that each $X_n \subset X$ is locally compact for all $n \in \mathbb{N}$. For ease in notation, we shall denote the restriction of the Lagrangian to $\mathcal{F}^{(n)} \times \mathcal{F}^{(n)}$ by $\mathcal{L}^{(n)}$. Thus for every $n \in \mathbb{N}$, we are given a second-countable, locally compact Hausdorff space $\mathcal{F}^{(n)} \subset \mathcal{F}$ together with a symmetric, lower semi-continuous Lagrangian

$$\mathcal{L}^{(n)} : \mathcal{F}^{(n)} \times \mathcal{F}^{(n)} \to \mathbb{R}^+_0 ,$$

which is strictly positive on the diagonal. Henceforth for every $n \in \mathbb{N}$ we are exactly in the σ-locally compact setting as worked out in [**66**] (see Chapter 3).

In the following, we *additionally assume that* $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}^+_0$ is continuous and of bounded range (see Definition 4.3.7). We again consider the above exhaustion $(X_n)_{n\in\mathbb{N}}$ of X by finite-dimensional subsets with $X_n \subset X_{n+1}$ for all $n \in \mathbb{N}$. Let us point out that each $(X_n, \|.\|)$ is a finite-dimensional normed vector space, and all norms on X_n are equivalent. Due to the Heine-Borel theorem [**5**] or [**28**, Chapter I, Theorem 4], each closed ball $\overline{B_r(x)} \subset X_n$ is compact for all $r > 0$ and $x \in X_n$. As a consequence, each bounded set $A \subset \mathcal{F}^{(n)}$ is contained in some compact ball $B := \overline{B_r(x)} \subset \mathcal{F}^{(n)}$. Definition 4.3.7 yields the existence of a compact set $B' \subset \mathcal{F}^{(n)}$ such that $\mathcal{L}^{(n)}(x,y) = 0$ for all $x \in B$ and $y \notin B'$. These considerations show that, whenever $\mathcal{L}$ is continuous and of bounded range, for every $n \in \mathbb{N}$ the restricted Lagrangian $\mathcal{L}^{(n)}$ is continuous and of compact range (see [**66**, Definition 3.3] or Definition 4.2.4). Therefore, by virtue of Theorem 4.2.6 (also see [**66**, Theorem 4.2]), for each $n \in \mathbb{N}$ there is a regular Borel measure ρ_n on $\mathcal{F}^{(n)}$ such that the following EL equations hold,

$$\ell_n|_{\text{supp}\,\rho_n} \equiv \inf_{x\in\mathcal{F}^{(n)}} \ell_n(x) = 0 ,$$

where $\ell_n \in C(\mathcal{F}) = C(\mathcal{F}, \mathbb{R})$ is defined by

$$\ell_n(x) := \int_{\mathcal{F}^{(n)}} \mathcal{L}^{(n)}(x,y)\, d\rho_n(y) - 1 .$$

According to Theorem 4.2.7 (cf. [**66**, Theorem 4.10]), each Borel measure $\rho_n \in \mathfrak{B}_{\mathcal{F}^{(n)}}$ is a minimizer of the corresponding causal variational principle

$$\text{minimize} \qquad \mathcal{S}^{(n)} := \int_{\mathcal{F}^{(n)}} \int_{\mathcal{F}^{(n)}} \mathcal{L}^{(n)}\, d\rho(x)\, d\rho(y)$$

under variations of compact support [**66**] in the class of regular Borel measures on $\mathcal{F}^{(n)}$ with respect to the constraints (4.2.3).

We extend the measures ρ_n by zero on the whole topological space $\mathcal{F}$,

$$\rho^{[n]}(A) := \rho_n(A \cap \mathcal{F}^{(n)}) \qquad \text{for all } A \in \mathcal{B}(\mathcal{F})\,. \tag{4.3.3}$$

Thus

$$\ell^{[n]}|_{\text{supp}\,\rho^{[n]}} \equiv \inf_{x \in \mathcal{F}^{(n)}} \ell^{[n]}(x) = 0\,, \tag{4.3.4}$$

where the function $\ell^{[n]} \in C(\mathcal{F})$ is defined by

$$\ell^{[n]}(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{[n]}(y) - 1\,. \tag{4.3.5}$$

This gives rise to a sequence of regular Borel measures $(\rho^{[n]})_{n\in\mathbb{N}}$ on $\mathcal{F}$. In particular, whenever condition (iv) is satisfied for $\rho^{[n]}$ (see (4.2.2)), that is

$$\sup_{x\in\mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{[n]} < \infty \qquad \text{for all } n \in \mathbb{N}\,, \tag{4.3.6}$$

each measure $\rho^{[n]}$ is a minimizer on $\mathcal{F}^{(n)}$ under variations of finite volume [**66**] (see Definition 4.2.1 and Definition 4.2.3). In virtue of Theorem 4.2.10, the same holds true if the Lagrangian $\mathcal{L}^{(n)}$ decays in entropy for any $n \in \mathbb{N}$, provided that condition (4.3.6) is satisfied.

4.4. Construction of a Global Measure

In the following, let X be an infinite-dimensional, separable complex Banach space, and let $\mathcal{F} \subset X$ be a non-locally compact Polish space endowed with a corresponding metric d such that $(\mathcal{F}, d)$ is a separable, complete metric space. By $\mathcal{O}(\mathcal{F})$ and $\mathcal{P}(\mathcal{F})$ we denote the collection of open subsets of $\mathcal{F}$ and the power set of $\mathcal{F}$, respectively. Moreover, the collection of all compact subsets of $\mathcal{F}$ is represented by $\mathfrak{K}(\mathcal{F})$.

The goal of this section is to construct a global measure ρ based on the sequence of regular Borel measures $(\rho^{[n]})_{n\in\mathbb{N}}$ as obtained in §4.3.3. To this end, in a first step we construct a countable set $\mathcal{D} \subset \mathfrak{K}(\mathcal{F})$ consisting of compact subsets of $\mathcal{F}$ (§4.4.1). In a second step, we make use of the set $\mathcal{D}$ in order to obtain a measure ρ on $\mathcal{F}$ (§4.4.2) by a suitable construction process. In §4.4.3 we finally prove that, restricted to suitable relatively compact subsets of $\mathcal{F}$, the measure ρ is the weak limit of a subsequence of $(\rho^{[n]})_{n\in\mathbb{N}}$.

4.4.1. Construction of a Countable Collection of Compact Sets. To begin with, by separability of $\mathcal{F}$ there exists a countable dense subset $E := \{x_j : j \in \mathbb{N}\}$ such that, for every $n \in \mathbb{N}$ the set $E^{(n)} := E \cap \mathcal{F}^{(n)}$ is dense in $\mathcal{F}^{(n)}$.[3] We denote its elements by $x_j^{(n)} \in E^{(n)}$ with $j, n \in \mathbb{N}$. Moreover, since $\mathcal{F}^{(n)}$ is locally compact, for all $j, k, n \in \mathbb{N}$ there is a compact neighborhood $V_{j,k}^{(n)} \subset \mathcal{F}^{(n)}$ of $x_j^{(n)} \in E^{(n)}$ such that $V_{j,k}^{(n)} \subset B_{1/k}(x_j^{(n)})$

[3]Since $\mathcal{F}$ is separable, there exists a countable set $E^{(0)} \subset \mathcal{F}$ being dense in $\mathcal{F}$. Similarly, for each $i \in \mathbb{N}$ there are countable sets $E^{(i)}$ which are dense in $\mathcal{F}^{(i)}$. As a consequence, the set $E := \bigcup_{i=0}^{\infty} E^{(i)}$ has the desired properties.

and each $V_{j,k}^{(n)}$ being the closure of its interior (in the topology of $\mathcal{F}^{(n)}$, where the interior of a set V shall be denoted by V°).[4] This gives rise to the set

$$\mathcal{V}^{(1)} := \left\{ V_{j,k}^{(n)} : j,k,n \in \mathbb{N} \right\} .$$

Denoting the union of $\mathcal{V}^{(1)}$ and the empty set $\varnothing$ by $\tilde{\mathcal{D}}^{(1)}$, and making use of the fact that a countable union of countable sets is countable (see e.g. [**75**, Section 2]), we conclude that $\tilde{\mathcal{D}}^{(1)}$ is countable. Therefore, applying Cantor's diagonal argument and proceeding iteratively, we conclude that

$$\tilde{\mathcal{D}}^{(i)} := \left\{ D \cup \tilde{D} : D, \tilde{D} \in \tilde{\mathcal{D}}^{(i-1)} \right\}$$

is countable for every $i \in \mathbb{N}$ with $i \geq 2$. As a consequence, the set

$$\mathcal{D} := \bigcup_{i=1}^{\infty} \tilde{\mathcal{D}}^{(i)}$$

is countable; we denote its members by $(D_m)_{m \in \mathbb{N}}$. In particular, each $D \in \mathcal{D}$ is a compact subset of $\mathcal{F}$. Moreover, for every $n \in \mathbb{N}$ we introduce $\mathcal{D}^{(n)} \subset \mathcal{P}(\mathcal{F}^{(n)})$ by

$$\mathcal{D}^{(n)} := \left\{ D \in \mathcal{D} : D \subset \mathcal{F}^{(n)} \right\} . \tag{4.4.1}$$

4.4.2. Construction of a Regular Global Measure. In order to construct a global measure on $\mathcal{F}$, we proceed similarly to [**66**] by selecting suitable subsequences of the sequence $(\rho^{[n]})_{n \in \mathbb{N}}$ restricted to compact subsets $D \in \mathcal{D}$. This allows us to construct a regular measure ρ on the whole space $\mathcal{F}$ (see Theorem 4.4.3 below). In Section 4.5 we will show that, under suitable assumptions, the measure ρ is indeed a minimizer of the causal variational principle. In analogy to [**66**, Lemma 4.1], let us first state the following result.

LEMMA 4.4.1. *Assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is lower semi-continuous and strictly positive on the diagonal* (4.3.1). *Furthermore, let $(\rho^{[n]})_{n \in \mathbb{N}}$ be a sequence of measures $\rho^{[n]} : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ such that, for every $x \in \operatorname{supp} \rho^{[n]}$,*

$$\int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{[n]}(y) = 1 \qquad \text{for all } n \in \mathbb{N} .$$

Then for every compact subset $K \subset \mathcal{F}$ there is a constant $C_K > 0$ such that

$$\rho^{[n]}(K) \leq C_K \qquad \text{for all } n \in \mathbb{N} .$$

PROOF. This statement is proven exactly as [**66**, Lemma 4.1]. □

Next, we apply Lemma 4.4.1 to the compact sets $D \in \mathcal{D}$. More precisely, restricting the sequence $(\rho^{[n]})_{n \in \mathbb{N}}$ as obtained in (4.3.3) (cf. §4.3.3) to the compact set $D_1 \in \mathcal{D}$, the resulting sequence $(\rho^{[n]}|_{D_1})_{n \in \mathbb{N}}$ is bounded (due to Lemma 4.4.1) as well as uniformly tight (for the definition see [**16**, Definition 8.6.1]). Since compact subsets of Polish spaces are again Polish, Prohorov's theorem (see for instance [**16**, Theorem 8.6.2] or [**40**, Satz VIII.4.23]) implies that a subsequence of $(\rho^{[n]}|_{D_1})_{n \in \mathbb{N}}$ converges weakly on D_1. Denoting the corresponding subsequence by $(\rho^{[1,n_k]})_{k \in \mathbb{N}}$ and considering its restriction to $D_2 \in \mathcal{D}$, the same arguments as before yield the existence of a weakly convergent

[4] For simplicity, one may consider $V_{j,k}^{(n)} = \overline{B_{1/(2k)}(x_j^{(n)}) \cap \mathcal{F}^{(n)}}$ for all $j,k,n \in \mathbb{N}$.

subsequence $(\rho^{[2,n_k]})_{k\in\mathbb{N}}$ on D_2. Proceeding iteratively, we denote the resulting diagonal sequence by

$$\rho^{(k)} := \rho^{[k,n_k]} \qquad \text{for all } k \in \mathbb{N}\,. \tag{4.4.2}$$

Thus by construction, for every $m \in \mathbb{N}$ the sequence $(\rho^{(k)}|_{D_m})_{k\in\mathbb{N}}$ converges weakly to some measure $\rho_{D_m} : \mathcal{B}(D_m) \to [0,\infty)$,

$$\rho^{(k)}|_{D_m} \rightharpoonup \rho_{D_m}\,. \tag{4.4.3}$$

In particular,

$$\lim_{k\to\infty} \rho^{(k)}(D_m) = \rho_{K_m}(D_m) \qquad \text{for all } m \in \mathbb{N}\,.$$

We point out that each measure $\rho^{(k)}$ is a minimizer on $\mathcal{F}^{(n_k)}$. For this reason, we restrict attention to the finite-dimensional exhaustion $(\mathcal{F}^{(k)})_{k\in\mathbb{N}}$, where for notational simplicity by $\mathcal{F}^{(k)}$ we denote the sets $\mathcal{F}^{(n_k)}$ for all $k \in \mathbb{N}$. Note that the sequence constructed in (4.4.2) above in general does *not* converge weakly on *arbitrary* compact subsets, but only restricted to compact sets $D \in \mathcal{D}$ (cf. (4.4.3)). In [**66**], this problem was resolved by deriving vague convergence of the sequence $(\rho^{(n)})_{n\in\mathbb{N}}$ to some global measure ρ. In order to obtain a similar situation, let us state the following result.

PROPOSITION 4.4.2. *The set function $\varphi : \mathcal{D} \to [0,\infty)$ defined by*

$$\varphi(D) := \lim_{k\to\infty} \rho^{(k)}(D) < \infty \qquad \textit{for any } D \in \mathcal{D} \tag{4.4.4}$$

has the following properties:

(1) *$\varphi(D_1) \le \varphi(D_2)$ for all $D_1, D_2 \in \mathcal{D}$ with $D_1 \subset D_2$,*
(2) *$\varphi(D_1 \cup D_2) \le \varphi(D_1) + \varphi(D_2)$ for all $D_1, D_2 \in \mathcal{D}$, and*
(3) *$\varphi(D_1 \cup D_2) = \varphi(D_1) + \varphi(D_2)$ for all $D_1, D_2 \in \mathcal{D}$ with $D_1 \cap D_2 = \varnothing$.*

PROOF. Given $D_1, D_2 \in \mathcal{D}$ with $D_1 \subset D_2$, property (1) follows from

$$\varphi(D_1) = \lim_{k\to\infty} \int_{\mathcal{F}} d\rho^{(k)}|_{D_1} \le \lim_{k\to\infty} \int_{\mathcal{F}} d\rho^{(k)}|_{D_2} = \varphi(D_2)\,.$$

Next, for all $D_1, D_2 \in \mathcal{D}$, construction of $\mathcal{D}$ yields $D_1 \cup D_2 \in \mathcal{D}$. Thus property (2) is a consequence of

$$\begin{aligned}\varphi(D_1 \cup D_2) &= \lim_{k\to\infty} \int_{\mathcal{F}} d\rho^{(k)}|_{D_1\cup D_2} \le \lim_{k\to\infty} \int_{\mathcal{F}} d\rho^{(k)}|_{D_1} + \lim_{k\to\infty} \int_{\mathcal{F}} d\rho^{(k)}|_{D_2} \\ &= \varphi(D_1) + \varphi(D_2)\,.\end{aligned}$$

Similarly, for all $D_1, D_2 \in \mathcal{D}$ with $D_1 \cap D_2 = \varnothing$ we obtain

$$\varphi(D_1 \cup D_2) = \lim_{k\to\infty} \rho^{(k)}(D_1 \cup D_2) = \lim_{k\to\infty} \rho^{(k)}(D_1) + \lim_{k\to\infty} \rho^{(k)}(D_2) = \varphi(D_1) + \varphi(D_2)\,,$$

which proves property (3). □

In order to construct a global measure ρ on $\mathcal{F}$, we proceed in analogy to the proof of [**40**, Satz VIII.4.22]. We point out that, since the underlying topological space $\mathcal{F}$ is non-locally compact, we cannot employ the Riesz representation theorem as in [**66**], and Riesz representation theorems on more general Hausdorff spaces as presented in [**100**, Section 16] do not seem applicable at this stage. Nevertheless, we obtain the following result.

THEOREM 4.4.3. *Introducing the set function $\varphi : \mathcal{D} \to [0, \infty)$ by (4.4.4) and defining the set functions $\mu : \mathcal{O}(\mathcal{F}) \to [0, \infty]$ and $\eta : \mathcal{P}(\mathcal{F}) \to [0, +\infty]$ by*

$$\begin{aligned} \mu(U) &:= \sup \{\varphi(D) : D \subset U,\ D \in \mathcal{D}\ \} && \text{for all } U \subset \mathcal{F} \text{ open}, \\ \eta(A) &:= \inf \{\mu(U) : A \subset U,\ U \subset \mathcal{F} \text{ open}\} && \text{for any } A \in \mathcal{P}(\mathcal{F}), \end{aligned} \tag{4.4.5}$$

then the restriction

$$\rho := \eta|_{\mathcal{B}(\mathcal{F})} \tag{4.4.6}$$

defines a (possibly trivial) measure on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$. In particular,

$$\rho(D) = \varphi(D) = \lim_{k \to \infty} \rho^{(k)}(D) \qquad \text{for any } D \in \mathcal{D}. \tag{4.4.7}$$

PROOF. Let us first point out that, by construction, the set function $\varphi : \mathcal{D} \to [0, \infty)$ defined by (4.4.4) has the following properties:

(1) $\varphi(D_1) \leq \varphi(D_2)$ for all $D_1, D_2 \in \mathcal{D}$ with $D_1 \subset D_2$,
(2) $\varphi(D_1 \cup D_2) \leq \varphi(D_1) + \varphi(D_2)$ for all $D_1, D_2 \in \mathcal{D}$, and
(3) $\varphi(D_1 \cup D_2) = \varphi(D_1) + \varphi(D_2)$ for all $D_1, D_2 \in \mathcal{D}$ with $D_1 \cap D_2 = \varnothing$.

Indeed, properties (1)–(3) are a consequence of Proposition 4.4.2. Moreover, $\varphi(\varnothing) = 0$ (since $\varnothing \in \mathcal{D}$).

Next, similarly to the proof of [**40**, Satz VIII.4.22], our goal is to show that η is an outer measure,[5] and that every Borel set $B \in \mathcal{B}(\mathcal{F})$ is η-measurable. This shall be done in the following by proving that

$$\eta \text{ is an outer measure, and each closed set } A \subset \mathcal{F} \text{ is } \eta\text{-measurable}. \tag{4.4.8}$$

Denoting the σ-algebra of η-measurable sets by $\mathfrak{A}_\eta$, statement (4.4.8) implies that the Borel σ-algebra $\mathcal{B}(\mathcal{F})$ is contained in $\mathfrak{A}_\eta$, i.e. $\mathcal{B}(\mathcal{F}) \subset \mathfrak{A}_\eta$. Therefore, in view of Carathéodory's theorem (see e.g. [**40**, Satz II.4.4]), the restriction

$$\rho := \eta|_{\mathcal{B}(\mathcal{F})}$$

defines a measure on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$. Thus it suffices to prove (4.4.8), which shall be done in several steps in the remainder of the proof.

(a) *Let $A \subset U \subset \mathcal{F}$ with A closed and U open. Whenever $A \subset D$ for some $D \in \mathcal{D}$, then there exists $E \in \mathcal{D}$ with $A \subset E \subset U$.*

Proof: Since each $D \in \mathcal{D}$ is compact, the closed set $A \subset D$ is compact as well. Moreover, we have $D \subset \mathcal{F}^{(n)}$ for sufficiently large $n \in \mathbb{N}$. Since $\mathcal{F}^{(n)}$ is locally compact, for every $x \in A$ there exists $V_x \in \mathcal{D}$ such that $x \in V_x^\circ \subset V_x \subset U$. Since A is compact, the set $E := \bigcup_{j=1}^N V_{x_j} \in \mathcal{D}$ for some integer $N = N(A)$ has the desired property.

(b) *Whenever $U, V \subset \mathcal{F}$ open, $\mu(U \cup V) \leq \mu(U) + \mu(V)$.*

---

[5] Given a set X, a set function $\eta : \mathcal{P}(X) \to \overline{\mathbb{R}} := [-\infty, +\infty]$ is said to be an *outer measure* if it has the following properties (see e.g. [**40**, Definition II.4.1]):

(i) $\eta(\varnothing) = 0$.
(ii) For all $A \subset B \subset X$ holds $\eta(A) \leq \eta(B)$ (monotonicity).
(iii) For every sequence $(A_n)_{n \in \mathbb{N}}$ of subsets of X holds $\eta\big(\bigcup_{n=1}^\infty A_n\big) \leq \sum_{n=1}^\infty \eta(A_n)$ (σ-subadditivity).

Proof: Without loss of generality, let $U \neq \mathcal{F} \neq V$ and $\mu(U), \mu(V) < \infty$ (otherwise the inequality is true). For this reason, let $U, V \subset \mathcal{F}$ be open sets with $U^c \neq \varnothing \neq V^c$ and $D \subset U \cup V$ for $D \in \mathcal{D}$. We then consider the closed sets

$$A := \{x \in D : d(x, U^c) \geq d(x, V^c)\} \subset D\,,$$
$$B := \{x \in D : d(x, U^c) \leq d(x, V^c)\} \subset D\,.$$

Obviously, $A \subset U$ and $B \subset V$. Assuming conversely that $x \in A \setminus U$, we conclude that $x \in V$, and therefore $d(x, U^c) = 0 < d(x, V^c)$ because V^c is closed, giving rise to the contradiction that $x \notin A$. Similarly, we conclude that $B \subset V$. Since $A \subset D$, by virtue of (a) there exists $E \in \mathcal{D}$ with $A \subset E \subset U$. Similarly, there exists $F \in \mathcal{D}$ in such a way that $B \subset F \subset V$, and $D = A \cup B \subset E \cup F$. Hence (1) and (2) yield

$$\varphi(D) \leq \varphi(E \cup F) \leq \varphi(E) + \varphi(F) \leq \mu(U) + \mu(V)\,.$$

Taking the supremum over all $D \in \mathcal{D}$ with $D \subset U \cup V$ gives (b).

(c) *For all $n \in \mathbb{N}$ and $U_n \subset \mathcal{F}$ open, $\mu\big(\bigcup_{n=1}^{\infty} U_n\big) \leq \sum_{n=1}^{\infty} \mu(U_n)$.*

Proof: Let $D \in \mathcal{D}$ with $D \subset \bigcup_{n=1}^{\infty} U_n$. Then by compactness of D there exists $p \in \mathbb{N}$ such that $D \subset \bigcup_{n=1}^{p} U_n$. Applying (b) inductively, we conclude that

$$\varphi(D) \leq \mu\left(\bigcup_{n=1}^{p} U_n\right) \leq \sum_{n=1}^{p} \mu(U_n) \leq \sum_{n=1}^{\infty} \mu(U_n)\,.$$

Since $D \in \mathcal{D}$ with $D \subset \bigcup_{n=1}^{\infty} U_n$ is arbitrary, we obtain (c).

(d) *η is an outer measure.*

Proof: As seen before, $\varphi(\varnothing) = 0$, and monotonicity of η is a consequence of (1)–(3) and (4.4.5). In order to prove σ-subadditivity, let $\varepsilon > 0$ and $M_n \subset \mathcal{F}$ with $\eta(M_n) < \infty$ for all $n \in \mathbb{N}$. In view of (4.4.5), for every $n \in \mathbb{N}$ there exists an open set $U_n \supset M_n$ with $\mu(U_n) \leq \eta(M_n) + 2^{-n}\,\varepsilon$. Making use of (4.4.5) and applying (c) yields

$$\eta\left(\bigcup_{n=1}^{\infty} M_n\right) \leq \mu\left(\bigcup_{n=1}^{\infty} U_n\right) \leq \sum_{n=1}^{\infty} \mu(U_n) \leq \sum_{n=1}^{\infty} \eta(M_n) + \varepsilon\,.$$

Since $\varepsilon > 0$ is arbitrary, we obtain (d).

(e) *Each closed set $A \subset \mathcal{F}$ is η-measurable.*

Proof: By definition of measurability (cf. [**40**, Definition II.4.2]), we need to show that, for all $Q \subset \mathcal{F}$,

$$\eta(Q) \geq \eta(Q \cap A) + \eta(Q \cap A^c)\,. \tag{4.4.9}$$

Without loss of generality we may assume that $\eta(Q) < \infty$. We first prove (4.4.9) in the case that $Q = U \subset \mathcal{F}$ is *open*. To this end, let $\varepsilon > 0$ be arbitrary. Given $A \subset \mathcal{F}$ closed, the set $U \cap A^c$ is open and $\mu(U \cap A^c) = \eta(U \cap A^c) < \infty$. In view of (4.4.5), there exists $D \in \mathcal{D}$ with $\varphi(D) \geq \mu(U \cap A^c) - \varepsilon$. Next, since $U \cap D^c$ is open, we may choose $E \in \mathcal{D}$ with $E \subset U \cap D^c$ and $\varphi(E) \geq \mu(U \cap D^c) - \varepsilon$. Since D, E are disjoint and $D \cup E \subset U$, from (1), (3), (4.4.5) and the fact that $U \cap D^c \supset U \cap A$ we infer that

$$\begin{aligned}\mu(U) &\geq \varphi(D \cup E) = \varphi(D) + \varphi(E)\\ &\geq \mu(U \cap A^c) + \mu(U \cap D^c) - 2\varepsilon\\ &\geq \eta(U \cap A) + \mu(U \cap A^c) - 2\varepsilon\,.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we obtain (4.4.9) for $Q = U$ open.

Given arbitrary $Q \subset \mathcal{F}$ with $\eta(Q) < \infty$, for $\varepsilon > 0$ arbitrary we choose $U \supset Q$ open with $\eta(Q) \geq \eta(U) - \varepsilon$ according to (4.4.5). Then the latter inequality yields

$$\begin{aligned} \eta(Q) &\geq \eta(U) - \varepsilon \geq \eta(U \cap A) + \eta(U \cap A^c) - \varepsilon \\ &\geq \eta(Q \cap A) + \eta(Q \cap A^c) - \varepsilon \,, \end{aligned}$$

proving (4.4.9).

As a consequence, the set function η is an outer measure according to (d), and each closed set $A \subset \mathcal{F}$ is η-measurable in view of (e). This yields (4.4.8), which completes the proof. □

The next result shows that the measure ρ given by (4.4.6) is regular [**40**].

LEMMA 4.4.4. *Let $\rho : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ be the measure defined by* (4.4.6)*. Then every open subset of $\mathcal{F}$ is inner regular. Moreover, the measure ρ is regular.*

PROOF. Let us first prove that every open subset of $\mathcal{F}$ is inner regular in view of (4.4.5). Namely, considering arbitrary $U \in \mathcal{O}(\mathcal{F})$ and $K \in \mathfrak{K}(U)$, according to (4.4.5) and (4.4.7) we obtain

$$\begin{aligned} \rho(K) \leq \rho(U) &= \sup \{\varphi(D) : D \in \mathcal{D},\, D \subset U\} = \sup \{\rho(D) : D \in \mathcal{D},\, D \subset U\} \\ &\leq \sup \{\rho(K) : K \subset U \text{ compact}\} \,. \end{aligned}$$

Taking the supremum on the left hand side yields

$$\rho(U) = \sup \{\rho(K) : K \subset U \text{ compact}\} \qquad \text{for all } U \in \mathcal{O}(\mathcal{F}) \,.$$

From this we conclude that every open set $U \in \mathcal{O}(\mathcal{F})$ is inner regular (in the sense of [**40**, Definition VIII.1.1]).

In view of (4.4.5), we are given $\rho(U) = \mu(U)$ for any $U \in \mathcal{O}(\mathcal{F})$, and the above considerations show that every open set is inner regular. From this we conclude that the measure ρ is regular in the sense of [**40**, Definition VIII.1.1]. □

As a matter of fact, in general it seems possible the regular measure ρ obtained in Theorem 4.4.3 to be zero. Nevertheless, the following remark gives a sufficient condition for the measure ρ defined by (4.4.6) to be non-zero.

REMARK 4.4.5. *Let $(\mathcal{F}^{(n)})_{n \in \mathbb{N}}$ be a finite-dimensional approximation of $\mathcal{F}$ (for details see §4.3.3). By construction of $\rho^{(n)}$ we are given $\operatorname{supp} \rho^{(n)} \subset \mathcal{F}^{(n)}$ for every $n \in \mathbb{N}$. Assuming that the Lagrangian is bounded and of bounded range, for every $x \in \mathcal{F}$ and $\delta > 0$ there exists $B_x \subset \mathcal{F}$ bounded and closed such that $\mathcal{L}(\tilde{x}, y) = 0$ for all $\tilde{x} \in B_\delta(x)$ and all $y \notin B_x$. Furthermore, in view of boundedness of the Lagrangian we introduce the upper bound $\mathfrak{C} < \infty$ by*

$$\mathfrak{C} := \sup_{x,y \in \mathcal{F}} \mathcal{L}(x, y) > 0 \,.$$

Thus for any $n \in \mathbb{N}$ we deduce that $\mathcal{L}^{(n)}(\tilde{x}, y) = 0$ for all $\tilde{x} \in B_\delta^{(n)}(x) := B_\delta(x) \cap \mathcal{F}^{(n)}$ and all $y \notin B_x^{(n)} := B_x \cap \mathcal{F}^{(n)}$. Hence the EL equations (4.3.4) *and* (4.3.5) *imply that*

$$1 \leq \int_{\mathcal{F}^{(n)}} \mathcal{L}^{(n)}(\tilde{x}, y)\, d\rho^{(n)} = \int_{B_x^{(n)}} \mathcal{L}^{(n)}(\tilde{x}, y)\, d\rho^{(n)} \leq \sup_{y \in B_x^{(n)}} \mathcal{L}^{(n)}(\tilde{x}, y)\, \rho^{(n)}(B_x^{(n)})$$

for every $n \in \mathbb{N}$. Thus positivity (4.3.1) *yields*

$$\rho^{(n)}(B_x^{(n)}) \geq \mathfrak{C}^{-1} > 0 \qquad \text{for all } n \in \mathbb{N} \,.$$

For each $n \in \mathbb{N}$ and arbitrary $\varepsilon > 0$, by regularity of $\rho^{(n)}$ there exists $D_n \in \mathcal{D}$ such that $\rho^{(n)}(D_n) > \mathcal{C}^{-1} - \varepsilon$. Moreover, $\hat{D}_N := \bigcup_{n=1}^{N} D_n \in \mathcal{D}$ for every $N \in \mathbb{N}$. Whenever there exists $N \in \mathbb{N}$ such that $\rho^{(n)}(\hat{D}_N) \geq c$ for almost all $n \in \mathbb{N}$ and some $c > 0$, then the measure ρ defined by (4.4.6) *is non-zero. If this holds true for an infinite number of disjoints sets $(\hat{D}_{N_i})_{i\in\mathbb{N}}$, the measure ρ possibly has infinite total volume.*

Next, in agreement with [**100**, Theorem 16.7] and the remark thereafter, it is not clear whether ρ as given by (4.4.6) is a locally finite measure. Nevertheless, the following results provide sufficient conditions for ρ as obtained in (4.4.6) to be locally finite.

LEMMA 4.4.6. *Let $\rho : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ be defined by* (4.4.6)*. Assuming that $\rho(K) < \infty$ for all $K \in \mathfrak{K}(\mathcal{F})$, then the measure ρ is locally finite and thus a Borel measure in the sense of* [**76**]*. In this case, ρ is regular and moderate.*

PROOF. We point out that the space $\mathcal{F}$ is first-countable. Thus under the assumption that $\rho(K) < \infty$ for all $K \in \mathfrak{K}(\mathcal{F})$, the statement that ρ is locally finite is a consequence of Lemma 4.4.4 and [**40**, Folgerungen VIII.1.2 (d)]. The last statement follows from Ulam's theorem [**40**, Theorem VIII.1.16]. □

REMARK 4.4.7. *We point out that, if $\mathcal{F}$ in Definition 4.3.4* were *locally compact, then the measure ρ constructed in the proof of Theorem 4.4.3* would be *locally finite, i.e. a Borel measure in the sense of* [**76**]*. Namely, whenever $x \in \mathcal{F}$, there exists a compact neighborhood V of x. Hence Lemma 4.4.1 implies that $\rho^{(n)}(V) \leq C_V$ for all $n \in \mathbb{N}$ and some $C_V > 0$. Choosing $U_x \subset V$ open with $x \in U$, we conclude that*

$$\rho(U_x) = \sup\{\varphi(D) : D \subset U_x,\, D \in \mathcal{D}\} \leq \sup\left\{\lim_{n\to\infty} \rho^{(n)}(D) : D \subset V,\, D \in \mathcal{D}\right\} \leq C_V$$

as desired.

In the remainder of this subsection, we discuss properties (iii) and (iv) in §4.2.2. Neither condition (iii) nor condition (iv) do hold in general, but the following result establishes a connection between condition (iv) in §4.2.2 and locally finite measures.

LEMMA 4.4.8. *Assume that $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is lower semi-continuous, symmetric and strictly positive on the diagonal* (4.3.1)*, and let ρ be a measure on $\mathcal{B}(\mathcal{F})$. Under the assumption that condition* (iv) *in §4.2.2 holds (see* (4.2.2)*), i.e.*

$$\sup_{x\in\mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x, y)\, d\rho(y) < \infty\,,$$

the measure ρ is locally finite (i.e. condition (iii) *in §4.2.2 is satisfied).*

PROOF. Assume conversely that there exists $x \in \mathcal{F}$ such that $\rho(U) = \infty$ for any open neighborhood U of x. Then $\mathcal{L}(x, x) > 0$ due to strict positivity on the diagonal (4.3.1). Moreover, by lower semi-continuity of the Lagrangian there is an open neighborhood U_x of x such that $\mathcal{L}(x, y) > \mathcal{L}(x, x)/2 > 0$ for all $y \in U_x$. Consequently,

$$\int_{\mathcal{F}} \mathcal{L}(x, y)\, d\rho(y) \geq \int_{U_x} \mathcal{L}(x, y)\, d\rho(y) > \mathcal{L}(x, x)/2\, \rho(U_x) = \infty$$

in contradiction to condition (iv) in §4.2.2. □

4.4.3. Convergence on Relatively Compact Subsets. In Section 4.5 below our goal is to show that, under suitable assumptions, the measure ρ as given by (4.4.6) is a minimizer under variations of finite volume (see Definition 4.3.5). To this end, we provide some useful tools which shall be worked out in the remainder of this section. For clarity, we point out that for every $D \in \mathcal{D}$ there is some $n' \in \mathbb{N}$ such that $D^\circ \neq \varnothing$ in the relative topology of $\mathcal{F}^{(n)}$ for all $n \leq n'$ and $D^\circ = \varnothing$ in the relative topology of $\mathcal{F}^{(n)}$ for all $n > n'$. Considering the restriction $\rho|_{D^\circ}$ shall always be understood in the sense of a restriction to the interior of D in the relative topology of $\mathcal{F}^{(n')}$. In order to derive weak convergence on suitable relatively compact sets (see Lemma 4.4.11 below), we require some properties of the measures ρ_D with $D \in \mathcal{D}$ as obtained in (4.4.3).

Given $D, E \in \mathcal{D}$ with $D \subset E$ and denoting the interior of D by D°, we claim that

$$\int_D f \, d\rho_D = \int_E f \, d\rho_E \qquad \text{for all } f \in C_c(D^\circ)\,. \tag{4.4.10}$$

Namely, in view of $C_c(D^\circ) \subset C_b(D) \cap C_b(E)$, weak convergence (4.4.3) yields

$$\int_D f \, d\rho_D = \lim_{k\to\infty} \int_D f \, d\rho^{(k)}|_D \overset{(\star)}{=} \lim_{k\to\infty} \int_E f \, d\rho^{(k)}|_E = \int_E f \, d\rho_E$$

for all $f \in C_c(D^\circ)$, where in $(\star)$ we made use of the fact that $\operatorname{supp} f \subset D^\circ \subset E$.

This allows us to prove the following result.

PROPOSITION 4.4.9. *Whenever $D \in \mathcal{D}$ and $\Omega \subset D^\circ$ open, then*

$$\sup \Big\{ \rho_{\tilde{D}}|_{D^\circ}(\tilde{D}) : \tilde{D} \subset \Omega,\ \tilde{D} \in \mathcal{D} \Big\} = \sup \Big\{ \rho_D|_{D^\circ}(\tilde{D}) : \tilde{D} \subset \Omega,\ \tilde{D} \in \mathcal{D} \Big\}\,. \tag{4.4.11}$$

PROOF. In order to prove (4.4.11), we need to show that for all $E, F \in \mathcal{D}$ with E, $F \subset \Omega$ there exist sets $\tilde{E}, \tilde{F} \in \mathcal{D}$ with $\tilde{E}, \tilde{F} \subset \Omega$ such that

$$\rho_E|_{D^\circ}(E) \leq \rho_D|_{D^\circ}(\tilde{E}) \qquad \text{and} \qquad \rho_D|_{D^\circ}(F) \leq \rho_{\tilde{F}}|_{D^\circ}(\tilde{F})\,.$$

Whenever $E \subset \Omega$ compact, there exists $\eta \in C_c(\Omega; [0,1])$ with $\eta|_E \equiv 1$ and $\tilde{E} \in \mathcal{D}$ with $\operatorname{supp}\eta \subset \tilde{E}^\circ \subset \Omega$, and by weak convergence (4.4.3) we are given

$$\rho_E|_{D^\circ}(E) = \int_E d\rho_E = \lim_{n\to\infty} \int_{\mathcal{F}} d\rho^{(n)}|_E \leq \lim_{n\to\infty} \int_{\mathcal{F}} \eta \, d\rho^{(n)}|_D = \int_{\mathcal{F}} \eta \, d\rho_D \leq \rho_D|_{D^\circ}(\tilde{E})\,.$$

On the other hand, whenever $F \in \mathcal{D}$ with $F \subset \Omega$, there is $\eta \in C_c(\Omega; [0,1])$ with $\eta|_F \equiv 1$ as well as $\tilde{F} \in \mathcal{D}$ with $\operatorname{supp}\eta \subset \tilde{F}^\circ$ and $\tilde{F}^\circ \subset \Omega$. Thus by (4.4.10) we obtain

$$\rho_D|_{D^\circ}(F) = \int_F d\rho_D \leq \int_D \eta \, d\rho_D = \int_{\tilde{F}} \eta \, d\rho_{\tilde{F}} \leq \rho_{\tilde{F}}|_{D^\circ}(\tilde{F})\,,$$

which completes the proof. □

LEMMA 4.4.10. *For every $D \in \mathcal{D}$, the measures $\rho|_{D^\circ}$ and $\rho_D|_{D^\circ}$ coincide. Moreover,*

$$\rho^{(n)}|_{D^\circ} \overset{v}{\to} \rho|_{D^\circ} \qquad \textit{vaguely}\,. \tag{4.4.12}$$

PROOF. According to (4.4.3) and (4.4.4), the measure $\rho_D|_{D^\circ} : \mathcal{B}(D^\circ) \to [0,\infty)$ is finite for any $D \in \mathcal{D}$. As a consequence, it is locally finite and thus Borel in the sense of [**76**]. Moreover, since open subsets of Polish spaces are Polish (see [**9**, §26]), it is regular in view of Ulam's theorem [**40**, Satz VIII.1.16] (alternatively, regularity follows by [**16**, Corollary 7.1.9] and the fact that each metrizable space is perfectly normal [**4**, Corollary 3.21]). Similar arguments yield regularity of $\rho|_{D^\circ}$ for any $D \in \mathcal{D}$. Thus for

any $D \in \mathcal{D}$, we may approximate arbitrary Borel sets $A \in \mathcal{B}(D^\circ)$ by compact sets from inside,

$$\rho_D|_{D^\circ}(A) = \sup\{\rho_D|_{D^\circ}(K) : K \subset A \text{ compact}\}\ ,$$
$$\rho|_{D^\circ}(A) = \sup\{\rho|_{D^\circ}(K) : K \subset A \text{ compact}\}\ .$$

Whenever $D \in \mathcal{D}$ and $\Omega \subset D$ open, for each $K \subset \Omega$ compact there exists $\tilde{D} \in \mathcal{D}$ such that $K \subset \tilde{D} \subset \Omega$ by construction of $\mathcal{D}$. From this we conclude that

$$\rho_D|_{D^\circ}(\Omega) = \sup\{\rho_D|_{D^\circ}(K) : K \subset \Omega \text{ compact}\} = \sup\left\{\rho_D|_{D^\circ}(\tilde{D}) : \tilde{D} \subset \Omega,\, \tilde{D} \in \mathcal{D}\right\}\ ,$$
$$\rho|_{D^\circ}(\Omega) = \sup\{\rho|_{D^\circ}(K) : K \subset \Omega \text{ compact}\} = \sup\left\{\rho|_{D^\circ}(\tilde{D}) : \tilde{D} \subset \Omega,\, \tilde{D} \in \mathcal{D}\right\}\ .$$

Moreover, for any $\tilde{D} \in \mathcal{D}$, by (4.4.3) and (4.4.7) we obtain

$$\varphi(\tilde{D}) = \lim_{k\to\infty} \rho^{(k)}(\tilde{D}) = \lim_{k\to\infty} \rho^{(k)}|_{\tilde{D}}(\tilde{D}) = \rho_{\tilde{D}}(\tilde{D})\,. \tag{4.4.13}$$

Given $D \in \mathcal{D}$ and $\Omega \subset D^\circ$ open, we conclude that $\rho|_{D^\circ}$ as well as $\rho_D|_{D^\circ}$ are regular finite Borel measures on $\mathcal{B}(D^\circ)$, implying that

$$\begin{aligned}\rho|_{D^\circ}(\Omega) &= \sup\left\{\varphi(\tilde{D}) : \tilde{D} \subset \Omega,\, \tilde{D} \in \mathcal{D}\right\} \overset{(4.4.13)}{=} \sup\left\{\rho_{\tilde{D}}(\tilde{D}) : \tilde{D} \subset \Omega,\, \tilde{D} \in \mathcal{D}\right\} \\ &= \sup\left\{\rho_{\tilde{D}}|_{D^\circ}(\tilde{D}) : \tilde{D} \subset \Omega,\, \tilde{D} \in \mathcal{D}\right\} \overset{(4.4.11)}{=} \sup\left\{\rho_D|_{D^\circ}(\tilde{D}) : \tilde{D} \subset \Omega,\, \tilde{D} \in \mathcal{D}\right\} \\ &= \rho_D|_{D^\circ}(\Omega)\,.\end{aligned}$$

As a consequence, the measures $\rho|_{D^\circ}$ and $\rho_D|_{D^\circ}$ coincide on all open sets $\Omega \subset \mathcal{D}^\circ$. Making use of [**16**, Lemma 7.1.2], we conclude that $\rho|_{D^\circ}$ and $\rho_D|_{D^\circ}$ already coincide on all Borel sets, i.e.

$$\rho|_{D^\circ} = \rho_D|_{D^\circ} \qquad \text{for all } D \in \mathcal{D}\,. \tag{4.4.14}$$

Given $f \in C_c(D^\circ)$, in view of (4.4.3) and (4.4.14) we thus obtain

$$\lim_{n\to\infty} \int_{\mathcal{F}} f\, d\rho^{(n)}|_{D^\circ} = \lim_{n\to\infty} \int_{\mathcal{F}} f\, d\rho^{(n)}|_D = \int_{\mathcal{F}} f\, d\rho_D = \int_{\mathcal{F}} f\, d\rho_D|_{D^\circ} = \int_{\mathcal{F}} f\, d\rho|_{D^\circ}\,.$$

Since $f \in C_c(D^\circ)$ was arbitrary, we obtain vague convergence

$$\rho^{(n)}|_{D^\circ} \xrightarrow{v} \rho|_{D^\circ}\,.$$

This completes the proof. □

Having proved vague convergence on open subsets of $D \in \mathcal{D}$, the following result even yields weak convergence on suitable relatively compact subsets $V \subset \mathcal{F}$ (so-called *continuity sets*, cf. [**16**, Section 8.2]).

LEMMA 4.4.11. *For every $D \in \mathcal{D}$ there exists $E \in \mathcal{D}$ with $D \subset E^\circ$ as well as a relatively compact, open subset $V \subset E^\circ$ with $D \subset V$ such that*

$$\rho^{(n)}|_V \rightharpoonup \rho|_V \qquad \textit{weakly}\,. \tag{4.4.15}$$

Similarly, whenever $D \in \mathcal{D}^{(n)}$ and $U \supset D$ open, there exists a relatively compact, open subset $V \subset \mathcal{F}^{(n)}$ with $D \subset V \subset U$ such that (4.4.15) *holds.*

PROOF. Given $D \in \mathcal{D}$, by construction of $\mathcal{D}$ we know that D is compact and thus contained in the interior of some $E \in \mathcal{D}$ with $\rho(E) < \infty$ in view of (4.4.4) and (4.4.7). Therefore, $\rho|_E : \mathcal{B}(E) \to [0, \infty)$ is a non-negative finite Borel measure, where $\mathcal{B}(E)$ denotes the Borel σ-algebra on the topological space E. Since E is metrizable, it is completely regular, implying that the class $\Gamma_{\rho|_E}$ of all Borel sets in E with boundaries of $\rho|_E$-measure zero contains a base (consisting of open sets) of the topology of E (for details see [**16**, Proposition 8.2.8]). Since E is compact, the set $D \subset E$ can be covered by finitely many relatively compact, open sets $V_1, \ldots, V_N \subset E^\circ$ in $\Gamma_{\rho|_E}$. By construction, the closure of the set $V := \bigcup_{i=1}^N V_i \subset E^\circ$ is compact. Considering the restriction $\rho^{(n)}|_V$, for any $f \in C_c(V)$ we obtain

$$\lim_{n\to\infty} \int_V f \, d\rho^{(n)}|_V = \lim_{n\to\infty} \int_{\mathcal{F}} f \, d\rho^{(n)}|_{E^\circ} \overset{(4.4.12)}{=} \int_{\mathcal{F}} f \, d\rho|_{E^\circ} = \int_V f \, d\rho|_V \,,$$

proving vague convergence

$$\rho^{(n)}|_V \overset{v}{\to} \rho|_V \,. \tag{4.4.16}$$

Since $\Gamma_{\rho|_E}$ is a subalgebra in $\mathcal{B}(E)$ (see [**16**, Proposition 8.2.8]), the set V is also contained in $\Gamma_{\rho|_E}$, implying that $\rho(\partial V) = \rho|_E(\partial V) = 0$. Note that the measures $(\rho^{(n)}|_{E^\circ})$ and $\rho|_{E^\circ}$ are regular Borel measures and thus Radon [**143**]. Therefore, making use of vague convergence (4.4.12) and applying [**9**, Theorem 30.2], for the relatively compact, open set $V \subset E^\circ$ and the compact set $\overline{V} \subset E^\circ$ we obtain

$$\rho(V) = \rho(\overline{V}) \geq \limsup_{n\to\infty} \rho^{(n)}(\overline{V}) \geq \limsup_{n\to\infty} \rho^{(n)}(V) \geq \liminf_{n\to\infty} \rho^{(n)}(V) \geq \rho(V) \,,$$

proving that

$$\rho(V) = \lim_{n\to\infty} \rho^{(n)}(V) \,. \tag{4.4.17}$$

Let us point out that, for each $n \in \mathbb{N}$, the measure $\rho^{(n)}|_V / \rho^{(n)}(V)$ is normalized in the sense that $\rho^{(n)}|_V(V)/\rho^{(n)}(V) = 1$ (cf. [**66**, §3.2]). Furthermore, applying vague convergence (4.4.16) as well as (4.4.17), for any $f \in C_c(V)$ we are given

$$\lim_{n\to\infty} \int_V f \, d\rho^{(n)}|_V / \rho^{(n)}(V) = \int_V f \, d\rho|_V / \rho(V) \,.$$

As a consequence, the sequence of normalized measures $(\rho^{(n)}|_V / \rho^{(n)}(V))_{n\in\mathbb{N}}$ converges vaguely to the normalized measure $\rho|_V/\rho(V)$, and from [**9**, Corollary 30.9] we deduce that $(\rho^{(n)}|_V / \rho^{(n)}(V))_{n\in\mathbb{N}}$ converges *weakly* to the normalized measure $\rho|_V/\rho(V)$. Thus in view of

$$\begin{aligned} \lim_{n\to\infty} \int_V f \, d\rho^{(n)}|_V &= \lim_{n\to\infty} \rho^{(n)}(V) \int_V f \, d\rho^{(n)}|_V / \rho^{(n)}(V) = \rho(V) \int_V f \, d\rho^{(n)}|_V / \rho(V) \\ &= \int_V f \, d\rho^{(n)}|_V \end{aligned}$$

for any $f \in C_b(V)$, we finally obtain weak convergence $\rho^{(n)}|_V \rightharpoonup \rho|_V$. □

By contrast to [**66**], it is not reasonable to consider vague convergence $\rho^{(n)} \overset{v}{\to} \rho$ in view of [**100**, Exercise 14.4].

4.5. Minimizers for Lagrangians of Bounded Range

4.5.1. Preliminaries. This section is devoted to the proof that, under suitable assumptions, the measure ρ as defined in (4.4.6) is a minimizer of the causal variational principle (4.3.2) under variations of finite volume (see Definition 4.3.5). This is accomplished in §4.5.4. To this end, we proceed in several steps. In this subsection, we introduce an additional assumption on the measure ρ obtained in Theorem 4.4.3 (see (4.5.1)). Afterwards we prove that ρ is a minimizer on suitable compact subsets (see §4.5.2 and §4.5.3). Assuming that $\rho \neq 0$ is locally finite, we finally show that ρ satisfies corresponding Euler-Lagrange (EL) equations (see §4.5.5), which have the same structure as the EL equations obtained in [**66**]. Throughout this section, we shall assume that the Lagrangian is of bounded range (see Definition 4.3.7).

In order to prove that ρ is a minimizer, we impose the following condition:

(B) For any $\varepsilon > 0$ and $B \subset \mathcal{F}$ bounded, there exists $N \in \mathbb{N}$ such that

$$\rho(B \setminus B^{(n)}) < \varepsilon \qquad \text{for all } n \geq N\,, \tag{4.5.1}$$

where $B^{(n)} := B \cap \mathcal{F}^{(n)}$.

LEMMA 4.5.1. *Assume that the measure ρ defined by* (4.4.6) *satisfies condition (B). Then the measure ρ is locally finite, and any bounded subset of $\mathcal{F}$ has finite ρ-measure.*

PROOF. Assuming that $B \subset \mathcal{F}$ is bounded, in view of condition (B) there exists $N \in \mathbb{N}$ such that $\rho(B \setminus B^{(n)}) < \varepsilon$ for all $n \geq N$, where $B^{(n)} := B \cap \mathcal{F}^{(n)}$ is relatively compact in view of [**28**, Theorem I.4]. For this reason, $B^{(n)}$ can be covered by a finite number of compact sets $D_1, \ldots, D_L$ with $D_i \in \mathcal{D}^{(n)}$ for all $i = 1, \ldots, L$, where $\mathcal{D}^{(n)}$ is given by (4.4.1). From (4.4.4) we obtain $\rho(B^{(n)}) < \infty$, implying that

$$\rho(B) = \rho(B \setminus B^{(n)}) + \rho(B^{(n)}) < \infty\,.$$

Since each compact set $K \subset \mathcal{F}$ is bounded, the measure ρ is locally finite in view of Lemma 4.4.6. □

4.5.2. Minimizers under Variations of Finite-Dimensional Compact Support. Before proving our first existence result, we point out that the restricted Lagrangian

$$\mathcal{L}^{(n)} = \mathcal{L}|_{\mathcal{F}^{(n)} \times \mathcal{F}^{(n)}} : \mathcal{F}^{(n)} \times \mathcal{F}^{(n)} \to \mathbb{R}_0^+$$

is of compact range (see Definition 4.2.4 and §4.3.3) for every $n \in \mathbb{N}$. Therefore, for all $j, k, n \in \mathbb{N}$, there exist compact subsets $(V_{j,k}^{(n)})' \subset \mathcal{F}^{(n)}$ such that $\mathcal{L}(x, y) = 0$ for all $x \in V_{j,k}^{(n)}$ and $y \in \mathcal{F}^{(n)} \setminus (V_{j,k}^{(n)})'$. By construction of $\mathcal{D}$ (see §4.4.1), the set $(V_{j,k}^{(n)})'$ can be covered by a finite number of sets $(V_{j'_\ell,k'_\ell}^{(n)})_{\ell=1,\ldots,L}$ in $\mathcal{D}$, whose union is also contained in $\mathcal{D}$. For this reason, we may assume that $(V_{j,k}^{(n)})' \in \mathcal{D}$ for all $j, k, n \in \mathbb{N}$.

After these preparations, we are now in the position to state our first existence result.

PROPOSITION 4.5.2. *Assume that the Lagrangian $\mathcal{L} \in C_b(\mathcal{F} \times \mathcal{F}; \mathbb{R}_0^+)$ is of bounded range, and that condition* (4.3.6) *holds. Moreover, assume that the measure ρ defined by* (4.4.6) *satisfies condition* (B) *in Section 4.5. Then ρ is a* ***minimizer under variations in*** *$\mathcal{D}$ in the sense that*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0$$

whenever $\tilde{\rho}$ satisfying (4.2.3) *is a regular measure on $\mathcal{B}(\mathcal{F})$ with* $\mathrm{supp}(\tilde{\rho} - \rho) \in \mathcal{D}$.

PROOF. Assume that $\tilde{\rho} : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ is a measure on the Borel σ-algebra of $\mathcal{F}$ with $D := \text{supp}(\tilde{\rho} - \rho) \in \mathcal{D}$ such that (4.2.3) is satisfied, i.e.

$$0 < \tilde{\rho}(D) = \rho(D) < \infty \,.$$

Since $D \in \mathcal{D}$ is compact, the difference (4.2.4) is well-defined in view of [**66**, §4.3],

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) &= \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \\ &\quad + \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \,.\end{aligned}$$

Making use of the symmetry of the Lagrangian and applying Fubini's theorem, we can write this expression as

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) &= 2 \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \,.\end{aligned}$$

Since $D \in \mathcal{D}$ is a bounded subset of $\mathcal{F}$, Definition 4.3.7 yields the existence of some bounded set $W \subset \mathcal{F}$ with $\mathcal{L}(x,y) = 0$ for all $x \in D$ and $y \in \mathcal{F} \setminus W$. For arbitrary $\tilde{\varepsilon} > 0$, by virtue of condition (B) in Section 4.5 (see (4.5.1)) there is some integer $n' = n'(D)$ such that $\rho(W \setminus W^{(n)}) < \tilde{\varepsilon}$ for all $n \geq n'$. We choose $n' \in \mathbb{N}$ sufficiently large and let $\tilde{D} \in \mathcal{D}$ with $\tilde{D} \subset \mathcal{F}^{(n')}$ and $\tilde{D} \supset D$. As explained at the beginning of §4.5.2, there exists $\tilde{D}' \in \mathcal{D}$ with $\tilde{D}' \subset \mathcal{F}^{(n')}$ and $\mathcal{L}(x,y) = 0$ for all $x \in \tilde{D}$ and $y \in \mathcal{F}^{(n')} \setminus \tilde{D}'$. In particular, $\mathcal{L}(x,y) = 0$ for all $x \in D$ and $y \in \mathcal{F}^{(n')} \setminus \tilde{D}'$. In view of Lemma 4.4.11, there exist relatively compact open sets $V, V' \subset \mathcal{F}^{(n')}$ with $V \supset \tilde{D}$ and $V' \supset \tilde{D}'$ in such a way that $\mathcal{L}(x,y) = 0$ for all $x \in V$ and $y \in \mathcal{F}^{(n')} \setminus V'$, and $\rho^{(n)}|_V \rightharpoonup \rho|_V$ as well as $\rho^{(n)}|_{V'} \rightharpoonup \rho|_{V'}$. These considerations give rise to

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) &= 2 \int_V d(\tilde{\rho} - \rho)(x) \int_{V'} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + 2 \int_V d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F} \setminus V'} d\rho(y)\, \mathcal{L}(x,y) + \int_V d(\tilde{\rho} - \rho)(x) \int_V d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) \,,\end{aligned}$$

where the expression

$$\int_V d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F} \setminus V'} d\rho(y)\, \mathcal{L}(x,y) = \int_D d(\tilde{\rho} - \rho)(x) \int_{W \setminus W^{(n')}} d\rho(y)\, \mathcal{L}(x,y)$$

can be chosen arbitrarily small for sufficiently large $n' \in \mathbb{N}$. Thus for $\varepsilon > 0$ arbitrary, we may arrange that

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) &\geq 2 \int_V d(\tilde{\rho} - \rho)(x) \int_{V'} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + \int_V d(\tilde{\rho} - \rho)(x) \int_V d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y) - \varepsilon \,.\end{aligned}$$

Next, in analogy to the proof of [**66**, Theorem 4.9], for any $n \in \mathbb{N}$ we introduce the measures $\tilde{\rho}_n : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ by

$$\tilde{\rho}_n := \begin{cases} c_n\, \tilde{\rho} & \text{on } V \\ \rho^{(n)} & \text{on } \mathcal{F} \setminus V \end{cases} \qquad \text{with} \qquad c_n := \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)} \qquad \text{for all } n \in \mathbb{N} \,.$$

Since ρ and $\tilde{\rho}$ coincide on $\mathcal{F} \setminus D$, by virtue of (4.4.17) (see Lemma 4.4.11) we obtain

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} \frac{\rho^{(n)}(V)}{\tilde{\rho}(V)} = \frac{\rho(V)}{\tilde{\rho}(V)} = \frac{\rho(V \setminus D) + \rho(D)}{\tilde{\rho}(V \setminus D) + \tilde{\rho}(D)} = 1\,.$$

Making use of the fact that $V \subset \mathcal{F}$ is separable (see for instance [**4**, Corollary 3.5]) and applying [**13**, Theorem 2.8] in a similar fashion to the proof of [**66**, Theorem 4.9], we thus arrive at

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \lim_{n\to\infty} \Bigg[& 2 \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_{V'} d\rho^{(n)}(y)\, \mathcal{L}(x,y) \\ & + \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(x) \int_V d\big(c_n\, \tilde{\rho} - \rho^{(n)}\big)(y)\, \mathcal{L}(x,y) \Bigg] - 2\varepsilon\,.\end{aligned}$$

In view of the fact that $\tilde{\rho}_n$ and $\rho^{(n)}$ coincide on $\mathcal{F}^{(n)} \setminus V$ for sufficiently large $n \in \mathbb{N}$ and $\mathcal{L}(x,y) = 0$ for all $x \in V$ and $y \notin V'$, the difference $\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)$ can finally be estimated by

$$\begin{aligned}\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho) \geq \lim_{n\to\infty} \Bigg[& 2 \int_{\mathcal{F}^{(n)}} d\big(\tilde{\rho}_n - \rho^{(n)}\big)(x) \int_{\mathcal{F}^{(n)}} d\rho^{(n)}(y)\, \mathcal{L}(x,y) \\ & + \int_{\mathcal{F}^{(n)}} d\big(\tilde{\rho}_n - \rho^{(n)}\big)(x) \int_{\mathcal{F}^{(n)}} d\big(\tilde{\rho}_n - \rho^{(n)}\big)(y)\, \mathcal{L}(x,y) \Bigg] - 2\varepsilon\,.\end{aligned}$$

Since $\rho^{(n)}$ is a minimizer on $\mathcal{F}^{(n)}$ for every $n \in \mathbb{N}$ (see §4.3.3 and §4.4.2), we are given

$$\big(\mathcal{S}_{\mathcal{F}^{(n)}}(\tilde{\rho}_n) - \mathcal{S}_{\mathcal{F}^{(n)}}(\rho^{(n)})\big) \geq 0 \qquad \text{for all } n \in \mathbb{N}\,. \tag{4.5.2}$$

Taking the limit $n \to \infty$ on the left hand side of (4.5.2), one obtains exactly the above expression in square brackets for $\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)$. Since $\varepsilon > 0$ is arbitrary, this implies that

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

Hence ρ is a minimizer under variations in $\mathcal{D}$. □

Our next goal is to extend the previous result to minimizers under variations of finite-dimensional compact support, which is defined as follows.

Definition 4.5.3. *A regular measure ρ on $\mathcal{B}(\mathcal{F})$ is said to be a* ***minimizer under variations of finite-dimensional compact support*** *if the inequality*

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0$$

holds for any regular measure $\tilde{\rho}$ on $\mathcal{B}(\mathcal{F})$ satisfying (4.2.3) *with the following property: There exists $n' \in \mathbb{N}$ such that, for all $n \geq n'$,*

$$\operatorname{supp}\big(\tilde{\rho} - \rho\big) \subset \mathcal{F}^{(n)} \text{ compact} \qquad \text{and} \qquad \operatorname{supp}\big(\tilde{\rho} - \rho\big) \cap \mathcal{F} \setminus \mathcal{F}^{(n)} = \varnothing\,.$$

Based on this definition, we may state the following existence result.

Proposition 4.5.4. *Assume that $\mathcal{L} \in C_b(\mathcal{F} \times \mathcal{F}; \mathbb{R}_0^+)$ is of bounded range, and assume that condition* (4.3.6) *holds. Furthermore, assume that the measure ρ defined by* (4.4.6) *satisfies condition* (B) *in Section 4.5. Then ρ is a minimizer under variations of finite-dimensional compact support.*

Proof. Assuming that $\tilde{\rho}$ is a variation of finite-dimensional compact support, there exists $n' \in \mathbb{N}$ such that

$$K_\sharp := \operatorname{supp}\big(\tilde{\rho} - \rho\big) \subset \mathcal{F}^{(n)}$$

is compact for all $n \geq n'$, implying that $\tilde{\rho}(K_\sharp) = \rho(K_\sharp) < \infty$ (in view of Lemma 4.5.1). According to Lemma 4.4.4, the measure ρ given by (4.4.6) is regular, and by regularity of ρ and $\tilde{\rho}$, for $\tilde{\varepsilon} > 0$ arbitrary we may choose $U \supset K_\sharp$ open such that

$$\rho(U \setminus K_\sharp) < \tilde{\varepsilon} \qquad \text{and} \qquad \tilde{\rho}(U \setminus K_\sharp) < \tilde{\varepsilon}\,.$$

By construction of $\mathcal{D}$, there is some compact set $D \in \mathcal{D}$ such that $K_\sharp \subset D \subset U^{(n')}$, where $U^{(n')} := U \cap \mathcal{F}^{(n')}$. In particular,

$$\rho(D \setminus K_\sharp) < \tilde{\varepsilon} \qquad \text{and} \qquad \tilde{\rho}(D \setminus K_\sharp) < \tilde{\varepsilon}\,.$$

Applying Lemma 4.4.11 yields the existence of some relatively compact set $V \subset \mathcal{F}$ with $D \subset V \subset U^{(n')}$ such that (4.4.15) holds. Moreover,

$$\rho(V \setminus K_\sharp) < \tilde{\varepsilon} \qquad \text{and} \qquad \tilde{\rho}(V \setminus K_\sharp) < \tilde{\varepsilon}\,.$$

Since $\tilde{\rho} - \rho$ is a signed measure of finite total variation and compact support, the difference (4.2.4) is well-defined (cf. [**66**, §4.3]),

$$\begin{aligned}(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)) &= 2 \int_{K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + \int_{K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{K_\sharp} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y)\,.\end{aligned}$$

By adding and subtracting the terms

$$2 \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{K_\sharp} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y)$$

as well as

$$\int_V d(\tilde{\rho}-\rho)(x) \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y)\,,$$

one easily verifies that

$$\begin{aligned}\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) &= 2 \int_V d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ &\quad + \int_V d(\tilde{\rho}-\rho)(x) \int_V d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y) - \bigg[\int_V d(\tilde{\rho}-\rho)(x) \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y) \\ &\quad + 2 \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{K_\sharp} d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y) \bigg]\,.\end{aligned}$$

Since the Lagrangian $\mathcal{L}$ is of bounded range (see Definition 4.3.7), its restriction $\mathcal{L}^{(n')}$ is of compact range, implying that $\mathcal{L}(x,y) = 0$ for all $x \in V$ and $y \notin V'$ for some relatively compact, open set $V' \subset \mathcal{F}^{(n')}$ such that (4.4.15) holds. Choosing the open set $U \supset K_\sharp$ suitably, one thus can arrange that

$$\left| \int_{V \setminus K_\sharp} d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \right| \leq \underbrace{\sup_{x,y \in V'} \mathcal{L}(x,y)\, \rho(V')}_{<\infty} \underbrace{\big(|\tilde{\rho}(V \setminus K_\sharp)| + |\rho(V \setminus K_\sharp)| \big)}_{<\tilde{\varepsilon}}$$

is arbitrarily small. Applying similar arguments to all summands of the above term in square brackets, one obtains the estimate

$$\big(\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho)\big) \geq \Big\{ 2\int_V d(\tilde\rho-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_V d(\tilde\rho-\rho)(x) \int_V d(\tilde\rho-\rho)(y)\, \mathcal{L}(x,y) \Big\} - \varepsilon$$

for any given $\varepsilon > 0$. Proceeding in analogy to the proof of Proposition 4.5.2 by applying weak convergence (4.4.15) together with [**13**, Theorem 2.8] (for details we refer to the proof of [**66**, Theorem 4.9]), one can show that the term in curly brackets is greater than or equal to zero, up to an arbitrarily small error term. Since $\varepsilon > 0$ was chosen arbitrarily, we finally arrive at

$$\big(\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho)\big) \geq 0\,,$$

which proves the claim. □

4.5.3. Existence of Minimizers under Variations of Compact Support. Having derived the above preparatory results in §4.5.2, we are now in the position to deal with minimizers under variations of compact support (see Definition 4.3.6). For clarity, we point out that, whenever $\tilde\rho$ is a variation of compact support of the measure ρ satisfying (4.2.3), the condition $|\tilde\rho-\rho| < \infty$ yields $\rho(K) = \tilde\rho(K) < \infty$, where the compact set $K \subset \mathcal{F}$ is defined by $K := \mathrm{supp}(\tilde\rho-\rho)$ (for details see Definition 4.2.1 and the explanations in §4.2.2).

LEMMA 4.5.5. *Assume that $\mathcal{L} \in C_b(\mathcal{F}\times\mathcal{F};\mathbb{R}^+_0)$ is of bounded range, and assume that condition* (4.3.6) *holds. Moreover, assume that the measure ρ defined by* (4.4.6) *satisfies condition* (B) *in Section 4.5. Then ρ is a minimizer under variations of compact support.*

PROOF. Let $\tilde\rho$ be a variation of compact support. Then the set $K := \mathrm{supp}(\tilde\rho-\rho) \subset \mathcal{F}$ is compact, and $\rho(K) = \tilde\rho(K) < \infty$ according to (4.2.3). Given $\tilde\varepsilon > 0$ arbitrary, by regularity of ρ, $\tilde\rho$ there exists $U \supset K$ open such that $\rho(U), \tilde\rho(U) < \infty$ and

$$\rho(U\setminus K) < \tilde\varepsilon \qquad \text{and} \qquad \tilde\rho(U\setminus K) < \tilde\varepsilon\,.$$

Moreover, in view of (4.4.5), there exists $D \in \mathfrak{D}$ such that

$$\rho(U\setminus D) < \tilde\varepsilon \qquad \text{and} \qquad \tilde\rho(U\setminus D) < \tilde\varepsilon\,.$$

Since $K \subset \mathcal{F}$ is compact, the difference (4.2.4) is well-defined (cf. [**66**, §4.3]),

$$\big(\mathcal{S}(\tilde\rho)-\mathcal{S}(\rho)\big) = 2\int_K d(\tilde\rho-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_K d(\tilde\rho-\rho)(x) \int_K d(\tilde\rho-\rho)(y)\, \mathcal{L}(x,y)\,.$$

Note that, by definition of $\mathfrak{D}$, each $D \in \mathfrak{D}$ is the finite union of finite-dimensional subsets of $(\mathcal{F}^{(n)})_{n\in\mathbb{N}}$ (cf. §4.4.1). As a consequence, for each $D \in \mathfrak{D}$ there exists $n' \in \mathbb{N}$ such that $D \subset \mathcal{F}^{(n)}$ for all $n \geq n'$. Moreover, by construction of $\mathfrak{D}$ there exists $E \in \mathfrak{D}$ such that $D \subset E^\circ \subset U^{(n)}$, where $U^{(n)} = U \cap \mathcal{F}^{(n)}$, and according to Lemma 4.4.11 there exists a relatively compact, open set $V \subset E^\circ$ such that $D \subset V$. In particular,

$$\rho(U\setminus V) < \tilde\varepsilon\,, \qquad \tilde\rho(U\setminus V) < \tilde\varepsilon\,.$$

Thus adding and subtracting the terms

$$2 \int_{U \setminus K} d(\tilde{\rho} - \rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) + \int_{U \setminus K} d(\tilde{\rho} - \rho)(x) \int_K d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)$$

as well as

$$\int_U d(\tilde{\rho} - \rho)(x) \int_{U \setminus K} d(\tilde{\rho} - \rho)(y)\, \mathcal{L}(x,y)\,,$$

and proceeding in analogy to the proof of Proposition 4.5.4, we conclude that

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,.$$

Hence ρ is indeed a minimizer under variations of compact support. □

4.5.4. Existence of Minimizers under Variations of Finite Volume. We now proceed similarly to [**66**, §4.4] in order to prove the existence of minimizers under variations of finite volume (see Definition 4.3.5). For the difference of the actions (4.2.4) to be well-defined, we require the additional property (iv) in §4.2.2, i.e.

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty\,.$$

Then we can state the following result.

THEOREM 4.5.6. *Assume that $\mathcal{L} \in C_b(\mathcal{F} \times \mathcal{F}; \mathbb{R}_0^+)$ is of bounded range, and assume that condition* (4.3.6) *holds. Furthermore, assume that the measure ρ defined by* (4.4.6) *satisfies condition* (B) *in Section 4.5. Then ρ is a minimizer under variations of finite volume.*

PROOF. Assume that $\tilde{\rho}$ be a variation of finite volume satisfying (4.2.3). Introducing the set $B := \operatorname{supp}(\tilde{\rho} - \rho)$, we thus obtain $\rho(B) = \tilde{\rho}(B) < \infty$. Given $\tilde{\varepsilon} > 0$ arbitrary, by regularity of ρ, $\tilde{\rho}$ there exists $U \supset B$ open such that

$$\rho(U \setminus B) < \tilde{\varepsilon} \qquad \text{and} \qquad \tilde{\rho}(U \setminus B) < \tilde{\varepsilon}\,.$$

Making use of the additional assumption that condition (iv) in §4.2.2 is satisfied, the difference of the actions (4.2.4) is well-defined. Therefore, proceeding in analogy to the proof of Lemma 4.5.5 finally gives rise to

$$\big(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)\big) \geq 0\,,$$

which implies that ρ is a minimizer under variations of finite volume. □

The remainder of this section is devoted to the derivation of the corresponding EL equations for minimizers under variations of finite volume (§4.5.5).

4.5.5. Derivation of the Euler-Lagrange Equations. The strategy in [**66**] was to derive the EL equations in order to prove the existence of minimizers under variations of finite volume. However, proceeding similarly to [**66**, §4.2] does not seem promising in the infinite-dimensional setting. For this reason, we rather proceed in the opposite direction by first proving the existence of minimizers and then deriving the EL equations. More precisely, under the assumptions that condition (iv) in §4.2.2 as well as condition (B) in Section 4.5 are satisfied, Theorem 4.5.6 shows that the measure ρ given by (4.4.6) is a minimizer under variations of finite volume (see Definition 4.3.5). Under these assumptions, Lemma 4.4.8 implies that ρ is locally finite. This allows us to proceed similarly to the proof of [**65**, Lemma 2.3], thus giving rise to corresponding EL equations. For convenience, let us state the latter result in greater generality.

THEOREM 4.5.7 (**The Euler-Lagrange equations**). *Let $\mathcal{F}$ be topological Hausdorff space, let ρ be a Borel measure on $\mathcal{F}$ (in the sense of* [**76**]*, i.e. a locally finite measure) and assume that $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is symmetric and lower semi-continuous. If $\rho \neq 0$ is a minimizer of the causal variational principle* (4.3.2), (4.2.3) *under variations of finite volume, then the* Euler-Lagrange equations

$$\ell|_{\text{supp}\,\rho} \equiv \inf_{x \in \mathcal{F}} \ell(x) \tag{4.5.3}$$

hold, where the mapping $\ell : \mathcal{F} \to [0, \infty)$ is defined by

$$\ell(x) := \int_{\mathcal{F}} \mathcal{L}(x, y)\, d\rho(y) - \mathfrak{s} \tag{4.5.4}$$

for some parameter $\mathfrak{s} \in \mathbb{R}$.

PROOF. Proceed in analogy to the proof of [**65**, Lemma 2.3]. □

For clarity, we point out that Theorem 4.5.7 requires that ρ is locally finite. Choosing the parameter $\mathfrak{s}$ suitably, one can arrange that the infimum in (4.5.3) vanishes:

LEMMA 4.5.8. *Assume that the measure ρ given by* (4.4.6) *is non-zero. Then, under the assumptions of Theorem 4.5.6, for a suitable choice of $\mathfrak{s} \geq 0$ in* (4.5.4) *the Euler-Lagrange equations* (4.5.3) *read*

$$\ell|_{\text{supp}\,\rho} \equiv \inf_{x \in \mathcal{F}} \ell(x) = 0\,. \tag{4.5.5}$$

PROOF. Assuming that $\rho \neq 0$, we conclude that $\text{supp}\,\rho \neq \varnothing$. Moreover, under the assumptions of Theorem 4.5.6, by Lemma 4.4.8 we know that ρ is locally finite. Next, from Theorem 4.5.6 and Theorem 4.5.7 we deduce that ρ satisfies the EL equations (4.5.3). By assumption, condition (iv) in §4.2.2 is satisfied, implying that

$$0 \leq \int_{\mathcal{F}} \mathcal{L}(x, y)\, d\rho(y) < \infty \qquad \text{for every } x \in \mathcal{F}\,.$$

This allows us to choose $\mathfrak{s} \geq 0$ such that (4.5.5) holds. □

Lemma 4.5.8 generalizes the results of [**65**, Section 2] and [**66**, Section 4] to the infinite-dimensional setting. It remains an open task to prove the existence of a Lagrangian of bounded range such that the measure ρ in (4.4.6) is non-zero and satisfies condition (iv) in §4.2.2 as well as condition (B) in Section 4.5.

4.6. Minimizers for Lagrangians Vanishing in Entropy

In Section 4.5 the results from [**66**, Section 4] were generalized to the non-locally compact setting. This raises the question whether it is possible also to weaken the assumption that the Lagrangian is of bounded range (see Definition 4.3.7) similarly to Lagrangians decaying in entropy as introduced in [**66**, Section 5]. It is precisely the objective of this section to analyze this question in detail. To this end, we first generalize the notion of Lagrangians decaying in entropy (§4.6.1). Afterwards we proceed similarly to [**66**, Section 5] and Section 4.5 to prove that, under suitable assumptions, the measure ρ obtained in Theorem 4.4.3 is a minimizer under variations of compact support and variations of finite volume (see Definition 4.3.6 and Definition 4.3.5).

4.6.1. Lagrangians Vanishing in Entropy. This subsection is devoted to generalize the notion of Lagrangians decaying in entropy as introduced in [**66**, Section 5]. More precisely, in order for the constructions in [**66**, Section 5] to work, the definition of Lagrangians decaying in entropy (see Definition 4.2.8 and [**66**, Definition 5.1]) requires an *unbounded* Heine-Borel metric. On the other hand, any Heine-Borel space (that is, a topological space endowed with a Heine-Borel metric) is σ-compact and locally compact, see [**162**]. In particular, every separable Heine-Borel space X is a second-countable, locally compact Hausdorff space, and hemicompact (see [**161**, Problem 17I]) in view of [**41**, Exercise 3.8.C]. Accordingly, there is a sequence $(K_n)_{n\in\mathbb{N}}$ of compact subsets of X with $K_n \subset K_{n+1}^\circ$ for every $n \in \mathbb{N}$ such that any compact set $K \subset X$ is contained in K_n for some $n \in \mathbb{N}$ and $X = \bigcup_{n=1}^\infty K_n$ (also see [**9**, Lemma 29.8]). Moreover, in view of [**9**, Theorem 31.5], the space X is Polish. These considerations motivate the following procedure.

For any second-countable, locally compact Hausdorff topological space X, assume that the Lagrangian $\mathcal{L} : X \times X \to \mathbb{R}_0^+$ is continuous, symmetric and positive on the diagonal (4.3.1). Moreover, assume that the measure $\tilde{\rho}$ on $\mathcal{B}(X)$ is obtained similarly to the constructions in [**66**, §4.1]. In order for the constructions in [**66**, Section 5] to apply, one requires that, for any $x \in X$ and arbitrary $\varepsilon > 0$, there exists $K_{x,\varepsilon} \subset X$ compact such that

$$\int_{X\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\tilde{\rho}(y) < \varepsilon\,.$$

The results in [**162**] imply that for any separable, locally compact metric space (X,d) there is a Heine-Borel metric d_{HB} on X which generates the same topology. In order to also allow for *bounded* Heine-Borel metrics in [**66**, Section 5], it seems preferable to not specify the set $K_{x,\varepsilon}$ in [**66**, eq. (5.2)] and the calculations thereafter in terms of a possibly bounded Heine-Borel metric d_{HB}. For this reason, it seems preferable to formulate the calculations in [**66**, §5.1] purely in terms of compact subsets rather than in terms of (closed) balls (whose diameter depends on the corresponding Heine-Borel metric). To this end, for any topological space Y, we denote the set of all functions $f : Y \to \mathbb{R}$ by $\mathbb{F}(Y)$, and let $\mathbb{F}^+(Y)$ be the subset of non-negative such functions. Given a second-countable, locally compact Hausdorff space X, by $\mathbb{F}_0^+(X)$ we denote the subset of non-negative functions *vanishing at infinity* in the sense that, for any $\varepsilon > 0$, there exists $K \subset X$ compact with $f|_{X\setminus K} < \varepsilon$ (for continuous functions vanishing at infinity we refer to [**40**, §VIII.2]).

This allows us to generalize the definition of Lagrangians decaying in entropy by restating condition (c) in Definition 4.2.8 in the following way: Given the compact exhaustion $(K_m)_{m\in\mathbb{N}}$ of X with $K_m \subset K_{m+1}^\circ$ for every $m \in \mathbb{N}$ and $X = \bigcup_{m=1}^\infty K_m$, for every $x \in X$ let $N = N(x)$ be the least integer such that $x \in K_m$ for all $m \geq N$. We now introduce the sets $(K_m(x))_{m\in\mathbb{N}}$ by

$$K_m(x) := K_{m+N-1} \qquad \text{for all } m \in \mathbb{N}\,. \tag{4.6.1}$$

Introducing *entropy* $E_x(K_m(x), \delta)$ according to §4.2.2 as the smallest number of balls of radius $\delta > 0$ covering $K_m(x)$, and replacing (c) in Definition 4.2.8 by (c'), we define Lagrangians vanishing in entropy as follows.

Definition 4.6.1. *Suppose that (X,d) is a second-countable, locally compact metric space. Then the Lagrangian $\mathcal{L} : X \times X \to \mathbb{R}_0^+$ is said to* **vanish in entropy** *if the following conditions are satisfied:*

(a) $c := \inf_{x\in X} \mathcal{L}(x,x) > 0$.

(b) *There is a compact set $K \subset X$ such that*

$$\delta := \inf_{x\in X\setminus K} \sup\Big\{ s\in\mathbb{R} \,:\, \mathcal{L}(x,y) \geq \frac{c}{2} \quad \textit{for all } y\in B_s(x)\Big\} > 0\,.$$

(c') *The Lagrangian has the following decay property: Given an exhaustion of X by compact subsets $(K_m)_{m\in\mathbb{N}}$, there exists $f : X\times X \to \mathbb{R}_0^+$ with $f(x,\cdot)\in \mathbb{F}_0^+(X)$ for every $x\in X$ such that, for every $x\in X$ and all $m\in\mathbb{N}$,*

$$\mathcal{L}(x,y) \leq 2^{-m}\frac{f(x,y)}{C_x(m,\delta)} \qquad \textit{for all } y\in K_m(x)\,,$$

where $(K_m(x))_{m\in\mathbb{N}}$ is defined by (4.6.1),

$$C_x(m,\delta) := C\, E_x(K_{m+2}(x),\delta) \qquad \textit{for all } x\in\mathcal{F},\ m\in\mathbb{N} \textit{ and } \delta>0$$

(with entropy $E_x(K_m(x),\delta)$ as introduced in §4.2.2), and the constant C is given by

$$C := 1+\frac{2}{c} < \infty\,.$$

As mentioned in [**66**], we may assume that $\delta = 1$ (otherwise we suitably rescale the corresponding metric on X). Let us point out that Definition 4.6.1, by contrast to Definition 4.2.8 (see [**66**, Definition 5.1]), does not require a Heine-Borel metric and thus allows for more general applications. Definition 4.2.8 can be considered as a special case of Definition 4.6.1.

Under the assumptions (a), (b), (c'), for any $x\in X$ and $\varepsilon>0$ there exists $K_{x,\varepsilon}\subset X$ compact such that

$$\int_{X\setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\tilde\rho(y) < \varepsilon\,.$$

To see this, we make use of the fact that $K_n \subset K_{n+1}^\circ$ for all $n\in\mathbb{N}$. Given $x\in X$ and arbitrary $\varepsilon>0$, there exists $\tilde K_{x,\varepsilon}\subset X$ compact with $f(x,y)<\varepsilon/6$ for all $y\notin \tilde K_{x,\varepsilon}$. Since X is hemicompact, there exists $n\in\mathbb{N}$ with $\tilde K_{x,\varepsilon}\subset K_n$. We denote the least such integer by $N_0 = N_0(x,\varepsilon)$. Then the compact set (cf. [**66**, eq. (5.2)])

$$K_{x,\varepsilon} := K_{N_0} \subset X$$

has the desired property:

$$\begin{aligned}\int_{X\setminus K_{N_0}} \mathcal{L}(x,y)\, d\tilde\rho(y) &= \sum_{m=N_0}^\infty \int_{K_{m+1}\setminus K_m} \mathcal{L}(x,y)\, d\tilde\rho(y)\\ &\leq \sum_{m=N_0}^\infty \sup_{y\in K_{m+1}} \mathcal{L}(x,y)\, \underbrace{\tilde\rho\,(K_{m+1}\setminus K_m)}_{\leq C_x(m,1)} \leq \sup_{y\in X\setminus K_{N_0}} f(x,y) \sum_{m=N_0}^\infty 2^{-m} < \varepsilon/3\,.\end{aligned}$$

By definition of $C_x(m,\delta)$, we are given

$$\int_{X\setminus K_{x,\varepsilon}} \mathcal{L}(\tilde x,y)\, d\tilde\rho(y) < \varepsilon/3$$

for all $\tilde x$ in a sufficiently small neighborhood of x.

Assuming that the Lagrangian is continuous, we proceed similarly to [**66**] to prove that the same is true for the measures $\tilde\rho^{(n)}$ as given by [**66**, eq. (4.5)] (where the measures $\tilde\rho^{(n)}$

originate in the same manner as in [**66**, §4.1]). More precisely, for any given $x \in X$ and $\varepsilon > 0$, we introduce the compact sets $A_m(x) \subset X$ by

$$A_m(x) := \overline{K_{m+1}(x) \setminus K_m(x)} \qquad \text{for all } m \geq N_0 = N_0(x, \varepsilon)\,.$$

Next, regularity of $\tilde{\rho}$ yields the existence of open sets $U_m(x) \supset A_m(x)$ with $U_m(x) \subset K_{m+1}(x) \setminus K_{m-1}(x)$ such that

$$\tilde{\rho}\big(U_m \setminus \big(K_{m+1}(x) \setminus K_m(x)\big)\big) < 2^{-m-1}\varepsilon/3 \qquad \text{for all } m \geq N_0$$

In view of [**4**, Lemma 2.92], for every $m \geq N_0$ there exists $\eta_m \in C_c(U_m(x); [0,1])$ such that $\eta_m|_{A_m(x)} \equiv 1$, implying that $\mathcal{L}(x,\cdot)\,\eta_m \in C_c(U_m(x))$ for all $m \geq N_0$. Repeating the arguments in [**66**], we finally arrive at [**66**, eq. (5.6)], i.e.

$$\int_{X \setminus K_{x,\varepsilon}} \mathcal{L}(\tilde{x}, y)\, d\tilde{\rho}^{(n)}(y) < \varepsilon \qquad \text{and} \qquad \int_{X \setminus K_{x,\varepsilon}} \mathcal{L}(\tilde{x}, y)\, d\tilde{\rho}(y) < \varepsilon$$

for all $\tilde{x}$ in a small neighborhood of x and sufficiently large $n \in \mathbb{N}$. As a consequence, all results in [**66**, Section 5] remain valid for Lagrangians decreasing in entropy.

The advantage of Definition 4.6.1 is that it applies to arbitrary second-countable, locally compact metric spaces. In particular, by contrast to Definition 4.2.8, it need not be endowed with an unbounded Heine-Borel metric. Furthermore, the concept of Lagrangians vanishing in entropy carries over to possibly non-locally compact metric spaces in the following way.

DEFINITION 4.6.2. *Given a metric space (X, d), the Lagrangian $\mathcal{L} : X \times X \to \mathbb{R}_0^+$ is said to **vanish in entropy** if, for any second-countable, locally compact topological Hausdorff space $Y \subset X$, its restriction $\mathcal{L}|_{Y \times Y} : Y \times Y \to \mathbb{R}_0^+$ vanishes in entropy with respect to the induced metric $d_Y := d|_{Y \times Y}$ (see Definition 4.6.1).*

4.6.2. Preparatory Results. After these preliminaries we return to causal variational principles in the non-locally compact setting (see Definition 4.3.4). Accordingly, let X be a separable infinite-dimensional complex Banach space and assume that $\mathcal{F} \subset X$ is a non-locally compact Polish subspace (with respect to the Fréchet metric d induced by the norm on X). Then $(\mathcal{F}, d)$ is a separable, complete metric space. In what follows we assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ vanishes in entropy (see Definition 4.6.2 and Definition 4.6.1). Considering the finite-dimensional exhaustion $\mathcal{F}^{(n)} \subset \mathcal{F}$ endowed with the induced metric $d_n := d|_{\mathcal{F}^{(n)} \times \mathcal{F}^{(n)}}$ for all $n \in \mathbb{N}$, the explanations in §4.3.3 imply that each closed bounded subset of $\mathcal{F}^{(n)}$ (with respect to d_n) is compact. Moreover, the restricted Lagrangians $\mathcal{L}^{(n)} : \mathcal{F}^{(n)} \times \mathcal{F}^{(n)} \to \mathbb{R}_0^+$ vanish in entropy (see Definition 4.6.1). As outlined in §4.6.1, all results in [**66**, Section 5] remain valid if we replace "decaying in entropy" by "vanishing in entropy." In particular, by applying [**66**, Theorem 5.8] we conclude that for each $n \in \mathbb{N}$, there is some regular Borel measure $\rho^{[n]}$ on $\mathcal{F}^{(n)}$ which is a minimizer of the corresponding action $\mathcal{S}^{(n)} := \mathcal{S}_{\mathcal{F}^{(n)}}$ under variations of compact support, where

$$\mathcal{S}_E(\rho) := \int_E d\rho(x) \int_E d\rho(y)\, \mathcal{L}(x, y)$$

for any $E \in \mathcal{B}(\mathcal{F})$ (cf. [**66**, §3.2]). Moreover, in view of [**66**, Theorem 5.5], for all $n \in \mathbb{N}$ the following Euler-Lagrange equations hold,

$$\ell^{[n]}|_{\operatorname{supp} \rho^{[n]}} \equiv \inf_{x \in \mathcal{F}} \ell^{[n]}(x) = 0\,,$$

where the mapping $\ell^{[n]} : \mathcal{F} \to \mathbb{R}$ is defined by

$$\ell^{[n]}(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho^{[n]}(y) - 1\,.$$

We point out that Lemma 4.4.1 is applicable to the sequence $(\rho^{[n]})_{n\in\mathbb{N}}$, implying that

$$\rho^{[n]}(K) \leq C_K \qquad \text{for all } n \in \mathbb{N}\,.$$

For this reason, we may proceed in analogy to Section 4.4 by introducing a countable set $\mathcal{D} \subset \mathfrak{K}(\mathcal{F})$ (see §4.4.1). Next, in analogy to §4.4.2 we iteratively restrict the sequence of measures $(\rho^{[n]})_{n\in\mathbb{N}}$ to $D_m \subset \mathcal{F}$ compact with $D_m \in \mathcal{D}$ for all $m \in \mathbb{N}$ and denote the resulting diagonal sequence by $(\rho^{(k)})_{k\in\mathbb{N}}$ (cf. (4.4.2)). Defining the corresponding set function $\varphi : \mathcal{D} \to [0,\infty)$ by (4.4.4) and proceeding in analogy to the proof of Theorem 4.4.3, we obtain a (possibly trivial) measure ρ on the Borel σ-algebra $\mathcal{B}(\mathcal{F})$. Lemma 4.4.4 yields that the resulting measure $\rho : \mathcal{B}(\mathcal{F}) \to [0,+\infty]$ is regular. Moreover, the useful results Lemma 4.4.10 and Lemma 4.4.11 still apply.

In analogy to Remark 4.4.5, the following remark yields a sufficient condition for the measure ρ obtained in Theorem 4.4.3 to be non-zero.

REMARK 4.6.3. *Let $(\mathcal{F}^{(n)})_{n\in\mathbb{N}}$ be a finite-dimensional approximation of $\mathcal{F}$ (for details see §4.3.3). Assuming that the Lagrangian is bounded and vanishes in entropy (see Definition 4.6.2), for every $x^{(n)} \in \mathcal{F}^{(n)}$ and $0 < \varepsilon < 1$ there exists $K^{(n)}_{x,\varepsilon} \subset \mathcal{F}^{(n)}$ compact such that*

$$\int_{\mathcal{F}\setminus K^{(n)}_{x,\varepsilon}} \mathcal{L}(x^{(n)},y)\, d\rho^{(n)}(y) < \varepsilon\,.$$

In view of boundedness of the Lagrangian we introduce the upper bound $\mathfrak{C} < \infty$ by

$$\mathfrak{C} := \sup_{x,y\in\mathcal{F}} \mathcal{L}(x,y) > 0\,.$$

Then the EL equations (4.3.4) *and* (4.3.5) *yield*

$$1 \leq \int_{\mathcal{F}} \mathcal{L}(x^{(n)},y)\, d\rho^{(n)} = \int_{K^{(n)}_{x,\varepsilon}} \mathcal{L}(x^{(n)},y)\, d\rho^{(n)} + \int_{\mathcal{F}\setminus K^{(n)}_{x,\varepsilon}} \mathcal{L}(x^{(n)},y)\, d\rho^{(n)}\,,$$

implying that

$$0 < \frac{1-\varepsilon}{\mathfrak{C}} \leq \rho^{(n)}(K^{(n)}_{x,\varepsilon}) \qquad \textit{for sufficiently large } n \in \mathbb{N}\,.$$

Without loss of generality we may assume that $K^{(n)}_{x,\varepsilon} \in \mathcal{D}$ for every $n \in \mathbb{N}$. Moreover, we are given $K_{x,\varepsilon}(N) := \bigcup_{n=1}^{N} K^{(n)}_{x,\varepsilon} \in \mathcal{D}$ for every $N \in \mathbb{N}$. Therefore, whenever there exists $N \in \mathbb{N}$ such that $\rho^{(n)}(K_{x,\varepsilon}(N)) \geq c$ for almost all $n \in \mathbb{N}$ and some $c > 0$, the measure ρ as defined by (4.4.6) *is non-zero. If this holds true for an infinite number of disjoints sets $(K_{x_i,\varepsilon}(N_i))_{i\in\mathbb{N}}$, the measure ρ possibly has infinite total volume.*

The remainder of this section is devoted to the proof that the measure ρ defined by (4.4.6) is, under suitable assumptions, a minimizer under variations of compact support as well as under variations of finite volume. For non-trivial minimizers, we shall derive the corresponding EL equations (see §4.6.4).

4.6.3. Existence of Minimizers. The aim of this subsection is to prove that, under suitable assumptions, the measure ρ defined by (4.4.6) is a minimizer of the causal variational principle (4.3.2), (4.2.3) under variations of finite volume (see Definition 4.3.5). To this end, we first show that ρ is a minimizer of the causal action under variations of compact support (see Definition 4.3.6).

In order to show that the measure ρ obtained in Theorem 4.4.3 is a minimizer under variations of compact support, we need to assume that condition (iv) in §4.2.2 holds (cf. [**66**, §5.4]), i.e.

$$\sup_{x \in \mathcal{F}} \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) < \infty\,.$$

Under the additional assumption that the measure ρ obtained in Theorem 4.4.3 also satisfies condition (B) in Section 4.5, we obtain the following existence result.

LEMMA 4.6.4. *Assume that the Lagrangian $\mathcal{L} \in C_b(\mathcal{F} \times \mathcal{F}; \mathbb{R}_0^+)$ vanishes in entropy, and that condition* (4.3.6) *holds. Furthermore, assume that the measure ρ obtained in* (4.4.6) *satisfies condition* (B) *in Section 4.5, and that condition* (iv) *in §4.2.2 holds. Then ρ is a minimizer under variations of compact support.*

PROOF. Assume that $\tilde{\rho} : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ is a regular Borel measure satisfying (4.2.3) such that $K := \text{supp}(\tilde{\rho} - \rho)$ is a compact subset of $\mathcal{F}$. Then the difference of actions (4.2.4) as given by

$$\begin{aligned}(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)) = 2 \int_K d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_K d(\tilde{\rho}-\rho)(x) \int_K d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y)\end{aligned}$$

is well-defined (see the explanations in [**66**, §4.3]). Assuming that condition (iv) in §4.2.2 holds, for any $x \in \mathcal{F}$ and $\tilde{\varepsilon} > 0$ there exists an integer $R = R(x, \tilde{\varepsilon})$ such that

$$\int_{\mathcal{F} \setminus B_R(x)} \mathcal{L}(x,y)\, d\rho(y) < \tilde{\varepsilon}/2\,.$$

By continuity of the Lagrangian, there is an open neighborhood U_x of x such that

$$\int_{\mathcal{F} \setminus B_R(x)} \mathcal{L}(z,y)\, d\rho(y) < \tilde{\varepsilon} \qquad \text{for all } z \in U_x\,.$$

Proceeding in analogy to the proof of [**66**, Theorem 5.8] by covering the compact set $K \subset \mathcal{F}$ by a finite number of such neighborhoods $U_{x_1}, \ldots, U_{x_L}$ and introducing the bounded set $B_K := \bigcup_{j=1}^L B_R(x_j)$, we conclude that

$$\int_{\mathcal{F} \setminus B_K} \mathcal{L}(x,y)\, d\rho(y) < \tilde{\varepsilon} \qquad \text{for all } x \in K\,.$$

This implies that, by choosing $\tilde{\varepsilon} > 0$ suitably, the last summand in the expression

$$\begin{aligned}(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)) = \Big[2 \int_K d(\tilde{\rho}-\rho)(x) \int_{B_K} d\rho(y)\, \mathcal{L}(x,y) \\ + \int_K d(\tilde{\rho}-\rho)(x) \int_K d(\tilde{\rho}-\rho)(y)\, \mathcal{L}(x,y) \Big] + 2 \int_K d(\tilde{\rho}-\rho)(x) \int_{\mathcal{F} \setminus B_K} d\rho(y)\, \mathcal{L}(x,y)\end{aligned}$$

is arbitrarily small. For this reason, it remains to consider the term in square brackets in more detail. Combining the facts that ρ satisfies condition (B) in Section 4.5 and

that $B_K \subset \mathcal{F}$ is bounded, Lemma 4.5.1 implies that $\rho(B_K) < \infty$. Proceeding similarly to the proof of Lemma 4.5.5, we deduce that the term in square brackets is bigger than or equal to zero, up to an arbitrarily small error term. This gives rise to

$$(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)) \geq 0\,,$$

which proves the claim. □

Proceeding similarly to the proof of Lemma 4.6.4, we obtain the following result.

THEOREM 4.6.5. *Assume that the Lagrangian $\mathcal{L} \in C_b(\mathcal{F}\times\mathcal{F};\mathbb{R}_0^+)$ vanishes in entropy, and that condition* (4.3.6) *holds. Moreover, assume that the measure ρ obtained in* (4.4.6) *satisfies condition* (B) *in Section 4.5, and that condition* (iv) *in §4.2.2 holds. Then ρ is a minimizer under variations of finite volume.*

PROOF. Assume that $\tilde{\rho} : \mathcal{B}(\mathcal{F}) \to [0,\infty]$ is a regular measure satisfying (4.2.3). By virtue of Lemma 4.5.1, the measure ρ is locally finite, implying that $\tilde{\rho}$ is also a locally finite measure (see the explanations after Definition 4.3.5). Introducing $B := \operatorname{supp}(\tilde{\rho}-\rho)$, we are given $\rho(B) = \tilde{\rho}(B) < \infty$. Since condition (iv) in §4.2.2 holds, the difference of actions (4.2.4) is well-defined and given by

$$\begin{aligned}(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)) = 2\int_B d(\tilde{\rho}-\rho)(x)\int_{\mathcal{F}} d\rho(y)\,\mathcal{L}(x,y)\\ + \int_B d(\tilde{\rho}-\rho)(x)\int_B d(\tilde{\rho}-\rho)(y)\,\mathcal{L}(x,y)\,.\end{aligned}$$

Making use of regularity of ρ and $\tilde{\rho}$, for arbitrary $\tilde{\varepsilon} > 0$ there is $U \supset B$ open such that

$$\rho(U \setminus B) < \tilde{\varepsilon} \qquad \text{and} \qquad \tilde{\rho}(U \setminus B) < \tilde{\varepsilon}\,.$$

Approximating U from inside by compact sets K such that

$$\rho(U \setminus K) < \tilde{\varepsilon} \qquad \text{and} \qquad \tilde{\rho}(U \setminus K) < \tilde{\varepsilon}$$

and proceeding in analogy to the proof of [**66**, Theorem 5.9] and Lemma 4.6.4, we finally may deduce that

$$(\mathcal{S}(\tilde{\rho}) - \mathcal{S}(\rho)) \geq 0\,,$$

which proves the claim. □

Theorem 4.6.5 concludes the existence theory in the non-locally compact setting.

4.6.4. Derivation of the Euler-Lagrange Equations. Under the assumptions of Theorem 4.6.5, for non-trivial measures $\rho \neq 0$ we are able to deduce the corresponding Euler-Lagrange equations. More precisely, in analogy to [**65**, Lemma 2.3] we obtain the following result.

THEOREM 4.6.6. *Assume that the Lagrangian $\mathcal{L} \in C_b(\mathcal{F}\times\mathcal{F};\mathbb{R}_0^+)$ vanishes in entropy (see Definition 4.6.2), and that condition* (4.3.6) *holds. Moreover, assume that the regular measure ρ obtained in* (4.4.6) *is non-zero and satisfies condition* (B) *in Section 4.5 as well as condition* (iv) *in §4.2.2. Then the following Euler-Lagrange equations hold,*

$$\ell|_{\operatorname{supp}\rho} \equiv \inf_{x\in\mathcal{F}} \ell(x) = 0\,, \tag{4.6.2}$$

where $\ell \in C(\mathcal{F})$ is defined by (4.5.4) *for a suitable parameter $\mathfrak{s} \in \mathbb{R}_0^+$.*

PROOF. Under the assumptions of Theorem 4.6.6, the measure ρ constructed in (4.4.6) is locally finite in view of Lemma 4.4.8 and a minimizer under variations of finite volume. Assuming that $\rho \neq 0$ and arguing similarly to the proof of Lemma 4.5.8, Theorem 4.5.7 gives rise to (4.6.2). □

Theorem 4.6.6 generalizes the results of [**66**, Section 5] to the infinite-dimensional setting. It remains an open task to prove the existence of Lagrangians vanishing in entropy such that the measure ρ given by (4.4.6) is non-zero and satisfies condition (iv) in §4.2.2 as well as condition (B) in Section 4.5.

4.7. Topological Properties of Spacetime

The goal of this section is to derive topological properties of spacetime and to work out a connection to dimension theory. To this end, we let $\mathcal{F}$ be a non-locally compact Polish space in the non-locally compact setting (see Definition 4.3.4). Under suitable assumptions on the Lagrangian (see Theorem 4.5.6 and Theorem 4.6.5), the measure ρ obtained in (4.4.6) is a minimizer of the corresponding variational principle (4.3.2). In order to obtain dimension-theoretical statements on its support, let us first recall some basic results from dimension theory (§4.7.1). Afterwards we specialize the setting by applying the obtained results to causal fermion systems (§4.7.2).

4.7.1. Dimension-Theoretical Preliminaries. To begin with, let us first point out that there are several notions of "dimension" of a topological space, among them the *small inductive dimension* ind, the *large inductive dimension* Ind, the *covering dimension* dim, the *Hausdorff dimension* $\dim_H$ and the *metric dimension* $\mu\dim$ (for details we refer to [**7**], [**42**], [**91**] and [**122**]). For a separable metric space X, the relation

$$\dim X \leq \dim_H X \tag{4.7.1}$$

holds in view of [**91**, Section VII.4]. Moreover, for every separable metrizable space X we have $\operatorname{ind} X = \operatorname{Ind} X = \dim X$ (see [**42**, Theorem 4.1.5]), and $\mu\dim Y = \dim Y$ for every compact metric space Y (see e.g. [**7**]).

For a metric space X, the *local dimension* $\dim_{\text{loc}} : X \to [0, \infty]$ is given by

$$\dim_{\text{loc}}(x) = \inf \{\dim_H(B_\varepsilon(x)) : \varepsilon > 0\}$$

(see [**27**, §2]), where

$$\dim_H(A) = \inf \{s \geq 0 : H^s(A) = 0\}$$

is the Hausdorff dimension of $A \subset X$, and H^s is the s-dimensional Hausdorff measure (the interested reader is referred to [**43**, Section 2.10], [**91**, Chapter VII] and [**137**]). Whenever X is a separable metric space, then one can show that

$$\dim_H(X) = \sup_{x \in X} \dim_{\text{loc}}(x)\,.$$

Moreover, if X is compact then the supremum is attained (cf. [**27**, Proposition 2.7]).

DEFINITION 4.7.1. *A normal space X is* locally finite-dimensional *if for every $x \in X$ there exists a normal open subspace U of X such that $x \in U$ and $\dim U < \infty$. See* [**42**, Section 5.5].

In order to apply the above preliminaries to minimizers of the causal variational principle, let us summarize some general topological properties of the support of a locally finite measure μ on a Polish space $\mathcal{F}$ in the next statement.

LEMMA 4.7.2. *Let X be a Polish space, and assume that μ is a locally finite measure on $\mathcal{B}(X)$. Then* $\operatorname{supp}\mu \subset X$ *is σ-compact, and there exists a locally finite-dimensional subspace F being dense in* $\operatorname{supp}\mu$. *Whenever* $\operatorname{supp}\mu$ *is hemicompact, then* $\operatorname{supp}\mu$ *is locally compact and thus locally finite-dimensional. Moreover, in the latter case there exists a (Heine-Borel) metric on* $\operatorname{supp}\mu$ *such that each bounded subset in* $\operatorname{supp}\mu$ *is finite-dimensional.*

PROOF. According to Lemma A.3.2, $\operatorname{supp}\mu$ is a σ-compact separable metric space, and there is a dense subset $F \subset \operatorname{supp}\mu$ such that each $x \in F$ is contained in a compact neighborhood N_x. In view of [**27**, Proposition 2.7] we conclude that $\dim_H N_x < \infty$ for every $x \in \operatorname{supp}\mu$. Thus Definition 4.7.1 together with (4.7.1) gives the first statement.

Whenever $\operatorname{supp}\mu$ is hemicompact, it is locally compact according to Lemma A.3.2. Thus each $x \in \operatorname{supp}\mu$ is contained in a compact neighborhood N_x. Making use of [**27**, Proposition 2.7] we deduce that $\dim_H N_x < \infty$. Since $x \in \operatorname{supp}\mu$ is arbitrary and the interior of N_x is open, from Definition 4.7.1 we obtain that $\operatorname{supp}\mu$ is locally finite-dimensional. Moreover, due to Lemma A.3.2, the space $\operatorname{supp}\mu$ can be endowed with a Heine-Borel metric. Accordingly, whenever $B \subset \operatorname{supp}\mu$ is bounded (with respect to the Heine-Borel metric), its closure is compact. Covering the resulting compact set by a finite number of compact neighborhoods, we conclude that $\dim_H(B) < \infty$. □

4.7.2. Application to Causal Fermion Systems. In the remainder of this section, we finally apply the previous results to the case of causal fermion systems. To this end, let $\mathcal{H}$ be an infinite-dimensional separable complex Hilbert space. For a given spin dimension $n \in \mathbb{N}$, the set $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ (for details see [**59**, Definition 1.1.1]) is a non-locally compact Polish space (see Theorem A.2.1 and Lemma 4.3.3). Assuming that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is symmetric, lower semi-continuous and strictly positive on the diagonal (4.3.1), we are exactly in the non-locally compact setting as introduced in §4.3.2 (see Definition 4.3.4). Therefore, by virtue of Theorem 4.4.3, there exists a regular measure $\rho : \mathcal{B}(\mathcal{F}) \to [0, \infty]$. Assuming in addition that the Lagrangian is continuous, bounded and of bounded range and that condition (B) in Section 4.5 is satisfied, the measure ρ is a minimizer of the causal variational principle (4.3.2), (4.2.3) under variations of compact support by virtue of Lemma 4.5.5. Under the additional assumption that condition (iv) in §4.2.2 is satisfied, the measure ρ is a minimizer of the causal variational principle under variations of finite volume due to Theorem 4.5.6. Under these assumptions, the same is true for Lagrangians vanishing in entropy (see Theorem 4.6.4 and Theorem 4.6.5). As a consequence, we are given a causal fermion system $(\mathcal{H}, \mathcal{F}, \rho)$, and *spacetime* M is defined as the support of the universal measure ρ,

$$M := \operatorname{supp}\rho\,.$$

Combining the results of Lemma 4.7.2 and Lemma A.3.2, we arrive at the following main results of this section.

THEOREM 4.7.3. *Assume that $\mathcal{L} \in C_b(\mathcal{F} \times \mathcal{F}; \mathbb{R}_0^+)$ is of bounded range or vanishes in entropy. Moreover, assume that the measure ρ defined by* (4.4.6) *satisfies condition* (B) *in Section 4.5. Then spacetime M is σ-compact and contains a locally finite-dimensional dense subspace. Under the additional assumption that spacetime M is hemicompact, it is a locally finite-dimensional, σ-locally compact Polish space.*

PROOF. Assuming that condition (B) in Section 4.5 is satisfied, the measure ρ is locally finite in view of Lemma 4.5.1. Henceforth the statement is a consequence of Lemma A.3.2 and Lemma 4.7.2. □

In [**12**] the question is raised whether the support of minimizing measures always is compact. Theorem 4.7.3 indicates that the support should in general at least be σ-compact.

Under the assumption that the measure ρ is locally finite (for sufficient conditions see Lemma 4.4.6, Lemma 4.4.8 and Lemma 4.5.1), we obtain the following result.

THEOREM 4.7.4. *Assume that the measure $\rho : \mathcal{B}(\mathcal{F}) \to [0, \infty]$ given by* (4.4.6) *is locally finite. Then the interior of spacetime $M = \operatorname{supp} \rho$ is empty (in the topology of $\mathcal{F}$).*

PROOF. Assume that $M^\circ \neq \varnothing$ in the topology of $\mathcal{F}$. Then $U^{\text{reg}} := M^\circ \cap \mathcal{F}^{\text{reg}}$ is open in the relative topology. Since $\mathcal{F}^{\text{reg}}$ is a Banach manifold (see [**67**]), it can be covered by an atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ for some index set A (cf. [**169**, Chapter 73]). In particular, each $x \in U^{\text{reg}}$ is contained in some open set U_α, whose image $V_\alpha := \phi_\alpha(U_\alpha)$ is open in some infinite-dimensional Banach space X_α. From Lemma A.3.2 we know that there exists a dense subset $F \subset \operatorname{supp} \rho$ such that each $x \in F$ has a compact neighborhood. Given $x \in F$ and choosing a compact neighborhood $N_x \subset U_\alpha$ for some $\alpha \in A$, from the fact that the mapping ϕ_α is a homeomorphism we conclude that $\phi_\alpha(N_x) \subset X_\alpha$ is a compact neighborhood of $\phi_\alpha(x) \in X_\alpha$ which contains a non-empty open subset in contradiction to [**100**, Exercise 14.3]. This gives the claim. □

Theorem 4.7.4 generalizes [**68**, Theorem 3.16] to the infinite-dimensional setting.

CHAPTER 5

Causal Variational Principles in the Homogeneous Setting

ABSTRACT. We derive Euler-Lagrange equations for causal variational principles in the homogeneous setting for minimizers in the class of strictly negative definite measures in the strong sense. Our method is to proceed in several steps: Under the assumption that the Lagrangian is continuously differentiable, the initial step is to compute the first variation of the causal action. Afterwards, we rewrite the first variation of the causal action in terms of momentum space. Analyzing the resulting expressions in terms of auxiliary variational principles, we then derive a convenient representation of the first variation of the causal action in terms of Lagrange multipliers. By restricting attention to appropriate variations of minimizers in the class of strictly negative definite measures in the strong sense, we finally deduce the desired Euler-Lagrange equations in momentum space.

5.1. Introduction

In the physical theory of causal fermion systems, spacetime and the structures therein are described by a minimizer of the so-called causal action principle (for an introduction to the physical background and the mathematical context, we refer the interested reader to Section 5.2, the textbook [**59**], the survey articles [**62, 64**] as well as the web platform [**1**]). Given a causal fermion system $(\mathcal{H}, \mathcal{F}, d\rho)$ together with a non-negative function

$$\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+ := [0, \infty)$$

(the *Lagrangian*), the causal action principle is to minimize the *action* $\mathcal{S}$ defined as the double integral over the Lagrangian

$$\mathcal{S}(\rho) = \int_{\mathcal{F}} d\rho(x) \int_{\mathcal{F}} d\rho(y)\, \mathcal{L}(x, y)$$

under variations of the measure $d\rho$ within the class of regular Borel measures on $\mathcal{F}$ under suitable side conditions. *Causal variational principles* evolved as a mathematical generalization of the causal action principle [**56, 65**]. The existence theory for minimizers of causal variational principles was developed in [**66, 110**], and the Euler-Lagrange (EL) equations in position space were derived. In the homogeneous setting, however, which is of particular physical interest by establishing a close connection to momentum space, neither the existence theory is developed nor the EL equations are derived. The aim of this chapter is to derive the EL equations in momentum space for causal variational principles in the *homogeneous* setting.

In order to put the chapter into the mathematical context, in [**53**] it was proposed to formulate physics by minimizing a new type of variational principle in spacetime. The suggestion in [**53**, Section 3.5] led to the causal action principle in discrete spacetime, which was first analyzed mathematically in [**54**]. A more general and systematic inquiry of causal variational principles on measure spaces was carried out in [**56**]. In [**56**, Section 3]

the existence of minimizers for variational principles in indefinite inner product spaces is proven in the special case that the total spacetime volume as well as the number of particles therein are finite. Under the additional assumption that the kernel of the fermionic projector is *homogeneous* in the sense that it only depends on the difference of two spacetime points, variational principles for homogeneous systems were considered in [**56**, Section 4] in order to deal with an infinite number of particles in an infinite spacetime volume. More precisely, the main advantage in the homogeneous setting is that it allows for Fourier methods, thus giving rise to a natural correspondence between position and momentum space. As a consequence, one is led to minimize the causal action by varying in the class of negative definite measures, and the existence of minimizers on bounded subsets of momentum space is proven in [**56**, Theorem 4.2].

Unfortunately, restricting attention to bounded subsets amounts to introducing a "cutoff in momentum space" which seems artificial and should be avoided. Accordingly, the aim is to develop the corresponding existence theory on whole momentum space. Roughly speaking, a possible strategy in the homogeneous setting is to proceed in the spirit of [**66**] by exhausting momentum space by a sequence of compact subsets; then the existence result in [**56**] yields minimizing measures on each compact subset. Having accomplished to derive the EL equations in momentum space, which is precisely the objective of this chapter, by suitably rescaling the resulting EL equations one can hope for constructing global minimizers. Further investigations regarding the existence theory of minimizers in the homogeneous setting shall be postponed to future projects.

The chapter is organized as follows. In Section 5.2, we first recall some fundamental definitions concerning indefinite inner product spaces and settle the notation (§5.2.1). Next, we give a short physical motivation (§5.2.2) and outline variational principles in infinite spacetime volume (§5.2.3). In Section 5.3 we introduce causal variational principles in the homogeneous setting. We first consider negative definite measures on whole momentum space (§5.3.1) and then introduce the more general class of operator-valued measures (§5.3.2). This allows us to state causal variational principles in the homogeneous setting with appropriate side conditions in sufficient generality in the context of calculus of variations (§5.3.3). In Section 5.4 we show that minimizing sequences are bounded. Unfortunately, it seems not possible to also infer that minimizing sequences are uniformly tight. For this reason, the remainder of the chapter is devoted to derive suitable EL equations in momentum space. The aim of Section 5.5 is to investigate the first variation of the causal action in the homogeneous setting. To this end, we first prove that the closed chain is Fréchet differentiable (§5.5.1). Subsequently, we deduce the first variation of the causal action under the crucial assumption that the Lagrangian is a continuously differentiable mapping (§5.5.2). We then prove some technical results (§5.5.3) in order to rewrite the first variation of the causal action in terms of momentum space (§5.5.4). In Section 5.6, the resulting expressions for the first variation of the causal action are examined in more detail. We start off with a few preliminary notes (§5.6.1); analyzing the relevant expressions in terms of auxiliary variational principles (§5.6.2) provides a convenient representation in terms of Lagrange multipliers (§5.6.3). In Section 5.7, we derive the EL equations for causal variational principles in the homogeneous setting. To this end, we first compile general assumptions (§5.7.1) and afterwards we introduce *strictly* negative definite measures *in the strong sense* (§5.7.2). Next, we consider appropriate variations of strictly negative definite minimizers in the strong sense of the causal variational principle with fixed support (§5.7.3) as well as variations which modify the support (§5.7.4). As a consequence, this allows us to deduce the main result of

this chapter (Theorem 5.7.10). In Section 5.8, we furnish the connection to homogeneous causal fermion systems. Following [**109**, Chapter 3], we first recall in which way a negative definite measure on momentum space gives rise to a corresponding causal fermion system (§5.8.1). Afterwards we put our results into the physical context (§5.8.2). In the appendix, we finally explain and justify the side conditions of our variational principle (Appendix A.4).

5.2. Physical Background and Mathematical Preliminaries

5.2.1. Mathematical Preliminaries and Notation. To begin with, let us compile some fundamental definitions which are of central relevance throughout this chapter. For details we refer the interested reader to [**17**], [**79**] and [**111**]. Unless specified otherwise, we always let $n \geq 1$ be a given integer.

DEFINITION 5.2.1. *A mapping $\prec . \mid . \succ : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ is called an **indefinite inner product** if the following conditions hold (cf.* [**79**, Definition 2.1]*):*

(i) $\prec y \mid \alpha x_1 + \beta x_2 \succ = \alpha \prec y \mid x_1 \succ + \beta \prec y \mid x_2 \succ$ *for all x_1, x_2, $y \in \mathbb{C}^n$, α, $\beta \in \mathbb{C}$.*
(ii) $\prec x \mid y \succ = \overline{\prec y \mid x \succ}$ *for all x, $y \in \mathbb{C}^n$.*
(iii) $\prec x \mid y \succ = 0$ *for all $y \in \mathbb{C}^n \Longrightarrow x = 0$.*

DEFINITION 5.2.2. *Let V be a finite-dimensional complex vector space, endowed with an indefinite inner product $\prec . \mid . \succ$. Then $(V, \prec . \mid . \succ)$ is called an **indefinite inner product space**.*

As usual, by $\mathrm{L}(V)$ we denote the set of (bounded) linear operators on a complex (finite-dimensional) vector space V of dimension $n \in \mathbb{N}$. The adjoint of $A \in \mathrm{L}(V)$ with respect to the Euclidean inner product $\langle . \mid . \rangle_{\mathbb{C}^n}$ on $V \simeq \mathbb{C}^n$ is denoted by $A^\dagger$. On the other hand, whenever $(V, \prec . \mid . \succ)$ is an indefinite inner product space, unitary matrices and the adjoint A^* (with respect to $\prec . \mid . \succ$) are defined as follows.

DEFINITION 5.2.3. *Let $\prec . \mid . \succ$ be an indefinite inner product on $V \simeq \mathbb{C}^n$, and let S be the associated invertible hermitian matrix determined by* [**79**, eq. (2.1.1)]*,*

$$\prec x \mid y \succ = \langle S\, x \mid y \rangle_{\mathbb{C}^n} \qquad \text{for all } x, y \in \mathbb{C}^n .$$

Then for every $A \in \mathrm{L}(V)$, the adjoint of A (with respect to $\prec . \mid . \succ$) is the unique matrix $A^* \in \mathrm{L}(V)$ which satisfies

$$\prec A\, x \mid y \succ = \prec x \mid A^* y \succ \qquad \text{for all } x, y \in V .$$

*A matrix $A \in \mathrm{L}(V)$ is called **self-adjoint (with respect to $\prec . \mid . \succ$)** if and only if $A = A^*$. In a similar fashion, an operator $U \in \mathrm{L}(V)$ is said to be **unitary (with respect to $\prec . \mid . \succ$)** if it is invertible and $U^{-1} = U^*$ (see* [**79**, Section 4.1]*).*

We remark that every non-negative matrix (with respect to $\prec . \mid . \succ$) is self-adjoint (with respect to $\prec . \mid . \succ$) and has a real spectrum (cf. [**79**, Theorem 5.7.2]). Moreover, the adjoint A^* of $A \in \mathrm{L}(V)$ satisfies the relation

$$A^* = S^{-1}\, A^\dagger\, S$$

in view of [**79**, eq. (4.1.3)] (where $A^\dagger$ denotes the adjoint with respect to $\langle . \mid . \rangle$ and A^* the adjoint with respect to $\prec . \mid . \succ$). For details concerning self-adjoint operators (with respect to $\prec . \mid . \succ$) we refer to [**111**] and the textbook [**17**].

In the remainder of this chapter we will restrict attention exclusively to indefinite inner product spaces $(V, \prec . \mid . \succ)$ with $V \simeq \mathbb{C}^{2n}$ for some $n \in \mathbb{N}$. It is convenient to

work with a fixed pseudo-orthonormal basis $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ of V in which the inner product has the standard representation with a signature matrix S,

$$\prec u \,|\, v \succ = \langle u \,|\, Sv \rangle \qquad \text{with} \qquad S = \text{diag}(\underbrace{1, \ldots, 1}_{n \text{ times}}, \underbrace{-1, \ldots, -1}_{n \text{ times}}) \,, \tag{5.2.1}$$

where $\langle \,.\,|\,.\,\rangle = \langle \,.\,|\,.\,\rangle_{\mathbb{C}^{2n}}$ denotes the standard inner product on $\mathbb{C}^{2n}$. The signature matrix can be regarded as an operator on V,

$$S = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \in \text{Symm}\, V \,, \tag{5.2.2}$$

where $\text{Symm}\, V$ denotes the set of symmetric matrices on V with respect to the spin scalar product (also cf. [**56**, proof of Lemma 3.4]). Without loss of generality we may assume that $\mathfrak{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^\mathsf{T}$ for all $i = 1, \ldots, 2n$.

Throughout this chapter, we denote Minkowski space by $\mathscr{M} \simeq \mathbb{R}^4$ and momentum space by $\hat{\mathscr{M}} \simeq \mathbb{R}^4$. Identifying $\hat{\mathscr{M}}$ with Minkowski space $\mathscr{M}$, the Minkowski inner product can be considered as a mapping

$$\langle .,. \rangle : \hat{\mathscr{M}} \times \mathscr{M} \to \mathbb{R} \,, \qquad (k, x) \mapsto \langle k, \xi \rangle = \eta_{\mu\nu}\, k^\mu\, \xi^\mu = k^0\, \xi^0 - \sum_{i=1}^3 k^i\, \xi^i$$

for all $\xi = (\xi^0, \xi^1, \xi^2, \xi^3) \in \mathscr{M}$ and $k = (k^0, k^1, k^2, k^3) \in \hat{\mathscr{M}}$ (with Minkowski metric η, where we employed Einstein's summation convention, cf. [**73**, Chapter 1]).

Moreover, by $\mathcal{B}(\mathscr{M})$ and $\mathcal{B}(\hat{\mathscr{M}})$ we denote the Borel σ-algebra on Minkowski and momentum space, respectively. The class of finite complex measures on $\hat{\mathscr{M}}$ is denoted by $\mathbf{M}_\mathbb{C}(\hat{\mathscr{M}})$. By $C_c(\hat{\mathscr{M}})$ we denote the set of continuous functions on $\hat{\mathscr{M}}$ with compact support, while $C_b(\hat{\mathscr{M}})$ and $C_0(\hat{\mathscr{M}})$ indicate the sets of continuous functions on $\hat{\mathscr{M}}$ which are bounded or vanish at infinity, respectively. Moreover, by $L^1_{\text{loc}}(\mathscr{M})$ we denote the set of locally integrable functions on $\mathscr{M}$ with respect to Lebesgue measure, denoted by $d\mu$. Unless otherwise specified, Borel measures are understood as locally finite measures on the Borel σ-algebra in the sense of [**76**]. A Borel measure is said to be regular if it is inner and outer regular. Inner regular Borel measures are referred to as Radon measures [**40**].

5.2.2. Physical Background and Motivation. The purpose of this subsection is to outline a few concepts of causal fermion systems and to explain how this chapter fits into the general physical context and the ongoing research program. The reader not interested in the physical background may skip this section.

To begin with, let us consider the Dirac equation in Minkowski space,

$$(i \partial\!\!\!/ - m)\, \psi = 0 \,,$$

where $\partial\!\!\!/ \equiv \partial_\mu \gamma^\mu$ denotes the so-called Feynman dagger, and $\gamma^\mu \in \mathbb{C}^{4\times 4}$ are the Dirac matrices with $\mu = 0, \ldots, 3$ (where we make use of the Einstein sum convention). Its solutions are called Dirac spinors, which are mappings $\psi : \mathscr{M} \to \mathbb{C}^4$, where $\mathscr{M}$ denotes Minkowski spacetime. Introducing the adjoint spinor $\overline{\psi} := \psi^\dagger \gamma^0$ (where the dagger denotes the adjoint with respect to the Euclidean scalar product $\langle \,.\,|\,.\,\rangle_{\mathbb{C}^4}$ on $\mathbb{C}^4$), setting $\prec \psi \,|\, \phi \succ \equiv \overline{\psi} \phi$ gives rise to an indefinite inner product of signature $(2,2)$ on $\mathbb{C}^4$ (for basic definitions see for instance [**79**]). For this reason, we refer to $\prec \,.\,|\,.\, \succ$ as *spin scalar product*. In a given reference frame, it is convenient to introduce a positive scalar product by

$$(\psi \,|\, \phi) := \int_{\mathbb{R}^3} \prec \psi(t, \vec{x}) \,|\, \gamma^0 \phi(t, \vec{x}) \succ \, d^3\vec{x} \,, \tag{5.2.3}$$

where $x = (t, \vec{x}) \in \mathbb{R} \times \mathbb{R}^3 \simeq \mathscr{M}$. The Hilbert space corresponding to $(\,.\,|\,.\,)$ as defined by (5.2.3) is given by $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C})^4$ (see [**53**, §1.2]).

After these quantum mechanical preliminaries, let us establish the connection to causal fermion systems. The theory of causal fermion systems is a recent approach to fundamental physics, originally motivated by the attempt to resolve shortcomings of relativistic quantum field theory. More precisely, the renormalization program is a method of getting along with infinite integrals in quantum field theory. Namely, due to ultraviolet divergences, perturbative quantum field theory is well-defined only after regularization, which is usually understood as a set of prescriptions for how to make divergent integrals finite. A common method is to introduce a suitable "cutoff" in momentum space, which often is associated to Planck energy $E_P \approx 1.22 \cdot 10^{28}$ eV. As a consequence of Heisenberg's uncertainty principle, the microscopic structure of spacetime for distances smaller than Planck length $\ell_P \approx 1.6 \cdot 10^{-35}$ m is completely unknown. Unfortunately, at present there is no consensus on what the correct mathematical model for "Planck scale physics" should be.

The simplest and maybe most natural approach is to assume that on the Planck scale, spacetime is no longer a continuum but becomes in some way "discrete." This is the starting point in the monograph [**53**]. The fermionic projector approach outlined in [**53**] is based on the belief that on a microscopic scale (like the Planck scale), spacetime should not be modelled by Minkowski space but should have a different, possibly discrete structure. Consequently, the Dirac equation in Minkowski space should not be considered as being fundamental, but it should be replaced by equations of different type. For such a more fundamental description, the scalar product (5.2.3) is problematic for the following reason: The scalar product (5.2.3) can be regarded as an integral over a Cauchy hypersurface in spacetime; however, in non-globally hyperbolic spacetimes Cauchy surfaces have no mathematical meaning.[1] A possible way out is to proceed as follows: For simplicity, one could still assume that spacetime looks like Minkowski space, but without taken the Dirac equation for granted. However, dropping the Dirac equation implies that we are no longer given current conservation on hypersurfaces. As a consequence, the integral (5.2.3) depends on the choice of the hypersurface, thus motivating to integrate over whole spacetime instead of hypersurfaces thereof (also cf. [**53**, preface to the second online edition]). Accordingly, integrating the spin scalar product $\prec \psi \,|\, \phi \succ$ for wave functions ψ, ϕ on Minkowski space $\mathscr{M}$ over whole Minkowski spacetime instead of Cauchy surfaces thereof gives rise to the expression (cf. [**53**, eq. (2.3.6)])

$$< \psi \,|\, \phi > := \int_{\mathscr{M}} \prec \psi(x) \,|\, \phi(x) \succ \, d\mu(x) \,, \tag{5.2.4}$$

where $d\mu$ denotes the Lebesgue measure on Minkowski space. As a result, the vector space H of all Dirac wave functions $\psi, \phi : \mathscr{M} \to \mathbb{C}^4$ endowed with the indefinite inner product (5.2.4) may be regarded as an (infinite-dimensional) indefinite inner product space $(H, < . \,|\, . >)$. In contrast to the inner product (5.2.3), which can be interpreted as the probability to detect the corresponding Dirac particle, the inner product (5.2.4) has no

[1] In general relativity, spacetime is described mathematically by a Lorentzian manifold (see [**8**]). As a fundamental example, we consider Minkowski spacetime endowed with the Minkowski inner product of signature $(+,-,-,-)$, thus giving rise to a Lorentzian manifold (see e.g. [**10**, Chapter 5] or [**123**]). Actually, it is a globally hyperbolic manifold (cf. [**125**, Chapter 14]). For this reason, it has the topology $\mathbb{R} \times \Sigma$ and hence can be foliated by a one-parameter family of smooth Cauchy surfaces (see for instance [**154**, Theorem 4.1.1]). In particular, the Dirac equation in Minkowski space admits unique global solutions if regular initial data are given on a fixed surface $\Sigma_t := \{(t, \mathbf{x}) \in \mathbb{R}^4 : \mathbf{x} \in \mathbb{R}^3\}$ (cf. [**126**, §2.1]).

immediate physical interpretation [**51**]. Nevertheless, the advantage of the indefinite inner product (5.2.4) is that it can easily be generalized for instance to a discrete spacetime if one simply replaces the integral in (5.2.4) by a sum over all spacetime points. These considerations suggest to regard the indefinite inner product (5.2.4) as being more fundamental than (5.2.3). Generalizing the above concept, the connection between indefinite inner product spaces and Hilbert spaces is established as follows: The above considerations motivate to consider an abstract indefinite inner product space $(H, <.|.>)$ as the starting point. Restricting attention to its negative definite subspace and taking the closure with respect to the indefinite inner product (5.2.4) gives rise to a Hilbert space $\mathcal{H}$ (this construction was first given in [**61**] and considered in more detail in [**109**]). The connection between the Hilbert space $\mathcal{H}$ and the underlying measure space $(M, d\mu)$ is outlined in more detail in [**53**, preface to the second online edition]. For the introduction of wave functions in the abstract setting we refer to [**59**, §1.1.4]. This procedure gives rise to causal fermion systems which are analyzed in detail in the textbook [**59**]. Concerning the construction of causal fermion systems in the homogeneous setting we refer to §5.8.1 below.

Let us finally point out the connection to the concept of the Dirac sea as introduced by Paul Dirac in his paper [**34**], where he assumes that

> "(...) *all the states of negative energy are occupied except perhaps a few of small velocity.* (...) *Only the small departure from exact uniformity, brought about by some of the negative-energy states being unoccupied, can we hope to observe.*"

Dirac made this picture precise in his paper [**35**] by introducing a relativistic density matrix $R(t, \vec{x}; t', \vec{x}')$ with $(t, \vec{x}), (t', \vec{x}') \in \mathbb{R} \times \mathbb{R}^3$ defined by

$$R(t, \vec{x}; t', \vec{x}') = \sum_{l \text{ occupied}} \Psi_l(t, \vec{x})\, \overline{\Psi_l(t', \vec{x}')}\,.$$

In analogy to Dirac's original idea, in [**50**] the kernel of the fermionic projector is introduced as the sum over all occupied wave functions

$$P(x, y) = -\sum_{l \text{ occupied}} \Psi_l(x)\, \overline{\Psi_l(y)}$$

for spacetime points $x, y \in \mathscr{M}$ as outlined in [**57**]. A straightforward calculation shows that (see e.g. [**60**, §4.1]) the kernel of the fermionic projector takes the form

$$P(x, y) = \int_{\hat{\mathscr{M}}} \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)}\,. \tag{5.2.5}$$

It is referred to as the (unregularized) *kernel of the fermionic projector of the vacuum* (cf. [**59**, eq. (1.2.20) and eq. (1.2.23)] as well as [**53**, eq. (4.1.1)]; this object already appears in [**48**]). We also refer to (5.2.5) as a *completely filled Dirac sea*.

5.2.3. Variational Principles in Infinite Spacetime Volume. The kernel of the fermionic projector (5.2.5) is the starting point for the analysis in [**56**, Section 4]. In order to deal with systems containing an infinite number of particles in an infinite spacetime volume, the main simplification in [**56**] is to assume that the kernel of the fermionic projector (5.2.5) is *homogeneous* in the sense that $P(x, y)$ only depends on the difference vector $y - x$ for all spacetime points $x, y \in \mathscr{M}$. As a consequence, for all $x, y \in \mathscr{M}$ the

fermionic projector (5.2.5) can be written as a Fourier transform,

$$P(x,y) = \int_{\hat{\mathscr{M}}} \frac{d^4k}{(2\pi)^4}\, \hat{P}(k)\, e^{i\langle k, y-x\rangle}\,. \tag{5.2.6}$$

We remark that in [**53**], the underlying homogeneity assumption $P(x,y) = P(y-x)$ for all $x, y \in \mathscr{M}$ is referred to as "homogeneous regularization of the vacuum" (cf. [**53**, eq. (4.1.2)] and the explanations thereafter; also see [**59**, Assumption 3.3.1]).

Thus in order to describe a completely filled Dirac sea in the vacuum (5.2.5), the Fourier transform $\hat{P}(k)$ takes the form

$$\hat{P}(k) = (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\,,$$

where δ denotes Dirac's delta distribution and Θ is the Heaviside function (for details we refer to [**55**, §6] and [**56**]). More generally, in order to obtain a suitable measure-theoretic framework, we regard $\hat{P}(k)\, d^4k/(2\pi)^4$ as a Borel measure $d\nu$ on $\hat{\mathscr{M}}$, taking values in $\mathrm{L}(V)$. In particular, the measure

$$d\nu(k) = (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, d^4k \tag{5.2.7}$$

has the remarkable property that $-d\nu$ is positive in the sense that

$$\prec u \,|\, -\nu(\Omega)\, u \succ \geq 0 \qquad \text{for all } u \in \mathbb{C}^4 \text{ and each } \Omega \in \mathcal{B}(\hat{\mathscr{M}}) \tag{5.2.8}$$

with respect to the spin scalar product $\prec . \,|\, . \succ$ on $\mathbb{C}^4$ introduced in §5.2.2.[2]

Generalizing $(\mathbb{C}^4, \prec . \,|\, . \succ)$ to some indefinite inner product space $(V, \prec . \,|\, . \succ)$ of dimension $2n$, the above observations motivate the following definition (see [**56**, Definition 4.1]):

DEFINITION 5.2.4. *Consider a regular Borel measure $d\nu$ on a bounded set $\hat{K} \subset \hat{\mathscr{M}}$ taking values in $\mathrm{L}(V)$ with the following properties:*

(i) *For every $u \in V$, the measure $d \prec u \,|\, \nu u \succ$ is a finite real measure.*
(ii) *For every Borel set $\Omega \subset \hat{K}$, the operator $-\nu(\Omega) \in \mathrm{L}(V)$ is positive* (5.2.8).

---

[2] In order to see this, we make use of the fact that the Dirac matrices anti-commute, i.e.

$$\gamma^\mu \gamma^\nu = -\gamma^\nu \gamma^\mu \qquad \text{whenever } \mu \neq \nu \text{ (where } \mu, \nu \in \{0, \ldots, 3\})\,.$$

Thus for every $k \in \hat{\mathscr{M}}$ with $k = (k^0, \vec{k})$, the operators $p_\pm(\vec{k})$ given by [**126**, eq. (2.13)] satisfy

$$p_\pm(\vec{k})\, \gamma^0 = \frac{\not{k} + m}{2k^0}\, \gamma^0|_{k^0 = \pm\omega(\vec{k})} = \gamma^0\, p_\pm(-\vec{k})$$

with $\omega(\vec{k}) = \sqrt{\vec{k}^2 + m^2}$. Applying the fact that $p_\pm(\vec{k})$ is idempotent and symmetric with respect to the Euclidean scalar product $\langle . \,|\, . \rangle_{\mathbb{C}^4}$ on $\mathbb{C}^4$ (cf. [**126**, Proposition 2.14]), the calculation

$$\prec u \,|\, (\not{k} + m)\, u \succ = 2k^0 \prec u \,|\, \gamma^0 p_\pm(-\vec{k})\, u \succ = 2k^0 \langle u \,|\, p_\pm(-\vec{k})^2\, u \rangle_{\mathbb{C}^4} = 2k^0 \langle p_\pm(-\vec{k})\, u \,|\, p_\pm(-\vec{k})\, u \rangle_{\mathbb{C}^4}$$

for any $u \in \mathbb{C}^4$ implies that

$$\prec \cdot \,|\, (\not{k} + m) \cdot \succ \quad \text{is} \quad \begin{cases} \text{positive semidefinite} & \text{if } k^0 > 0 \\ \text{negative semidefinite} & \text{if } k^0 < 0\,. \end{cases}$$

Introducing the set $\Omega^- = \Omega \cap \{k^0 < 0 : k = (k^0, \vec{k}) \in \hat{\mathscr{M}}\}$ for any $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$, for all $u \in V$ we obtain

$$\begin{aligned} \prec u \,|\, -\nu(\Omega)\, u \succ &= \prec u \,|\, -\int_\Omega (k_j\, \gamma^j + m)\, \delta\left(\langle k, k\rangle - m^2\right)\, \Theta(-k^0)\, d^4k\, u \succ \\ &= \int_{\Omega^-} \underbrace{-\prec u \,|\, (\not{k} + m) u \succ}_{\geq\, 0}\, \delta\left(\langle k, k\rangle - m^2\right)\, d^4k \geq 0\,. \end{aligned}$$

Therefore, positivity (5.2.8) is a consequence of the corresponding behavior of the operator $(\not{k} + m)$.

Then $d\nu$ is called a ***negative definite measure*** *on $\hat{K}$ with values in* $\mathrm{L}(V)$.

In terms of a negative definite measure $d\nu$, the *kernel of the fermionic projector* is then introduced by

$$P(\xi) := \int_{\hat{K}} e^{i\langle p,\xi\rangle}\, d\nu(p)\,, \qquad \text{where } \xi := y - x \text{ for all } x, y \in \mathscr{M}\,.$$

Following the procedure in [**56**, Section 3], for every $\xi \in \mathscr{M}$ the *closed chain* is defined by $A(\xi) := P(\xi)\, P(-\xi)$. According to [**56**, eq. (3.7)], the spectral weight $|A|$ of an operator $A \in \mathrm{L}(V)$ is given by the sum of the absolute value of the eigenvalues of A,

$$|A| = \sum_{i=1}^{2n} |\lambda_i|\,,$$

where λ_i denote the eigenvalues of A, counted with algebraic multiplicities. In analogy to [**56**, eq. (3.8)], for every $\xi \in \mathscr{M}$ the Lagrangian is introduced via

$$\mathcal{L}[A(\xi)] := |A(\xi)^2| - \frac{1}{2n} |A(\xi)|^2\,.$$

Defining the functionals $\mathcal{S}$ and $\mathcal{T}$ by [**56**, eq. (4.5)],

$$\mathcal{S}[\nu] := \int_{\mathscr{M}} \mathcal{L}[A(\xi)]\, d\mu(\xi)\,, \qquad \mathcal{T}[\nu] := \int_{\mathscr{M}} |A(\xi)|^2\, d\mu(\xi)\,,$$

the main result in [**56**, Section 4] can be stated as follows (see [**56**, Theorem 4.2]):

THEOREM 5.2.5. *Let $(d\nu_k)_{k\in\mathbb{N}}$ be a sequence of negative definite measures on the bounded set $\hat{K} \subset \hat{\mathscr{M}}$ such that the functional $\mathcal{T}$ is bounded by some constant $C > 0$, i.e.*

$$\mathcal{T}[\nu_k] \leq C \qquad \text{for all } k \in \mathbb{N}\,.$$

Then there is a subsequence $(d\nu_{k_\ell})_{\ell\in\mathbb{N}}$ as well as a sequence of unitary transformations $(U_\ell)_{\ell\in\mathbb{N}}$ on V (with respect to $\prec . \,|\, . \succ$) such that the measures $U_\ell\, d\nu_{k_\ell}\, U_\ell^{-1}$ converge weakly to a negative definite measure $d\nu$ with the properties

$$\mathcal{T}[\nu] \leq \liminf_{k\to\infty} \mathcal{T}[\nu_k]\,, \qquad \mathcal{S}[\nu] \leq \liminf_{k\to\infty} \mathcal{S}[\nu_k]\,.$$

Theorem 5.2.5 is stated as a compactness result. Applying it to a minimal sequence yields statements similar to [**56**, Theorem 2.2 and Theorem 2.3], asserting that the functional $\mathcal{S}$ attains its minimum. Restricting to bounded subsets $\hat{K} \subset \hat{\mathscr{M}}$, however, can be thought of as introducing a cutoff in momentum space which seems artificial and should be avoided. The first step towards this goal is to proceed in the spirit of [**66**] by exhausting the whole space $\hat{\mathscr{M}}$ by compact subsets and to essentially apply the EL equations on each compact subset. In a second step, one could hope for constructing a global minimizer.

5.3. Causal Variational Principles in the Homogeneous Setting

After the preliminaries stated in Section 5.2, we are now going to introduce causal variational principles in the homogeneous setting as considered in [**109**, Chapter 5]. Starting point in what follows is again the kernel of the fermionic projector (5.2.5), which for all spacetime points $x, y \in \mathscr{M}$ is given by

$$P(x,y) = \int_{\hat{\mathscr{M}}} \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)}\,.$$

In analogy to §5.2.3, we essentially assume that the kernel of the fermionic projector is *homogeneous* in the sense that $P(x,y)$ only depends on the difference vector $y-x$ for all spacetime points $x, y \in \mathscr{M}$. Following the reasoning in §5.2.3, the kernel of the fermionic projector can be written as Fourier transform (5.2.6). In particular, the measure given by (5.2.7) yields positivity (5.2.8).

5.3.1. Negative Definite Measures. We now generalize the indefinite inner product space $(\mathbb{C}^4, \prec . \mid . \succ)$ to indefinite inner product spaces $(V, \prec . \mid . \succ)$ of dimension $2n$. In generalization of Definition 5.2.4, we allow for negative definite measures on whole momentum space $\hat{\mathscr{M}}$ instead of bounded subsets thereof.

DEFINITION 5.3.1. *A vector-valued measure $d\nu$ on $\mathcal{B}(\hat{\mathscr{M}})$ taking values in $\mathrm{L}(V)$ is called a* ***negative definite measure*** *on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ if $d \prec u \mid -\nu\, u \succ$ is a positive finite Borel measure on $\mathcal{B}(\hat{\mathscr{M}})$ for every $u \in V$.*

Every negative definite measure $d\nu$ has by definition the property that, for any Borel set $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$,

$$\prec u \mid -\nu(\Omega)\, u \succ \geq 0 \qquad \text{for all } u \in V\,, \tag{5.3.1}$$

thus generalizing (5.2.8). Moreover, given a negative definite measure $d\nu$, the complex measure $d \prec u \mid \nu\, v \succ \in \mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ is defined by polarization for all $u, v \in V$,

$$\begin{aligned} d \prec u \mid \nu\, v \succ := \frac{1}{4}\big\{ & d \prec u+v \mid \nu\,(u+v) \succ + i\, d \prec u+iv \mid \nu\,(u+iv) \succ \\ & - d \prec u-v \mid \nu\,(u-v) \succ - i\, d \prec u-iv \mid \nu\,(u-iv) \succ \big\} \end{aligned} \tag{5.3.2}$$

(see e.g. [**79**, eq. (2.2.6)], also cf. [**132**, Section VIII.3]). Following [**109**, Definition A.16], we define integration with respect to negative definite measures as follows:

DEFINITION 5.3.2. *Let $(V, \prec . \mid . \succ)$ be an indefinite inner product space and let $d\nu$ be a negative definite measure. Moreover, let $f : \hat{\mathscr{M}} \to \mathbb{C}$ be a bounded Borel measurable function. For all $u, v \in V$, integration with respect to $d\nu$ is defined by*

$$\prec u \mid \left(\int_{\hat{\mathscr{M}}} f(k)\, d\nu(k) \right) v \succ := \int_{\hat{\mathscr{M}}} f(k)\, d \prec u \mid \nu(k)\, v \succ \;.$$

A similar definition in terms of operator-valued measures is stated below (see Definition 5.3.8). For a connection to spectral theory we refer to [**113**, Chapter 31].

Unless stated otherwise, in the remainder of this chapter we let $(V, \prec . \mid . \succ)$ be an indefinite inner product space of dimension $2n$ and signature (n, n) for some given integer $n \geq 1$. We consider Minkowski space $\mathscr{M} \simeq \mathbb{R}^4$ (also referred to as *position space*) as the group of translations in $\mathbb{R}^4$, and let $\hat{\mathscr{M}} \simeq \mathbb{R}^4$ be momentum space. We denote the set of finite real and complex measures on $\hat{\mathscr{M}}$ by $\mathbf{M}(\hat{\mathscr{M}})$ and $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$, respectively. Whenever V and $\hat{\mathscr{M}}$ are understood, the class of all negative definite measures on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ is denoted by $\mathfrak{Ndm}$. By $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ we always denote a pseudo-orthonormal basis of V satisfying (5.2.1).

5.3.2. Operator-Valued Measures. In order to deal with causal variational principles in the homogeneous setting in sufficient generality, this subsection is devoted to put the definition of negative definite measures (see Definition 5.3.1) into the context of calculus of variations. More precisely, as explained in §5.2.3, the variational principle as introduced in [**56**, Section 4] is to minimize the causal action $\mathcal{S}$ in the class of

negative definite measures. Unfortunately, in view of (5.3.1), the set of negative definite measures does not form a vector space, whereas in the calculus of variations one usually considers functionals on a real, locally convex vector space (for details we refer to [**168**, Section 43.2]). Hence in order to obtain a suitable framework, we first introduce operator-valued measures, which can be regarded as a generalization of negative definite measures, thus providing the basic structures required for the calculus of variations (see Lemma 5.3.5 below). Concerning the connection to vector-valued measures we refer to [**29**].

Then operator-valued measures on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ are introduced as a generalization of negative definite measures (see Definition 5.3.1) in the following way:

DEFINITION 5.3.3. *A vector-valued measure $d\omega$ on $\mathcal{B}(\hat{\mathscr{M}})$ taking values in $\mathrm{L}(V)$ is called an* ***operator-valued measure*** *on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ if $d\prec u \,|\, \omega\, v \succ$ is a finite complex measure in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ for all $u, v \in V$.*

Whenever $\hat{\mathscr{M}}$ and V are understood, the class of operator-valued measures on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ shall be denoted by $\mathfrak{Ovm}$.

In what follows, the variation of an operator-valued measure plays a central role:

DEFINITION 5.3.4. *Given an operator-valued measure $d\omega \in \mathfrak{Ovm}$, the* ***variation*** *of $d\omega$, denoted by $d|\omega|$, is defined by*

$$d|\omega| := \sum_{i,j=1}^{2n} d\,|\prec \mathfrak{e}_i \,|\, \omega\, \mathfrak{e}_j \succ| \;,$$

where $d\,|\,.\,|$ denotes the variation of a complex measure. Moreover, the ***total variation*** *of $d\omega$, denoted by $d\|\omega\|$, is given by*

$$d\|\omega\| := d|\omega|(\hat{\mathscr{M}}) = \sum_{i,j=1}^{2n} d\,|\prec \mathfrak{e}_i \,|\, \omega\, \mathfrak{e}_j \succ|\,(\hat{\mathscr{M}})\,. \tag{5.3.3}$$

We point out that the variation as given by Definition 5.3.4 crucially depends on the pseudo-orthogonal $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ basis of V. Nevertheless, the set of operator-valued measures $\mathfrak{Ovm}$ is a Banach space with respect to the total variation:

LEMMA 5.3.5. *The total variation $d\|\cdot\|$ given by* (5.3.3) *defines a norm on $\mathfrak{Ovm}$ in such a way that $(\mathfrak{Ovm}, d\|\cdot\|)$ is a complex Banach space. In particular, $(\mathfrak{Ovm}, d\|\cdot\|)$ is a real, locally convex vector space.*

PROOF. For the first part of the statement see the proof of [**109**, Corollary 5.3]. In order to show that $\mathfrak{Ovm}$ is a Banach space, let us consider a Cauchy sequence of operator-valued measures $(d\omega_k)_{k\in\mathbb{N}}$ with respect to the norm (5.3.3), that is, $d\|\omega_k - \omega_m\| \to 0$ as $k, m \to \infty$. Our task is to prove that its limit, denoted by $d\omega$, exists and that $d\omega$ is contained in $\mathfrak{Ovm}$. Assuming that $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ is a pseudo-orthonormal basis of V satisfying (5.2.1), from (5.3.3) we deduce that

$$\lim_{k,m\to\infty} d\|\prec \mathfrak{e}_i \,|\, (\omega_k - \omega_m)\, \mathfrak{e}_j \succ\| = 0 \qquad \text{for all } i, j = 1, \ldots, 2n\,.$$

Consequently, each sequence $(d \prec \mathfrak{e}_i \,|\, \omega_k\, \mathfrak{e}_j \succ)_{k\in\mathbb{N}}$ is a Cauchy sequence of complex measures in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ for all $i, j \in \{1, \ldots, 2n\}$. Since $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ is a complex Banach space with respect to the total variation $d\|\cdot\|$ in virtue of [**40**, Aufgabe VII.1.7], there is a

complex measure $d\omega_{i,j} \in \mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$, being the unique limit of $(d \prec \mathfrak{e}_i \,|\, \omega_k\, \mathfrak{e}_j \succ)_{k \in \mathbb{N}}$ for all $i, j \in \{1, \ldots, 2n\}$.

Next, for all $i, j \in \{1, \ldots, 2n\}$, the complex measures $d\omega_{i,j}$ in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ give rise to an operator-valued measure $d\omega$ on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ in such a way that, for all $i, j = 1, \ldots, 2n$, we are given $d \prec \mathfrak{e}_i \,|\, \omega\, \mathfrak{e}_j \succ = d\omega_{i,j}$. More precisely, defining the operator $\omega(\Omega) \in \mathrm{L}(V)$ for any $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$ by

$$\omega(\Omega) := \begin{pmatrix} \omega_{1,1}(\Omega) & \cdots & \omega_{1,2n}(\Omega) \\ \vdots & \ddots & \vdots \\ \omega_{n,1}(\Omega) & \cdots & \omega_{n,2n}(\Omega) \\ -\omega_{n+1,1}(\Omega) & \cdots & -\omega_{n+1,2n}(\Omega) \\ \vdots & \ddots & \vdots \\ -\omega_{1,2n}(\Omega) & \cdots & -\omega_{2n,2n}(\Omega) \end{pmatrix} \in \mathrm{L}(V)\,,$$

we obtain a mapping $d\omega \colon \mathcal{B}(\hat{\mathscr{M}}) \to \mathrm{L}(V)$ such that $d \prec \mathfrak{e}_i \,|\, \omega\, \mathfrak{e}_j \succ = d\omega_{i,j} \in \mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ for all $i, j \in \{1, \ldots, 2n\}$. Since $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ is a basis of V, for any $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$ and arbitrary elements $u = \sum_{i=1}^{2n} \alpha_i\, \mathfrak{e}_i$, $v = \sum_{j=1}^{2n} \beta_j\, \mathfrak{e}_j \in V$ we arrive at

$$\prec u \,|\, \omega(\Omega)\, v \succ = \sum_{i,j=1}^{2n} \overline{\alpha_i}\, \beta_j \prec \mathfrak{e}_i \,|\, \omega(\Omega)\, \mathfrak{e}_j \succ = \sum_{i,j=1}^{2n} \overline{\alpha_i}\, \beta_j\, \omega_{i,j}(\Omega)\,.$$

The fact that $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ is a complex Banach space implies that

$$d \prec u \,|\, \omega\, v \succ = \sum_{i,j=1}^{2n} \overline{\alpha_i}\, \beta_j\, d\omega_{i,j} \in \mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}}) \qquad \text{for all } u, w \in V\,.$$

This shows that $d\omega \in \mathfrak{Ovm}$ is an operator-valued measure in view of Definition 5.3.3. Thus $(\mathfrak{Ovm}, d\|\cdot\|)$ is a complex Banach space with respect to the norm $d\|\cdot\|$ defined by (5.3.3). Since each norm induces a corresponding Fréchet metric, $(\mathfrak{Ovm}, d\|\cdot\|)$ can be regarded as a metric space. In particular, each complex vector space is a real one, and each Banach space is locally convex. This completes the proof. □

Remark 5.3.6. *The set of negative definite measures $\mathfrak{Ndm}$ clearly is a subset of the vector space $\mathfrak{Ovm}$. However, $\mathfrak{Ndm}$ itself is* not *a vector space (see* [**109**, Remark 5.6]*), but a* cone*, i.e. a closed subset under multiplication with positive real numbers.*

Next, let us introduce the support of operator-valued measures as follows:

Definition 5.3.7. *We define the* ***support*** *of an operator-valued measure $d\omega$ in the class of operator-valued measures $\mathfrak{Ovm}$ as the support of its variation measure $d|\omega|$, i.e.*

$$\operatorname{supp} d\omega := \operatorname{supp} d|\omega| = \hat{\mathscr{M}} \setminus \bigcup \Big\{ U \subset \hat{\mathscr{M}} : U \text{ open and } d|\omega|(U) = 0 \Big\}\,.$$

Since $d|\omega|$ is a locally finite measure on a locally compact Polish space, we conclude that $d|\omega|$ is regular and has support, $d|\omega|(\hat{\mathscr{M}} \setminus \operatorname{supp} d|\omega|) = 0$.

In a similar fashion, following [**16**, Definition 7.1.5], an operator-valued measure $d\omega$ is called *regular* if and only if $d|\omega|$ is regular. Moreover, the measure $d\omega$ is said to be *tight* if for every $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset \hat{\mathscr{M}}$ such that $d|\omega|(\hat{\mathscr{M}} \setminus K_\varepsilon) < \varepsilon$ (cf. [**16**, Definition 7.1.4]).

DEFINITION 5.3.8. *In analogy to negative definite measures (see Definition 5.3.2), for any bounded Borel measurable function $f : \hat{\mathscr{M}} \to \mathbb{C}$ we define integration with respect to operator-valued measures $d\omega$ by*

$$\prec u \,|\, \left(\int_{\hat{\mathscr{M}}} f(k)\, d\omega(k) \right) v \succ := \int_{\hat{\mathscr{M}}} f(k)\, d \prec u \,|\, \omega(k)\, v \succ \qquad \text{for all } u, v \in V\,.$$

Let us finally state the definition of weak convergence of operator-valued measures.

DEFINITION 5.3.9. *We shall say that a sequence $(d\omega_k)_{k\in\mathbb{N}}$ of operator-valued measures* ***converges weakly*** *to some operator-valued measure $d\omega$ if and only if*

$$\lim_{k\to\infty} \int_{\hat{\mathscr{M}}} f\, d \prec u \,|\, \omega_k\, v \succ = \int_{\hat{\mathscr{M}}} f\, d \prec u \,|\, \omega\, v \succ \qquad \text{for all } u, v \in V \text{ and } f \in C_b(\hat{\mathscr{M}})\,.$$

We write symbolically $d\omega_k \rightharpoonup d\omega$.

Whenever $d\nu \in \mathfrak{Ndm}$ is a negative definite measure, we recall that, for all $u, v \in V$, the complex measure $d \prec u \,|\, \nu\, v \succ$ in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ is defined by polarization (5.3.2). Thus a sequence of negative definite measures $(d\nu)_{k\in\mathbb{N}}$ converges weakly to some negative definite measure $d\nu \in \mathfrak{Ndm}$ if and only if

$$\lim_{k\to\infty} \int_{\hat{\mathscr{M}}} f\, d \prec u \,|\, \nu_k\, u \succ = \int_{\hat{\mathscr{M}}} f\, d \prec u \,|\, \nu\, u \succ \qquad \text{for all } u \in V \text{ and } f \in C_b(\hat{\mathscr{M}})\,.$$

By polarization (5.3.2) we then conclude that

$$\lim_{k\to\infty} \int_{\hat{\mathscr{M}}} f\, d \prec u \,|\, \nu_k\, v \succ = \int_{\hat{\mathscr{M}}} f\, d \prec u \,|\, \nu\, v \succ \qquad \text{for all } u, v \in V \text{ and } f \in C_b(\hat{\mathscr{M}})$$

in accordance with Definition 5.3.9.

5.3.3. Causal Variational Principles. After these technical preliminaries, we now return to causal variational principles in the homogeneous setting. Motivated by (5.2.5), the fermionic projector $P(x, y)$ in the homogeneous setting takes the form

$$P(x, y) = \int_{\hat{\mathscr{M}}} e^{ik(y-x)}\, d\nu(k) \tag{5.3.4}$$

for all $x, y \in \mathscr{M}$, where the measure $d\nu$ is given by (5.2.7). Generalizing $d\nu$ according to §5.2.3 and §5.3.2 to operator-valued measures, for a given operator-valued measure $d\omega$ on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ and all $x, y \in \mathscr{M}$ we introduce the *kernel of the fermionic projector* by

$$P(x, y) : V \to V, \qquad P(x, y) := \int_{\hat{\mathscr{M}}} e^{ik(y-x)}\, d\omega(k)\,.$$

In order to emphasize the dependence on the operator-valued measure $d\omega$, we also write $P[\omega](x, y)$. As $P(x, y)$ is supposed to be *homogeneous*, only the difference of two spacetime points $x, y \in \mathscr{M}$ matters; denoting the difference vector by $\xi = y - x \in \mathscr{M}$, the kernel of the fermionic projector reads

$$P(\xi) : V \to V, \qquad P(\xi) = \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d\omega(k)\,. \tag{5.3.5}$$

We remark that, for every $\xi \in \mathscr{M}$, the adjoint of $P(\xi)$ with respect to the spin scalar product $\prec . \,|\, . \succ$ is given by (for details see [**109**, §3.4])

$$P(\xi)^* = P(-\xi)\,. \tag{5.3.6}$$

Next, we introduce the corresponding *fermionic operator* P by (cf. [**109**, eq. (3.7)])

$$(P\phi)(x) := \int_{\mathscr{M}} P(x,y)\, \phi(y)\, d\mu(y)\,.$$

In the homogeneous setting, the fermionic operator can be written as a convolution,

$$(P\phi)(x) = \int_{\mathscr{M}} P(x-y)\, \phi(y)\, d\mu(y)\,;$$

it is well-defined for instance on the Schwartz functions $\mathcal{S}(\mathscr{M}, V)$ (cf. [**109**, §5]).

We now proceed similarly to the basic ideas in [**53**, Chapter 3] in order to set up a variational principle. More explicitly, our goal is to introduce endomorphisms on the vector space V which allow us to consider the corresponding eigenvalues. Defining the so-called spectral weight as the sum of the absolute values of these eigenvalues gives rise to a Lagrangian, and integration yields the corresponding causal action. The first step in order to obtain suitable endomorphisms is to form the *closed chain*, which (as motivated by [**53**, §3.5]) for any $\xi \in \mathscr{M}$ is defined as the mapping

$$A(\xi) : V \to V, \qquad A(\xi) := P(\xi)\, P(-\xi)\,. \tag{5.3.7}$$

We also write $A[\omega](\xi)$ in order to clarify the dependence of the closed chain on the operator-valued measure $d\omega$. Next, given a linear operator $A : V \to V$, we define the *spectral weight* by

$$|A| := \sum_{i=1}^{2n} |\lambda_i|\,,$$

where by $(\lambda_i)_{i=1,\ldots,2n}$ we denote the eigenvalues of the operator A, counted with algebraic multiplicities. In this way, the spectral weight yields a connection between endomorphisms and scalar functionals.

In order to set up a real-valued variational principle on the class of operator-valued measures, for every $d\omega \in \mathfrak{Ovm}$ we introduce the *Lagrangian*

$$\mathcal{L}[\omega] : \mathscr{M} \to \mathbb{R}_0^+\,, \qquad \mathcal{L}[\omega](\xi) := \big|A[\omega](\xi)^2\big| - \frac{1}{2n}\,|A[\omega](\xi)|^2\,. \tag{5.3.8}$$

Defining the *causal action* $\mathcal{S} : \mathfrak{Ovm} \to \mathbb{R}_0^+ \cup \{+\infty\}$ by

$$\mathcal{S}(\omega) := \int_{\mathscr{M}} \mathcal{L}[\omega](\xi)\, d\mu(\xi)\,,$$

the *causal variational principle in the homogeneous setting* is to

minimize $\mathcal{S}(\nu)$ by suitably varying $d\nu$ in $\mathfrak{Ndm}$.

In order to exclude trivial minimizers, we impose the *trace constraint*

$$\mathrm{Tr}_V\big(\nu(\hat{\mathscr{M}})\big) = c \tag{5.3.9}$$

for some $c > 0$. Additionally, for $f > 0$ we shall either introduce the constraint

$$\mathrm{Tr}_V\big(-S\nu(\hat{\mathscr{M}})\big) \leq f \tag{5.3.10}$$

(where S denotes the signature matrix (5.2.2)) or the side condition

$$|\nu(\hat{\mathscr{M}})| \leq f\,. \tag{5.3.11}$$

A motivation for the constraints (5.3.9)–(5.3.11) can be found in Appendix A.4. For the connection to the boundedness constraint as considered in [**56**, Section 4] we refer to the Section 5.4 below (cf. Lemma 5.4.6).

DEFINITION 5.3.10. *Given a subset $N \subset \mathfrak{Ndm}$, the **causal variational principle in the homogeneous setting** is to*

$$\text{minimize } \mathcal{S}(\nu) \text{ by varying } d\nu \text{ in } N \subset \mathfrak{Ndm}\,.$$

Concerning the side conditions (5.3.9)–(5.3.11), the subset N takes the form

$$N = \{d\nu \in \mathfrak{Ndm} : d\nu \text{ satisfies conditions (5.3.9) and (5.3.10)}\} \qquad \text{or}$$
$$N = \{d\nu \in \mathfrak{Ndm} : d\nu \text{ satisfies conditions (5.3.9) and (5.3.11)}\}\,.$$

Motivated by [**168**, Definition 43.4], we define a minimizer for $\mathcal{S}$ as follows:

DEFINITION 5.3.11. *A negative definite measure $d\nu \in N$ is said to be a **minimizer** for the causal variational principle* (5.4.8) *if and only if*

$$\mathcal{S}(\tilde{\nu}) \geq \mathcal{S}(\nu) \qquad \text{for all } d\tilde{\nu} \in N\,.$$

For further details concerning the calculus of variations we refer to [**168**, Chapter 43].

5.4. Estimates for Minimizing Sequences

This section is intended to derive estimates for minimizing sequences in the case that certain side conditions are satisfied. In particular, under such constraints we deduce that minimizing sequences for causal variational principles in the homogeneous setting are uniformly tight. Unfortunately, this seems not to suffice for directly proving the existence of minimizing measures.

Let us assume that $(d\nu^{(k)})_{k\in\mathbb{N}}$ is a sequence of negative definite measures in $\mathfrak{Ndm}$, either satisfying

$$\mathrm{Tr}_V(-S\nu^{(k)}(\hat{\mathscr{M}})) \leq f \qquad \text{or} \qquad |\nu^{(k)}(\hat{\mathscr{M}})| \leq f$$

for all $k \in \mathbb{N}$ and some positive constant $f > 0$ (and $|\cdot|$ denotes the spectral weight). The aim of this subsection is to show that in both cases, there exists a sequence of unitary matrices $(U_k)_{k\in\mathbb{N}}$ in $\mathrm{L}(V)$ (with respect to $\prec \,.\, | \,.\, \succ$) such that the resulting sequence $(U_k\, d\nu^{(k)}\, U_k^{-1})_{k\in\mathbb{N}}$ is bounded in $\mathfrak{Ndm}$ (with respect to the norm (5.3.3)). In particular, whenever the first condition is imposed, it eventually turns out that one can choose $U_k = \mathbb{1}_V$ for all $k \in \mathbb{N}$. In preparation, let us state the following results:

PROPOSITION 5.4.1. *For all $B, C \in \mathrm{L}(V)$, the operator products BC and CB have the same spectrum.*

PROOF. Follow the arguments in [**54**, Section 3] or cf. [**53**, eq. (3.5.6)]. □

LEMMA 5.4.2. *Assume that $U \in \mathrm{L}(V)$ is unitary (with respect to $\prec \,.\, | \,.\, \succ$), and let $d\nu$ in $\mathfrak{Ndm}$. Then the operators $\nu(\hat{\mathscr{M}})$ and $U\,\nu(\hat{\mathscr{M}})\,U^{-1}$ have the same spectrum.*

PROOF. Applying Proposition 5.4.1, we infer that the operators

$$\nu(\hat{\mathscr{M}}) = (\nu(\hat{\mathscr{M}})\,U^{-1})\,U \qquad \text{and} \qquad U\,\nu(\hat{\mathscr{M}})\,U^{-1} = U\,(\nu(\hat{\mathscr{M}})\,U^{-1})$$

have the same spectrum for any unitary matrix U in $\mathrm{L}(V)$. □

COROLLARY 5.4.3. *For any negative definite measure $d\nu \in \mathfrak{Ndm}$ and arbitrary unitary transformations U on V (with respect to $\prec . \,|\, . \succ$),*

$$\mathcal{L}[U\nu U^{-1}] = \mathcal{L}[\nu] \qquad \text{and} \qquad \mathcal{S}(U\nu U^{-1}) = \mathcal{S}(\nu)\,.$$

PROOF. Introducing the kernel of the fermionic projector by (5.3.5) and making use of Definition 5.3.2, for all $u, w \in V$ and $\xi \in \mathscr{M}$ we obtain

$$\begin{aligned}
\prec u \,|\, P[U\nu U^{-1}](\xi)\, w \succ &= \prec u \,|\, \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d\left(U\nu U^{-1}\right)(k)\, w \succ \\
&= \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d \prec u \,|\, U\nu(k) U^{-1} w \succ = \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d \prec U^{-1} u \,|\, \nu(k) U^{-1} w \succ \\
&= \prec U^{-1} u \,|\, \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d\nu(k) U^{-1} w \succ = \prec u \,|\, U \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d\nu(k) U^{-1} w \succ \\
&= \prec u \,|\, U\, P[\nu](\xi) U^{-1} w \succ
\end{aligned}$$

for any negative definite measure $d\nu \in \mathfrak{Ndm}$ and any unitary matrix U (with respect to $\prec . \,|\, . \succ$). Thus non-degeneracy of the indefinite inner product implies that

$$P[U\nu U^{-1}] = U\, P[\nu]\, U^{-1}\,.$$

Henceforth, employing Lemma 5.4.2, we deduce that the spectral weight of the closed chain A defined by (5.3.7) is unaffected by unitary similarity, i.e.

$$\left|A[U\nu U^{-1}](\xi)\right| = \left|U\, A[\nu](\xi)\, U^{-1}\right| = |A[\nu](\xi)| \qquad \text{for all } \xi \in \mathscr{M}\,.$$

Analogously, for every $\xi \in \mathscr{M}$ we obtain

$$\left|A[U\nu U^{-1}](\xi)^2\right| = \left|\left(U\, A[\nu](\xi)\, U^{-1}\right)^2\right| = \left|A[\nu](\xi)^2\right|\,,$$

thus implying that

$$\mathcal{L}[U\nu U^{-1}](\xi) = \mathcal{L}[\nu](\xi) \qquad \text{for all } \xi \in \mathscr{M}$$

as well as $\mathcal{S}(U\nu U^{-1}) = \mathcal{S}(\nu)$. This completes the proof. □

We are now in the position to prove the following result:

LEMMA 5.4.4. *Let $f > 0$ and assume that $(d\nu^{(k)})_{k\in\mathbb{N}}$ is a sequence in $\mathfrak{Ndm}$ such that*

$$\mathrm{Tr}_V\left(-S\nu^{(k)}(\hat{\mathscr{M}})\right) \leq f \qquad \textit{for all } k \in \mathbb{N}$$

(where S denotes the signature matrix). Then there exists a positive constant $C > 0$ in such a way that $d\|\nu^{(k)}\| \leq C$ for all $k \in \mathbb{N}$, where $d\|\cdot\|$ denotes the total variation according to Definition 5.3.4.

PROOF. For convenience, we fix an arbitrary integer $k \in \mathbb{N}$ and let $d\nu = d\nu^{(k)}$. Next, we let $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ be a pseudo-orthonormal basis of V with signature matrix S such that (5.2.1) is satisfied. Then $d \prec \mathfrak{e}_i \,|\, \nu\, \mathfrak{e}_j \succ$ is a finite complex measure in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ for every $i, j \in \{1, \ldots, 2n\}$ according to Definition 5.3.3, i.e.

$$d\|\prec \mathfrak{e}_i \,|\, \nu\, \mathfrak{e}_j \succ\| = d|\prec \mathfrak{e}_i \,|\, \nu\, \mathfrak{e}_j \succ|(\hat{\mathscr{M}}) < \infty \qquad \text{for all } i, j = 1, \ldots, 2n\,.$$

Employing the definition of the total variation of complex measures and applying the Schwarz inequality (see e.g. [**109**, Lemma A.13] or [**79**, ineq. (2.3.9)]), we obtain

$$\begin{aligned} d\|\prec \mathfrak{e}_i \mid \nu\, \mathfrak{e}_j \succ\| &= \sup \sum_{n\in\mathbb{N}} |\prec \mathfrak{e}_i \mid \nu(E_n)\, \mathfrak{e}_j \succ| = \sup \sum_{n\in\mathbb{N}} |\prec \mathfrak{e}_i \mid -\nu(E_n)\, \mathfrak{e}_j \succ| \\ &\leq \sup \sum_{n\in\mathbb{N}} \sqrt{|\prec \mathfrak{e}_i \mid -\nu(E_n)\, \mathfrak{e}_i \succ|}\, \sqrt{|\prec \mathfrak{e}_j \mid -\nu(E_n)\, \mathfrak{e}_j \succ|}\,, \end{aligned}$$

where the supremum is taken over all partitions $(E_n)_{n\in\mathbb{N}}$ of $\hat{\mathscr{M}}$ (cf. [**139**, Chapter 6]). Applying Young's inequality (see e.g. [**5**, §1]), for all $i, j \in \{1, \ldots, 2n\}$ we arrive at

$$\begin{aligned} d\|\prec \mathfrak{e}_i \mid \nu\, \mathfrak{e}_j \succ\| &\leq \sup \sum_{n\in\mathbb{N}} \sqrt{|\prec \mathfrak{e}_i \mid -\nu(E_n)\, \mathfrak{e}_j \succ|}\, \sqrt{|\prec \mathfrak{e}_j \mid -\nu(E_n)\, \mathfrak{e}_j \succ|} \\ &\leq \frac{1}{2} \sup \sum_{n\in\mathbb{N}} \left(|\prec \mathfrak{e}_i \mid -\nu(E_n)\, \mathfrak{e}_i \succ| + |\prec \mathfrak{e}_j \mid -\nu(E_n)\, \mathfrak{e}_j \succ|\right) \\ &\leq \frac{1}{2} \left[\sup \sum_{n\in\mathbb{N}} |\prec \mathfrak{e}_i \mid -\nu(E_n)\, \mathfrak{e}_i \succ| + \sup \sum_{n\in\mathbb{N}} |\prec \mathfrak{e}_j \mid -\nu(E_n)\, \mathfrak{e}_j \succ| \right] \\ &= \frac{1}{2} \left(d\|\prec \mathfrak{e}_i \mid \nu\, \mathfrak{e}_i \succ\| + d\|\prec \mathfrak{e}_j \mid \nu\, \mathfrak{e}_j \succ\| \right) . \end{aligned}$$

Due to the fact that $d \prec \mathfrak{e}_i \mid -\nu\, \mathfrak{e}_i \succ$ is a positive measure for each $i \in \{1, \ldots, 2n\}$, the total variation $d\|\prec \mathfrak{e}_i \mid \nu\, \mathfrak{e}_j \succ\|$ is bounded by

$$d\|\prec \mathfrak{e}_i \mid \nu\, \mathfrak{e}_j \succ\| \leq \sum_{i=1}^{2n} d\|\prec \mathfrak{e}_i \mid \nu\, \mathfrak{e}_i \succ\| = \sum_{i=1}^{2n} \prec \mathfrak{e}_i \mid -\nu(\hat{\mathscr{M}})\, \mathfrak{e}_i \succ$$

for all $i, j \in \{1, \ldots, 2n\}$. The last expression can be estimated by

$$\sum_{i=1}^{2n} \prec \mathfrak{e}_i \mid -\nu(\hat{\mathscr{M}})\, \mathfrak{e}_i \succ = \sum_{i=1}^{2n} \langle \mathfrak{e}_i \mid -S\nu(\hat{\mathscr{M}})\, \mathfrak{e}_i \rangle = \mathrm{Tr}_V \left(-S\nu(\hat{\mathscr{M}}) \right) \leq f\,,$$

thus completing the proof. □

In the case that the spectral weight is bounded (in analogy to [**54**, Theorem 6.1]), we obtain the following result:

LEMMA 5.4.5. *Let $f > 0$ and assume that $(d\nu^{(k)})_{k\in\mathbb{N}}$ is a sequence in $\mathfrak{Ndm}$ such that*

$$|\nu^{(k)}(\hat{\mathscr{M}})| \leq f \qquad \text{for all } k \in \mathbb{N}$$

(where $|\cdot|$ denotes the spectral weight). Then there is a sequence $(U_k)_{k\in\mathbb{N}}$ of unitary operators on V (with respect to $\prec\, .\, |\, .\, \succ$) as well as a positive constant $C > 0$ such that $d\|U_k\, \nu^{(k)}\, U_k^{-1}\| \leq C$ for all $k \in \mathbb{N}$ (where $d\|\cdot\|$ denotes the total variation according to Definition 5.3.4).

PROOF. The basic idea is to make use of [**56**, Lemma 4.4]. For convenience, we fix an arbitrary integer $k \in \mathbb{N}$ and let $d\nu = d\nu^{(k)}$. Moreover, let $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ be a pseudo-orthonormal basis of V with signature matrix S such that (5.2.1) is satisfied (see for instance [**79**, §2.3] or [**109**, §3.3]). Since V is a finite-dimensional vector space, all norms

on $\mathrm{L}(V)$ are equivalent, and one of these norms is given by

$$\|A\|_1 = \max_{j=1,\ldots,2n} \sum_{i=1}^{2n} |\langle \mathfrak{e}_i \mid A\mathfrak{e}_j \rangle| \tag{5.4.1}$$

for any $A \in \mathrm{L}(V)$, where $|\,.\,|$ denotes the absolute value. Moreover, for any unitary matrix U in $\mathrm{L}(V)$ (with respect to $\prec\,.\mid.\,\succ$), we may introduce another pseudo-orthonormal basis $(\mathfrak{f}_j)_{j=1,\ldots,2n}$ by

$$\mathfrak{f}_i := U^{-1}\,\mathfrak{e}_i \qquad \text{for all } i = 1,\ldots,2n\,. \tag{5.4.2}$$

Making use of $U^* = U^{-1}$, for all $i,j = 1,\ldots,2n$ we obtain

$$d \prec \mathfrak{e}_i \mid U\,\nu\,U^{-1}\mathfrak{e}_j \succ = d \prec U^*\,\mathfrak{e}_i \mid \nu\,U^{-1}\,\mathfrak{e}_j \succ = d \prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_j \succ . \tag{5.4.3}$$

Since $d\nu$ is a negative definite measure, the operator $-\nu(\hat{\mathscr{M}})$ is positive (5.3.1). Thus in view of [**56**, Lemma 4.4], for any $\varepsilon > 0$ there is a unitary matrix $U = U(\varepsilon)$ in $\mathrm{L}(V)$ (with respect to $\prec\,.\mid.\,\succ$) so that $U\,\nu(\hat{\mathscr{M}})\,U^{-1}$ is diagonal, up to an arbitrarily small error term $\Delta\nu(\hat{\mathscr{M}})$ with $\|\Delta\nu(\hat{\mathscr{M}})\|_1 < \varepsilon$. Since $k \in \mathbb{N}$ is arbitrary, we thus obtain a sequence of negative definite measures $(U_k\,d\nu^{(k)}\,U_k^{-1})_{k\in\mathbb{N}}$.

Next, in order to prove that $(U_k\,d\nu^{(k)}\,U_k^{-1})_{k\in\mathbb{N}}$ is bounded with respect to the total variation defined by (5.3.3), for each $k \in \mathbb{N}$ we consider the basis $(\mathfrak{f}_i)_{i=1,\ldots,2n}$ given by (5.4.2) with respect to the unitary matrix $U = U_k$. Accordingly, each $d \prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_j \succ$ is a finite complex measure in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ in view of Definition 5.3.3,

$$d\|\prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_j \succ\| = d|\prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_j \succ|(\hat{\mathscr{M}}) < \infty \qquad \text{for all } i,j = 1,\ldots,2n\,.$$

Employing the definition of the total variation of complex measures and applying the Schwarz inequality in analogy to the proof of Lemma 5.4.4, we obtain

$$d\|\prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_j \succ\| \leq \sup \sum_{n\in\mathbb{N}} \sqrt{|\prec \mathfrak{f}_i \mid -\nu(E_n)\,\mathfrak{f}_i \succ|}\,\sqrt{|\prec \mathfrak{f}_j \mid -\nu(E_n)\,\mathfrak{f}_j \succ|}\,,$$

where the supremum is taken over all partitions $(E_n)_{n\in\mathbb{N}}$ of $\hat{\mathscr{M}}$ (cf. [**139**, Chapter 6]). Applying Young's inequality in analogy to the proof of Lemma 5.4.4, we arrive at

$$d\|\prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_j \succ\| \leq \frac{1}{2}\,(d\|\prec \mathfrak{f}_i \mid \nu\,\mathfrak{f}_i \succ\| + d\|\prec \mathfrak{f}_j \mid \nu\,\mathfrak{f}_j \succ\|)$$

for all $i,j \in \{1,\ldots,2n\}$. Since $S = S^{-1}$ and $U^* = S^{-1}\,U^\dagger\,S$ in view of [**79**, eq. (4.1.3)] (where $U^\dagger$ denotes the adjoint with respect to $\langle\,.\,|\,.\,\rangle$ and U^* the adjoint with respect to $\prec\,.\mid.\,\succ$), for all $i = 1,\ldots,2n$ we obtain

$$\begin{aligned}
d\|\prec \mathfrak{f}_i \mid -\nu\,\mathfrak{f}_i \succ\| &= \prec \mathfrak{f}_i \mid -\nu(\hat{\mathscr{M}})\,\mathfrak{f}_i \succ \leq \sum_{i,j=1}^{2n} \Big|\prec U^{-1}\,\mathfrak{e}_i \mid \nu(\hat{\mathscr{M}})\,U^{-1}\,\mathfrak{e}_j \succ\Big| \\
&\leq \sum_{i,j=1}^{2n} \Big|\prec SU^*\,SS\mathfrak{e}_i \mid S\nu(\hat{\mathscr{M}})\,U^{-1}\,\mathfrak{e}_j \succ\Big| = \sum_{i,j=1}^{2n} \Big|\langle U^\dagger\,\mathfrak{e}_i \mid \nu(\hat{\mathscr{M}})\,U^{-1}\,\mathfrak{e}_j\rangle\Big|\;|s_i| \\
&= \sum_{i,j=1}^{2n} \Big|\langle \mathfrak{e}_i \mid U\,\nu(\hat{\mathscr{M}})\,U^{-1}\,\mathfrak{e}_j\rangle\Big| \overset{(5.4.1)}{\leq} 2n\,\|U\,\nu(\hat{\mathscr{M}})\,U^{-1}\|_1\,,
\end{aligned}$$

where we made use of $S\mathfrak{e}_i = s_i\mathfrak{e}_i$ with $|s_i| = |\langle \mathfrak{e}_i \mid S\mathfrak{e}_i\rangle| = 1$ for all $i = 1,\ldots,2n$ and employed the fact that $d \prec \mathfrak{f}_i \mid -\nu\,\mathfrak{f}_i \succ$ is a positive measure for any $i \in \{1,\ldots,2n\}$.

Taken the previous results together, by (5.4.3) we obtain the inequality

$$d\|\prec \mathfrak{e}_i \,|\, U\,\nu\,U^{-1}\,\mathfrak{e}_i \succ\| = d\|\prec \mathfrak{f}_i \,|\, -\nu\,\mathfrak{f}_i \succ\| \leq 2n\,\|U\,\nu(\hat{\mathscr{M}})\,U^{-1}\|_1 \tag{5.4.4}$$

for all $i = 1, \ldots, 2n$. Thus it only remains to find an upper bound for $\|U\,\nu(\hat{\mathscr{M}})\,U^{-1}\|_1$ in terms of f by establishing a connection to the spectral weight $|\nu(\hat{\mathscr{M}})|$. To this end we exploit the fact that $U\,\nu(\hat{\mathscr{M}})\,U^{-1}$ is diagonal according to [**56**, Lemma 4.4], up to an arbitrarily small error term $\Delta\nu(\hat{\mathscr{M}})$,

$$U\,\nu(\hat{\mathscr{M}})\,U^{-1} = \operatorname{diag}\big(\tilde{\lambda}_1(U), \ldots, \tilde{\lambda}_{2n}(U)\big) + \Delta\nu(\hat{\mathscr{M}})\,.$$

Denoting the eigenvalues of $U\,\nu(\hat{\mathscr{M}})\,U^{-1}$ by $\lambda_i(U)$ for all $i = 1, \ldots, 2n$, by choosing the error term $\Delta\nu(\hat{\mathscr{M}})$ sufficiently small we can arrange that

$$\sum_{i=1}^{2n} |\tilde{\lambda}_i(U) - \lambda_i(U)| < \varepsilon \qquad \text{for any } \varepsilon > 0\,.$$

Since the off-diagonal elements $\|\Delta\nu(\hat{\mathscr{M}})\|_1 < \varepsilon$ are arbitrarily small, we thus obtain

$$\|U\,\nu(\hat{\mathscr{M}})\,U^{-1}\|_1 \leq \|\operatorname{diag}(\tilde{\lambda}_1(U), \ldots, \tilde{\lambda}_{2n}(U))\|_1 + \|\Delta\nu(\hat{\mathscr{M}})\|_1 \leq \sum_{i=1}^{2n} |\lambda_i(U)| + 2\varepsilon\,.$$

Applying Lemma 5.4.2, we conclude that $|\nu(\hat{\mathscr{M}})| = |U\,\nu(\hat{\mathscr{M}})\,U^{-1}|$ (where $|\cdot|$ denotes the spectral weight). Choosing $\varepsilon < 1/2$, we arrive at

$$\|U\,\nu(\hat{\mathscr{M}})\,U^{-1}\|_1 \leq |\nu(\hat{\mathscr{M}})| + 1 \leq f + 1\,.$$

Hence in view of Definition 5.3.4 and (5.4.4), we obtain

$$d\|U_k\,\nu^{(k)}\,U_k^{-1}\| = \sum_{i,j=1}^{2n} d\|\prec \mathfrak{e}_i \,|\, U_k\,\nu^{(k)}\,U_k^{-1}\,\mathfrak{e}_j \succ\| \leq (2n)^3\,(f+1) =: C\,.$$

This completes the proof. □

Let us finally establish a connection to the boundedness constraint as considered in [**56**, Section 4] (which originally was proposed in [**53**, eq. (3.5.10)] as a constraint for the causal action principle). In the homogeneous setting, for any operator-valued measure $d\omega \in \mathfrak{Ovm}$ we introduce the mapping $\mathfrak{t}[\omega] : \mathscr{M} \to \mathbb{R}_0^+$ by

$$\mathfrak{t}[\omega](\xi) := |A[\omega](\xi)|^2 \qquad \text{for all } \xi \in \mathscr{M}\,.$$

We then define the functional $\mathcal{T} : \mathfrak{Ovm} \to \mathbb{R}_0^+ \cup \{+\infty\}$ by

$$\mathcal{T}(\omega) := \int_{\mathscr{M}} \mathfrak{t}[\omega](\xi)\, d\mu(\xi) = \int_{\mathscr{M}} |A[\omega](\xi)|^2\, d\mu(\xi)\,.$$

Given $C > 0$, the corresponding *boundedness constraint* reads

$$\mathcal{T}(\omega) \leq C\,. \tag{5.4.5}$$

These considerations allow us to derive the following result.

LEMMA 5.4.6. *Whenever $d\nu \in \mathfrak{Ndm}$ is a negative definite measure satisfying the boundedness constraint* (5.4.5)*, it satisfies condition* (5.3.11) *for some constant $f > 0$.*

PROOF. Let us first note that $\mathfrak{t}[\nu] \in L^1_{\mathrm{loc}}(\mathscr{M})$ whenever $d\nu \in \mathfrak{Ndm}$ satisfies (5.4.5). In analogy to [**72**, Section 3.4], for every $f \in L^1_{\mathrm{loc}}(\mathscr{M})$ we then introduce

$$G_r f(x) := \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y)\, d\mu(y) \qquad \text{for all } x \in \mathscr{M} \text{ and } r > 0\,,$$

and by virtue of [**72**, Theorem 3.18] we know that

$$\lim_{r \to 0} G_r f(x) = f(x) \qquad \text{for almost every } x \in \mathscr{M}\,.$$

Since $\mu(B_\varepsilon(x)) > 0$ for every $x \in \mathscr{M}$ and arbitrary $\varepsilon > 0$, continuity of $A[\nu]$ yields the existence of $x_0 \in \mathscr{M}$ such that

$$|A[\nu](0)|^2 < \mathfrak{t}[\nu](x_0) + 1 = \lim_{\varepsilon \searrow 0} \int_{B_\varepsilon(x_0)} |A[\nu](\xi)|^2\, d\mu(\xi) + 1 \leq C + 1\,. \tag{5.4.6}$$

We now apply (5.4.6) in order to prove that $|\nu(\hat{\mathscr{M}})| < f$ for some constant $f > 0$. To this end, we essentially employ [**56**, Lemma 4.4]. More precisely, for any negative definite measure $d\nu$ and arbitrary $\varepsilon > 0$, there is a unitary operator $U \in \mathrm{L}(V)$ (with respect to $\prec . \,|\, . \succ$) such that

$$U\, \nu(\hat{\mathscr{M}})\, U^{-1} = -\operatorname{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{2n}) + \Delta\nu(\hat{\mathscr{M}})\,,$$

where the real parameters $\tilde{\lambda}_i$ $(i = 1, \ldots, 2n)$ are ordered according to [**56**, eq. (2.6)], and $\|\Delta\nu(\hat{\mathscr{M}})\| < \varepsilon$. Denoting by $\{.,.\}$ the anti-commutator, we thus obtain

$$\begin{aligned} U\, A[\nu](0)\, U^{-1} &= \big(U\, \nu(\hat{\mathscr{M}})\, U^{-1}\big)^2 \\ &= \operatorname{diag}\big(\tilde{\lambda}_1^2, \ldots, \tilde{\lambda}_{2n}^2\big) - \Big\{\operatorname{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{2n}), \Delta\nu(\hat{\mathscr{M}})\Big\} + \Delta\nu(\hat{\mathscr{M}})^2\,. \end{aligned}$$

Since $\|\nu(\hat{\mathscr{M}})\| < \infty$, the absolute values of $\tilde{\lambda}_i$ are bounded for all $i = 1, \ldots, 2n$; from this we conclude that the spectrum of $\operatorname{diag}\big(\tilde{\lambda}_1^2, \ldots, \tilde{\lambda}_{2n}^2\big)$ coincides with the spectrum of $A[\nu](0)$, up to an arbitrarily small error term (where we applied the fact that the spectra of $A[\nu](0)$ and $U\, A[\nu](0)\, U^{-1}$ coincide according to Lemma 5.4.2). In a similar fashion, one can show that the spectra of $\nu(\hat{\mathscr{M}})$ and $-\operatorname{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{2n})$ coincide, up to an arbitrarily small error term. Neglecting the error terms in what follows, we thus can arrange that

$$|\nu(\hat{\mathscr{M}})| \leq 2 \sum_{i=1}^{2n} |\tilde{\lambda}_i| \qquad \text{and} \qquad \sum_{i=1}^{2n} \tilde{\lambda}_i^2 \leq 2|A[\nu](0)|\,.$$

Employing Jensen's inequality, we conclude that

$$|\nu(\hat{\mathscr{M}})|^2 \leq 4 \left(\sum_{i=1}^{2n} |\tilde{\lambda}_i| \right)^2 \leq 8n \sum_{i=1}^{2n} |\tilde{\lambda}_i|^2 \leq 16n |A[\nu](0)|\,.$$

Applying (5.4.6), the boundedness constraint gives rise to the desired estimate

$$|\nu(\hat{\mathscr{M}})| < 4\sqrt{n\,(C+1)} =: f\,,$$

which completes the proof. □

COROLLARY 5.4.7. *Let $C > 0$ and assume that $(d\nu^{(k)})_{k \in \mathbb{N}}$ is a sequence in $\mathfrak{Ndm}$ such that*

$$\mathcal{T}(\nu^{(k)}) \leq C \qquad \textit{for all } k \in \mathbb{N}$$

Then there is a sequence $(U_k)_{k\in\mathbb{N}}$ of unitary operators on V (with respect to $\prec . \mid . \succ$) as well as a positive constant $C' > 0$ such that $d\|U_k \nu^{(k)} U_k^{-1}\| \leq C'$ for all $k \in \mathbb{N}$ (where $d\|\cdot\|$ denotes the total variation according to Definition 5.3.4).

PROOF. Combine Lemma 5.4.6 and Lemma 5.4.5. □

One possible strategy in order to prove the existence of minimizers is to show that the resulting subsequences according to Lemma 5.4.4 and Lemma 5.4.5 are uniformly tight as well. In this case, by applying Prohorov's theorem one immediately obtains the existence of a weakly converging subsequence, thus giving rise to the desired minimizer. Unfortunately, this strategy does not seem to work for the following reason: We do not know whether the resulting sequences contain uniformly tight subsequences. This shortcoming makes it necessary to proceed in a different way in the spirit of [**66**]. More explicitly, our goal is to first derive the EL equations and then to suitable apply them in order to deduce the existence of a global minimizer. It is precisely the objective in the remainder of this chapter to derive the EL equations under suitable assumptions. For technical simplicity, in what follows we shall restrict attention to the subsequent side conditions.

DEFINITION 5.4.8. *Given $c, f > 0$, we call $d\nu \in \mathfrak{Ndm}$ **admissible** if and only if*

$$\mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = c \qquad \text{and} \qquad \mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f\,. \tag{5.4.7}$$

Introducing the corresponding subset $N \subset \mathfrak{Ndm}$ by

$$N := \{d\nu \in \mathfrak{Ndm} : \text{property (5.4.7) is satisfied}\}\,,$$

*the **causal variational principle in the homogeneous setting** is to*

$$\text{minimize } \mathcal{S}(\nu) \text{ by varying } d\nu \text{ in } N \subset \mathfrak{Ndm}\,. \tag{5.4.8}$$

In agreement with [**168**, Definition 43.4], we define a minimizer for $\mathcal{S}$ as follows:

DEFINITION 5.4.9. *A negative definite measure $d\nu \in N$ is said to be a **minimizer** for the causal variational principle (5.4.8) if and only if*

$$\mathcal{S}(\tilde{\nu}) \geq \mathcal{S}(\nu) \qquad \text{for all } d\tilde{\nu} \in N\,.$$

5.5. Analysis of the Causal Action in the Homogeneous Setting

In order to eventually develop the existence theory of minimizers of causal variational principles in the homogeneous setting (see Definition 5.4.8), our strategy is basically to derive and essentially exploit the EL equations corresponding to minimizers on compact subsets which exist in view of [**56**, Theorem 4.2]. Accordingly, in what follows we shall always assume $d\nu$ is a negative definite measure in $\mathfrak{Ndm}$ which is a minimizer in the sense of Definition 5.4.9 (possibly restricted to a compact subset of momentum space). In order to specify the minimizer $d\nu$ in more detail and to realize our basic strategy, the remainder of this chapter is devoted to derive the corresponding EL equations. In a first step, the goal of this section is to analyze important properties of the causal action in the homogeneous setting. To begin with, we prove Fréchet differentiability of the closed chain (§5.5.1) and compute the first variation of the causal action (§5.5.2). Afterwards, we shall derive some useful results (§5.5.3) which allow us to express the first variation of the causal action in terms of momentum space (§5.5.4).

Unless specified otherwise, for given $n \in \mathbb{N}$ we always assume that $(V, \prec . \mid . \succ)$ is a finite-dimensional inner product space of dimension $2n$. By $\mathscr{M} \simeq \mathbb{R}^4$ we denote

Minkowski space and let $\hat{\mathscr{M}}$ be momentum space $\hat{\mathscr{M}}$. Moreover, by $\mathfrak{Ndm}$ and $\mathfrak{Ovm}$ we indicate the class of negative definite and operator-valued measures on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$, respectively.

For convenience, let us briefly recall some important definitions required in the following. Further details can be found in [**168**].

REMARK 5.5.1. *Let X be a real locally convex space, and let $F : D(F) \subset X \to \mathbb{R}$ be a functional (where $D(F)$ denotes the domain of F). Then the **first variation** of F at a given fixed interior point $u_0 \in D(F)$ in the direction $h \in X$ is given by*

$$\delta F(u_0; h) := \frac{dF(u_0 + th)}{dt}\Big|_{t=0} .$$

*Given $u_0 \in X$, the **Gâteaux derivative** of F at u_0, denoted by $F'(u_0) \in X^*$, exists if and only if $\delta F(u_0; h)$ exists and $h \mapsto \delta F(u_0; h)$ is continuous and linear. In this case,*

$$\delta F(u_0; h) = \langle F'(u_0), h\rangle \qquad \text{for all } h \in X$$

*(where $\langle ., .\rangle$ denotes duality). Whenever the Gâteaux derivative $F'(u)$ is defined for all u in a neighborhood $U(u_0)$ of u_0, and if $F' : U(u_0) \subset X \to X^*$ is continuous at u_0, then $F'(u_0)$ coincides with the **Fréchet derivative** of F (for details see* [**168**, Section 40.1]*).*

5.5.1. Properties of the Closed Chain. The purpose of this subsection is to show that the closed chain can be regarded as a Fréchet differentiable mapping on the class of operator-valued measures (Lemma 5.5.3). To clarify notation, by $\|\cdot\|_{\mathrm{L}(V)}$ we denote the operator norm on $\mathrm{L}(V)$, and $d\|\cdot\|$ indicates the total variation (5.3.3) on $\mathfrak{Ovm}$.

Let us begin with the following remark.

REMARK 5.5.2. *Choosing a gauge, for any $d\omega \in \mathfrak{Ovm}$ and $\xi \in \mathscr{M}$ we can represent the fermionic projector $P[\omega](\xi) \in \mathrm{L}(V)$ by a complex matrix. Whenever $d\|\tilde{\omega} - \omega\| < \varepsilon$ for operator-valued measures $d\omega, d\tilde{\omega} \in \mathfrak{Ovm}$, for arbitrary $\xi \in \mathscr{M}$ we arrive at*

$$\|P[\tilde{\omega}](\xi) - P[\omega](\xi)\|_{\mathrm{L}(V)} = \left\| \int_{\hat{\mathscr{M}}} e^{ip\xi}\, d\tilde{\omega}(p) - \int_{\hat{\mathscr{M}}} e^{ip\xi}\, d\omega(p) \right\|_{\mathrm{L}(V)} \leq d\,\|\tilde{\omega} - \omega\| < \varepsilon$$

by linearity of the integrator (cf. [**109**, Definition A.29]*). Therefore, small perturbations of $d\omega$ according to $d\|\tilde{\omega} - \omega\| < \varepsilon$ affect the closed chain $A(\xi)$ in the following way,*

$$\begin{aligned} \|A[\tilde{\omega}](\xi) - A[\omega](\xi)\|_{\mathrm{L}(V)} &= \|P[\tilde{\omega}](\xi)\, P[\tilde{\omega}](-\xi) - P[\omega](\xi)\, P[\omega](-\xi)\|_{\mathrm{L}(V)} \\ &\leq \varepsilon \Big(\|P[\tilde{\omega}](-\xi) - P[\omega](-\xi)\|_{\mathrm{L}(V)} + \|P[\omega](-\xi)\|_{\mathrm{L}(V)} \Big) + \varepsilon\, \|P[\omega](\xi)\|_{\mathrm{L}(V)} \\ &< \varepsilon^2 + \varepsilon\, \|P[\omega](-\xi)\|_{\mathrm{L}(V)} + \varepsilon\, \|P[\omega](\xi)\|_{\mathrm{L}(V)} \leq \varepsilon^2 + 2\varepsilon\, d\|\omega\| \end{aligned}$$

for any $\xi \in \mathscr{M}$. Since $\xi \in \mathscr{M}$ is arbitrary, the previous inequalities imply that

$$\begin{aligned} &\sup_{\xi \in \mathscr{M}} \|P[\tilde{\omega}](\xi) - P[\omega](\xi)\|_{\mathrm{L}(V)} < \varepsilon\,, \\ &\sup_{\xi \in \mathscr{M}} \|A[\tilde{\omega}](\xi) - A[\omega](\xi)\|_{\mathrm{L}(V)} \leq \varepsilon^2 + 2\varepsilon\, d\|\omega\| \end{aligned} \tag{5.5.1}$$

whenever $d\omega, d\tilde{\omega} \in \mathfrak{Ovm}$ satisfying $d\|\tilde{\omega} - \omega\| < \varepsilon$.

We employ Remark 5.5.2 to prove that the closed chain is Fréchet differentiable:

LEMMA 5.5.3. *For any $\xi \in \mathscr{M}$, the mapping*

$$A[\cdot](\xi) : \mathfrak{Ovm} \to \mathrm{L}(V)\,, \qquad d\omega \mapsto A[\omega](\xi)$$

is Fréchet differentiable, and its Fréchet derivative coincides with its first variation, denoted by $\delta A(\xi)$. Moreover, $\delta A(\xi)$ is self-adjoint in the sense that $\delta A(\xi)^* = \delta A(\xi)$ (with respect to the spin scalar product $\prec . \,|\, . \succ$).

PROOF. For any $\xi \in \mathscr{M}$, the mapping $P[\cdot](\xi) : \mathfrak{Ovm} \to \mathrm{L}(V)$ is linear,

$$P[\omega + \tau\eta](\xi) = \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d\big(\omega + \tau\eta\big)(k) = P[\omega](\xi) + \tau P[\eta](\xi)$$

for all $d\omega, d\eta \in \mathfrak{Ovm}$ and $\tau \in \mathbb{R}$. Thus we obtain

$$\begin{aligned} A[\omega + \tau\eta](\xi) - A[\omega](\xi) &= \big(P[\omega](\xi) + \tau P[\eta](\xi)\big)\big(P[\omega](-\xi) + \tau P[\eta](-\xi)\big) - A[\omega](\xi) \\ &= \tau P[\eta](\xi)\, P[\omega](-\xi) + \tau P[\omega](\xi)\, P[\eta](-\xi) + \tau^2 A[\eta](\xi)\,, \end{aligned}$$

implying that

$$\lim_{\tau\to 0} \frac{A[\omega + \tau\eta](\xi) - A[\omega](\xi)}{\tau} = P[\eta](\xi)\, P[\omega](-\xi) + P[\omega](\xi)\, P[\eta](-\xi)\,.$$

Given $\xi \in \mathscr{M}$ and $d\omega \in \mathfrak{Ovm}$, the mapping $A'[\omega](\xi) : \mathfrak{Ovm} \to \mathrm{L}(V)$ defined by

$$\mathfrak{Ovm} \ni d\eta \mapsto \big[A'[\omega](\xi)\big]\,(\eta) := P[\eta](\xi)\, P[\omega](-\xi) + P[\omega](\xi)\, P[\eta](-\xi)$$

is linear and continuous in the sense that for any $\varepsilon > 0$ there is $\delta > 0$ such that

$$\|A'[\omega](\xi) - A'[\eta](\xi)\|_{\mathrm{L}(V)} < \varepsilon \qquad \text{whenever} \qquad d\|\omega - \eta\| < \delta$$

by virtue of (5.5.1). From this we infer that $A'[\omega](\xi) : \mathfrak{Ovm} \to \mathrm{L}(V)$ is the Gâteaux derivative of the closed chain at $d\omega$ (cf. [**168**, §40]). Since $\mathfrak{Ovm}$ is a normed vector space (see Lemma 5.3.5) and A' is continuous, the closed chain is Fréchet differentiable; moreover, its Fréchet derivative coincides with its first variation (see [**168**, §40]),

$$\delta A[\omega; \eta](\xi) = \langle A'[\omega](\xi), \eta \rangle \qquad \text{for all } d\eta \in \mathfrak{Ovm}^*$$

(where $\mathfrak{Ovm}^*$ denotes the dual space of $\mathfrak{Ovm}$). The last statement follows from the fact that $P[\omega](\xi)^* = P[\omega](-\xi)$ in view of (5.3.6) for any $d\omega \in \mathfrak{Ovm}$ (with respect to the spin scalar product $\prec . \,|\, . \succ$). □

5.5.2. First Variation of the Causal Action. After these preparations, we can deal with the first variation of the causal action $\mathcal{S}$. For any $d\omega, d\eta \in \mathfrak{Ovm}$, we assume that

$$\delta\mathcal{S}(\omega; \eta) = \int_{\mathscr{M}} \delta\mathcal{L}[\omega; \eta](\xi)\, d\mu(\xi) \tag{5.5.2}$$

(for necessary and sufficient conditions we refer to [**95**]), which motivates to analyze variations of the Lagrangian in more detail. To this end, in what follows we consider the Lagrangian (5.3.8) as a mapping

$$\mathcal{L} : \mathrm{L}(V) \simeq \mathbb{C}^{2n\times 2n} \to \mathbb{R}_0^+\,, \qquad A \mapsto \mathcal{L}[A] = \big|A^2\big| - \frac{1}{2n}|A|^2\,. \tag{5.5.3}$$

In order to calculate the first variation of the causal action (and henceforth to infer the corresponding Euler-Lagrange equations), we make the following assumption:

ASSUMPTION 5.5.4. *Unless otherwise specified, in the remainder of this chapter we assume that* (5.5.2) *holds for all* $d\omega, d\eta \in \mathfrak{Dvm}$. *Furthermore, we assume that the Lagrangian* (5.5.3) *is a continuously differentiable mapping* $\mathrm{L}(V) \ni A \mapsto \mathcal{L}[A] \in \mathbb{R}$ *with respect to the real and imaginary part of* $A_{\alpha,\beta}$ *for all* $\alpha, \beta = 1, \ldots, 2n$ *and* $A \in \mathrm{L}(V)$.[3]

In order to explain this assumption, for any $\xi \in \mathscr{M}$ we rewrite $A(\xi)$ as

$$A(\xi) = P(\xi)\, P(-\xi) \overset{(5.3.6)}{=} P(\xi)\, P(\xi)^* \in \mathbb{C}^{2n \times 2n} .$$

Thus for any $\alpha, \beta \in \{1, \ldots, 2n\}$, the components of the closed chain are given by

$$\begin{aligned} \big(A(-\xi)\big)_{\alpha,\beta} &= \langle e_\alpha \mid A(-\xi)\, e_\beta \rangle_{\mathbb{C}^{2n}} \overset{(5.2.1)}{=} \prec e_\alpha \mid SP(-\xi)\, P(-\xi)^*\, e_\beta \succ \\ &= \langle P(\xi)\, S e_\alpha \mid SP(\xi)\, e_\beta \rangle_{\mathbb{C}^{2n}} = \sum_{j,k=1}^{2n} \overline{\big(P(\xi)\, S e_\alpha\big)_j}\, \big(SP(\xi)\, e_\beta\big)_k . \end{aligned}$$

Making use of $S e_\alpha = s_\alpha e_\alpha$ for all $\alpha = 1, \ldots, 2n$ in view of (5.2.1), we obtain

$$\begin{aligned} \big(A(-\xi)\big)_{\alpha,\beta} &= s_\alpha \sum_{j,k=1}^{2n} \overline{P(\xi)_{j,\alpha}}\, \big(SP(\xi)\big)_{k,\beta} \\ &= s_\alpha\, s_\beta \sum_{j,k=1}^{2n} \big(\operatorname{Re} P(\xi)_{j,\alpha} - i \operatorname{Im} P(\xi)_{j,\alpha}\big)\, \big(\operatorname{Re} P(\xi)_{k,\beta} + i \operatorname{Im} P(\xi)_{k,\beta}\big) \end{aligned}$$

for all $\alpha, \beta \in \{1, \ldots, 2n\}$ and $\xi \in \mathscr{M}$. Accordingly, for arbitrary $\xi \in \mathscr{M}$ the closed chain $A(-\xi)$ may be regarded as a function depending on

$$\operatorname{Re} P(\xi)_{\alpha,\beta} \qquad \text{and} \qquad \operatorname{Im} P(\xi)_{\alpha,\beta} \qquad \text{for all } \alpha, \beta = 1, \ldots, 2n ;$$

this explains our assumption that the Lagrangian is continuously differentiable with respect to $\operatorname{Re} P(\xi)_{\alpha,\beta}$ and $\operatorname{Im} P(\xi)_{\alpha,\beta}$ for all $\alpha, \beta = 1, \ldots, 2n$. Moreover, making use of Proposition 5.4.1, we conclude that

$$\Big|\big(P(\xi)\, P(\xi)^*\big)^2\Big| = \Big|\big(P(-\xi)\, P(-\xi)^*\big)^2\Big| \qquad \text{and} \qquad |P(\xi)\, P(\xi)^*|^2 = |P(-\xi)\, P(-\xi)^*|^2$$

for all $\xi \in \mathscr{M}$. As a consequence, for any $\xi \in \mathscr{M}$ we obtain the composed mapping

$$\mathcal{L} \circ A[\cdot](\xi) : \mathfrak{Dvm} \to \mathbb{R}_0^+ , \qquad d\omega \mapsto \Big|\big(P[\omega](\xi)\, P[\omega](\xi)^*\big)^2\Big| - \frac{1}{2n} |P[\omega](\xi)\, P[\omega](\xi)^*|^2 ,$$

depending on $\operatorname{Re} P(\xi)_{\alpha,\beta}$ and $\operatorname{Im} P(\xi)_{\alpha,\beta}$ for all $\alpha, \beta = 1, \ldots, 2n$.

The above considerations reveal that in order to analyze the first variation of the causal action one is led to differentiate the mapping $\mathcal{L} \circ A[\cdot](\xi) : \mathfrak{Dvm} \to \mathbb{R}_0^+$. More precisely, the chain rule for Gâteaux derivatives (see e.g. [**88**, Theorem 3.3.4]) yields

$$(\mathcal{L} \circ A)'(\omega; \eta) = \langle \mathcal{L}'(A(\omega)), A'(\omega; \eta) \rangle \qquad \text{for all } d\omega, d\eta \in \mathfrak{Dvm} .$$

From this we deduce that, for any $\xi \in \mathscr{M}$ and all $d\omega, d\eta \in \mathfrak{Dvm}$,

$$\delta(\mathcal{L} \circ A[\cdot](\xi))[\omega; \eta] = \langle \mathcal{L}'(A[\omega](\xi)), \delta A[\omega; \eta](\xi) \rangle .$$

We now make use of the fact that in finite dimensions the Fréchet derivative coincides with the usual derivative and apply [**23**, Chapter 3, Theorem 5] (which states that the Fréchet derivative of a continuously differentiable function between Euclidean spaces is

[3] In the recent article [**67**] it is shown that the Lagrangian $\mathcal{L}$ can be differentiated along smooth curves (for details see [**67**, §5.3]).

given by its Jacobian). As a convenient notation, let us define the following first-order linear partial differential operators in analogy to the usual Wirtinger derivatives,

$$\frac{\partial}{\partial z_j} := \frac{\partial}{\partial x_j} - i \frac{\partial}{\partial y_j} \qquad \text{and} \qquad \frac{\partial}{\partial \overline{z}_j} := \frac{\partial}{\partial x_j} + i \frac{\partial}{\partial y_j}\,, \tag{5.5.4}$$

where x_j, y_j are the underlying real coordinates of $\mathbb{C}^{2n}$ for all $j \in \{1, \ldots, 2n\}$ such that $z_j = x_j + i y_j$ (see [**84**]). Thus in the spirit of (5.5.4) we introduce the notation

$$\frac{\partial \mathcal{L}}{\partial (A[\omega](\xi))_{\beta,\alpha}} := \frac{\partial \mathcal{L}}{\partial \operatorname{Re} P[\omega](\xi)_{\beta,\alpha}} - i \frac{\partial \mathcal{L}}{\partial \operatorname{Im} P[\omega](\xi)_{\beta,\alpha}}$$

for all $\alpha, \beta = 1, \ldots, 2n$. Accordingly, the Fréchet derivative of $\mathcal{L}(A[\omega](\xi))$ reads

$$(\mathcal{L}'(A[\omega](\xi)))_{\alpha,\beta=1,\ldots,2n} = \left(\frac{\partial \mathcal{L}}{\partial (A[\omega](\xi))_{\beta,\alpha}} \right)_{\alpha,\beta=1,\ldots,2n}.$$

Abbreviating the functional derivative of the Lagrangian (cf. [**59**, eq. (3.5.15)]) by

$$\mathcal{M}_{\alpha,\beta} := \frac{\partial \mathcal{L}}{\partial (A[\omega](\xi))_{\beta,\alpha}} = \frac{\partial \mathcal{L}}{\partial \operatorname{Re} P[\omega](\xi)_{\beta,\alpha}} - i \frac{\partial \mathcal{L}}{\partial \operatorname{Im} P[\omega](\xi)_{\beta,\alpha}}$$

for all $\alpha, \beta = 1, \ldots, 2n$ and defining

$$\mathcal{M}(\xi) \equiv \mathcal{M}[A[\omega](\xi)] = \left(\frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial (A[\omega](\xi))_{\beta,\alpha}} \right)_{\alpha,\beta=1,\ldots,2n},$$

for arbitrary $\xi \in \mathscr{M}$ and $d\omega, d\eta \in \mathfrak{Obm}$ we obtain the expression

$$\begin{aligned} \delta(\mathcal{L} \circ A[\cdot](\xi))[\omega;\eta] &= \sum_{\alpha,\beta=1}^{2n} \mathcal{M}[A[\omega](\xi)]_{\alpha,\beta}\, \delta A[\omega;\eta](\xi)_{\beta,\alpha} \\ &= \sum_{\alpha=1}^{2n} \left(\mathcal{M}[A[\omega](\xi)]\, \delta A[\omega;\eta](\xi) \right)_{\alpha,\alpha} = \operatorname{Tr}_V \left(\mathcal{M}[A[\omega](\xi)]\, \delta A[\omega;\eta](\xi) \right). \end{aligned}$$

Moreover, applying the product rule

$$\delta A[\omega;\eta](\xi) = \delta P[\omega;\eta](\xi)\, P[\omega](-\xi) + P[\omega](\xi)\, \delta P[\omega;\eta](-\xi)$$

and making use of the fact that the trace is cyclic, we consequently arrive at

$$\delta(\mathcal{L} \circ A[\cdot](\xi))[\omega;\eta] = \operatorname{Tr}_V \left(\mathcal{M}(\xi)\, P(\xi)\, \delta P(-\xi) \right) + \operatorname{Tr}_V \left(P(-\xi)\, \mathcal{M}(\xi)\, \delta P(\xi) \right) \tag{5.5.5}$$

with $\mathcal{M} = \mathcal{M}[\omega]$, $P = P[\omega]$ and $\delta P = \delta P[\omega;\eta]$ for all $\xi \in \mathscr{M}$ and $d\omega, d\eta \in \mathfrak{Obm}$. Our goal is to further simplify the expression (5.5.5) in the fashion of [**53**, Section 5.2] by applying Proposition 5.5.6 below. In preparation, let us state the following remark.

REMARK 5.5.5. *As is verified by direct computation, a matrix $A \in \mathrm{L}(V)$ is symmetric with respect to the spin scalar product $\prec . \mid . \succ$ if and only if*

$$\overline{A_{\alpha,\beta}} = s_\alpha\, s_\beta\, A_{\beta,\alpha} \qquad \text{for all } \alpha, \beta \in \{1, \ldots, 2n\},$$

where the signature matrix S is given by (5.2.2) *with diagonal entries*

$$s_\alpha := \langle e_\alpha \mid S e_\alpha \rangle = \begin{cases} +1 & \text{if } \alpha \in \{1, \ldots, n\} \\ -1 & \text{if } \alpha \in \{n+1, \ldots, 2n\}, \end{cases}$$

and $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ denotes a pseudo-orthonormal basis of V satisfying (5.2.1).

PROPOSITION 5.5.6. *Assume that the Lagrangian is continuous differentiable. Then for any $\xi \in \mathscr{M}$, its functional derivative $\mathcal{M}(\xi)$ is symmetric, i.e. $\mathcal{M}(\xi)^* = \mathcal{M}(\xi)$ (with respect to the spin scalar product).*

PROOF. To begin with, for all $\xi \in \mathscr{M}$ and $d\omega \in \mathfrak{Dvm}$ we are given

$$\mathcal{M}[A[\omega](\xi)] = \left(\frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial \operatorname{Re} P[\omega](\xi)_{\beta,\alpha}} - i \frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial \operatorname{Im} P[\omega](\xi)_{\beta,\alpha}} \right)_{\alpha,\beta=1,\ldots,2n} .$$

Assuming that the mapping $\mathrm{L}(V) \ni A \mapsto \mathcal{L}[A] \in \mathbb{R}_0^+$ is continuously differentiable with respect to the real variables $\operatorname{Re} A_{\alpha,\beta}$ and $\operatorname{Im} A_{\alpha,\beta}$ for all $\alpha, \beta \in \{1, \ldots, 2n\}$, in virtue of [**98**, §0.1] we deduce that (also cf. [**53**, Section 5.4])

$$\overline{\left(\frac{\partial \mathcal{L}}{\partial A_{\alpha,\beta}} \right)} = \frac{\partial \overline{\mathcal{L}}}{\partial \overline{A}_{\alpha,\beta}} \qquad \text{for all } \alpha, \beta \in \{1, \ldots, 2n\} .$$

Since $\overline{\mathcal{L}} = \mathcal{L}$, the complex conjugate of the functional derivative reads

$$\overline{\mathcal{M}[A[\omega](\xi)]}_{\alpha,\beta} = \overline{\left(\frac{\partial \mathcal{L}}{\partial (A[\omega](\xi))_{\beta,\alpha}} \right)} \overset{(5.5.4)}{=} \frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial \operatorname{Re} P[\omega](\xi)_{\beta,\alpha}} + i \frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial \operatorname{Im} P[\omega](\xi)_{\beta,\alpha}} .$$

In view of $P(\xi)^* = P(-\xi)$, for all $\alpha, \beta \in \{1, \ldots, 2n\}$ and $\xi \in \mathscr{M}$ we are given

$$\begin{aligned} \overline{P(\xi)_{\alpha,\beta}} &= \overline{\langle e_\alpha \mid P(\xi)\, e_\beta \rangle} = \langle P(\xi)\, e_\beta \mid e_\alpha \rangle = \prec P(\xi)\, e_\beta \mid S e_\alpha \succ \\ &= s_\alpha \prec e_\beta \mid P(-\xi)\, e_\alpha \succ = s_\alpha\, s_\beta\, \langle e_\beta \mid P(-\xi)\, e_\alpha \rangle = s_\alpha\, s_\beta\, P(-\xi)_{\beta,\alpha} . \end{aligned}$$

As a consequence, for all $\xi \in \mathscr{M}$ and $\alpha, \beta \in \{1, \ldots, 2n\}$ we obtain

$$\operatorname{Re} P(-\xi)_{\beta,\alpha} = s_\alpha s_\beta \operatorname{Re} P(\xi)_{\alpha,\beta} , \qquad \operatorname{Im} P(-\xi)_{\beta,\alpha} = -s_\alpha\, s_\beta \operatorname{Im} P(\xi)_{\alpha,\beta} .$$

Making use of the fact that (cf. [**53**, eq. (3.5.7)])

$$\mathcal{L}[A(\xi)] = \mathcal{L}[P(\xi)\, P(-\xi)] = \mathcal{L}[P(-\xi)\, P(\xi)] = \mathcal{L}[A(-\xi)] \qquad \text{for all } \xi \in \mathscr{M} ,$$

we deduce that, for all $\xi \in \mathscr{M}$ and $d\omega \in \mathfrak{Dvm}$,

$$\mathcal{M}[A[\omega](\xi)] = \mathcal{M}[A[\omega](-\xi)] .$$

As a result, for the transposed matrix we obtain the equality

$$\begin{aligned} s_\alpha\, s_\beta\, \mathcal{M}[A[\omega](\xi)]_{\beta,\alpha} &= s_\alpha\, s_\beta \left(\frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial \operatorname{Re} P[\omega](\xi)_{\alpha,\beta}} - i \frac{\partial \mathcal{L}[A[\omega](\xi)]}{\partial \operatorname{Im} P[\omega](\xi)_{\alpha,\beta}} \right) \\ &= \frac{\partial \mathcal{L}[A[\omega](-\xi)]}{\partial \operatorname{Re} P[\omega](-\xi)_{\beta,\alpha}} + i \frac{\partial \mathcal{L}[A[\omega](-\xi)]}{\partial \operatorname{Im} P[\omega](-\xi)_{\beta,\alpha}} = \overline{\mathcal{M}[A[\omega](\xi)]}_{\alpha,\beta} . \end{aligned}$$

In view of Remark 5.5.5 we conclude that $\mathcal{M}(\xi)^* = \mathcal{M}(\xi)$ as desired. □

We now employ this result in order to simplify the expression (5.5.5) similarly to [**53**]. More precisely, introducing the operator

$$R(\xi) := \frac{1}{2} \mathcal{M}(\xi)\, P(\xi) \qquad \text{for all } \xi \in \mathscr{M}$$

and applying Proposition 5.5.6, the first variation (5.5.5) can be written as

$$\delta(\mathcal{L} \circ A[\cdot](\xi))[\omega; \eta] = 2 \operatorname{Tr}_V \Big(R(\xi)\, \delta P(-\xi) \Big) + 2 \operatorname{Tr}_V \Big(R(\xi)^*\, \delta P(\xi) \Big)$$

for all $\xi \in \mathscr{M}$ and $d\omega, d\eta \in \mathfrak{Dvm}$. For all $\alpha, \beta \in \{1, \ldots, 2n\}$, it is easily verified that

$$\overline{\delta P(\xi)_{\alpha,\beta}} = s_\alpha\, s_\beta\, \delta P(-\xi)_{\beta,\alpha} , \qquad \mathcal{M}(\xi)_{\alpha,\beta} = s_\alpha\, s_\beta\, \mathcal{M}(\xi)_{\beta,\alpha}$$

for all $\xi \in \mathscr{M}$. Furthermore, a straightforward calculation yields

$$\overline{\mathrm{Tr}_V\big(R(\xi)\,\delta P(-\xi)\big)} = \mathrm{Tr}_V\big(R(\xi)^*\,\delta P(\xi)\big)\,,$$

where the adjoint of $R(\xi)$ with respect to the spin scalar product reads

$$R(\xi)^* = \frac{1}{2} P(-\xi)\,\mathcal{M}(\xi) \qquad \text{for all } \xi \in \mathscr{M}\,. \tag{5.5.6}$$

Thus for any $\xi \in \mathscr{M}$ and $d\omega, d\eta \in \mathfrak{Dvm}$, the first variation (5.5.5) takes the form

$$\delta(\mathcal{L} \circ A[\cdot](\xi))[\omega;\eta] = 4\ \mathrm{Re}\,\mathrm{Tr}_V\big(R(\xi)\,\delta P(-\xi)\big)\,.$$

In order to get rid of the real part, it is convenient to introduce the operator

$$Q(\xi) := \frac{1}{2}\big(R(\xi) + R(-\xi)^*\big) = \frac{1}{4}\big(\mathcal{M}(\xi)\,P(\xi) + P(\xi)\,\mathcal{M}(-\xi)\big) \tag{5.5.7}$$

for every $\xi \in \mathscr{M}$ (cf. [**53**, eq. (5.2.6)]). Applying (5.5.5) as well as (5.5.6) and making use of the symmetry $\xi \leftrightarrow -\xi$, we finally obtain

$$\begin{aligned}
\int_{\mathscr{M}} \delta\mathcal{L}\,[A[\omega;\eta](\xi)]\ d\mu(\xi) &= \int_{\mathscr{M}} \mathrm{Tr}_V\Big(\mathcal{M}\big[A[\omega](\xi)\big]\,\delta A[\omega;\eta](\xi)\Big)\,d\mu(\xi) \\
&= 2\int_{\mathscr{M}} \left[\mathrm{Tr}_V\left(\frac{1}{2}\mathcal{M}(\xi)\,P(\xi)\,\delta P(-\xi)\right) + \mathrm{Tr}_V\left(\frac{1}{2}P(-\xi)\,\mathcal{M}(\xi)\,\delta P(\xi)\right)\right] d\mu(\xi) \\
&= 2\int_{\mathscr{M}} \mathrm{Tr}_V\big(R(\xi)\,\delta P(-\xi)\big)\,d\mu(\xi) + 2\int_{\mathscr{M}} \mathrm{Tr}_V\big(R(-\xi)^*\,\delta P(-\xi)\big)\,d\mu(\xi) \\
&= 4\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q(\xi)\,\delta P(-\xi)\big)\,d\mu(\xi)\,.
\end{aligned}$$

As a result, the first variation of the causal action has the compact form

$$\delta\mathcal{S}(\omega;\eta) = 4\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q[\omega](\xi)\,\delta P[\omega;\eta](-\xi)\big)\,d\mu(\xi) \tag{5.5.8}$$

for all $d\omega, d\eta \in \mathfrak{Dvm}$.

5.5.3. Technical Preliminaries. In what follows, our goal is to rewrite (5.5.8) in terms of momentum space. To this end, we require some technical preliminaries. Let us start with the next proposition (where $m \geq 1$ is a positive integer).

PROPOSITION 5.5.7. *Let $f \in C_c(\mathbb{R}^m)$ and $\varphi \in C_c^\infty(\mathbb{R}^m)$ mit $\varphi \geq 0$. Moreover, let $d\rho$ be a positive finite measure on $\mathbb{R}^m$ and for $\varepsilon > 0$ let $(\varphi_n)_{n\in\mathbb{N}}$ be the Dirac sequence defined by*

$$\varphi_n(x) := \varepsilon_n^{-m}\varphi\left(\frac{x}{\varepsilon_n}\right) \qquad \textit{with } \varepsilon_n = 1/n \textit{ for each } n \in \mathbb{N}\,.$$

Then there is a subsequence $(\varphi_{n_k})_{k\in\mathbb{N}}$ with the property that

$$\lim_{k\to\infty} \int_{\mathbb{R}^m} (\varphi_{n_k} * f)(y)\,d\rho(y) = \int_{\mathbb{R}^m} f(y)\,d\rho(y)\,.$$

PROOF. Introducing the functions $f_n := \varphi_n * f$ for every $n \in \mathbb{N}$, we basically apply Lebesgue's dominated convergence theorem [**16**, Theorem 2.8.1 and Corollary 2.8.6] to the sequence $(f_n)_{n\in\mathbb{N}}$ with respect to the measure $d\rho$. According to [**5**, Satz 2.14] we have $f_n \to f$ in $L^2(\mathbb{R}^m)$ as $n \to \infty$, and making use of $f \in C_c(\mathbb{R}^m)$ we infer that

$$\infty > \int_{\mathbb{R}^m} |f|^2\,d\mu = \lim_{n\to\infty} \int_{\mathbb{R}^m} |f_n|^2\,d\mu\,.$$

In order to show that $|f_n(x)| \leq \Phi(x)$ for some ρ-integrable function Φ for all $n \in \mathbb{N}$ and ρ-a.e. $x \in \mathbb{R}^m$, we first note that $\operatorname{supp} \varphi_n \subset \operatorname{supp} \varphi$ for all $n \in \mathbb{N}$. Thus in view of [**72**, Theorem 8.6] there is $K \subset \mathbb{R}^m$ compact such that $\operatorname{supp} f_n \subset K$ for all $n \in \mathbb{N}$. On the other hand, we are given $f_n \in C^\infty(\mathbb{R}^m)$ in virtue of [**5**, Satz 2.12]. From this we deduce that $f_n \in C^\infty_c(\mathbb{R}^m) \subset L^1(\mathbb{R}^m; d\rho)$ for all $n \in \mathbb{N}$. Since $\varphi \geq 0$, we are given

$$\|\varphi_n\|_{L^1(\mathbb{R}^m)} = \int_{\mathbb{R}^m} \varphi_n \, d\mu = 1 \qquad \text{for all } n \in \mathbb{N}\,.$$

Applying [**72**, Young's Inequality 8.7], we obtain

$$\|f_n\|_{L^\infty(\mathbb{R}^m)} \leq \|\varphi_n\|_{L^1(\mathbb{R}^m)} \, \|f\|_{L^\infty(\mathbb{R}^m)} = \|f\|_{L^\infty(\mathbb{R}^m)} \qquad \text{for all } n \in \mathbb{N}\,.$$

Since $f \in C_c(\mathbb{R}^m)$, there is $C > 0$ such that $\sup_{x \in \mathbb{R}^m} |f(x)| \leq C$. As a consequence,

$$|f_n(x)| \leq C \chi_K \qquad \text{for all } n \in \mathbb{N} \text{ and } \rho\text{-a.e. } x \in \mathbb{R}^m$$

(cf. [**6**, Section X.4]), where $C\chi_K \in L^1(\mathbb{R}^m; d\rho)$.

In order for Lebesgue's dominated convergence theorem [**16**, Theorem 2.8.1] to apply we require that $f_n(x) \to f(x)$ for ρ-almost all $x \in \mathbb{R}^m$. To this end, we essentially make use of [**104**, Corollary 16]. Let us recall that we are given a sequence $(f_n)_{n \in \mathbb{N}}$ of compactly supported functions (each of which thus vanishes at infinity) as well as a function $f \in C_c(\mathbb{R}^m)$ (which also vanishes at infinity) such that $\|f_n - f\|_{L^2(\mathbb{R}^m)} \to 0$ as $n \to \infty$ (for details see [**5**, Satz 2.14]), implying that $f_n(x) \to f(x)$ for μ-almost every $x \in \mathbb{R}^m$. Denoting the corresponding null set by E, we conclude that $f_n(x) \to f(x)$ for all $x \in \tilde{K} := K \setminus E$. However, in order to apply [**104**, Corollary 16] we need to show that $|f_n(x) - f(x)| \to 0$ *monotonically*. To this end we make use of the fact that $\tilde{K} \subset \mathbb{R}^m$ is separable (cf. [**4**, Lemma 3.5]) and let K' be a countable dense subset of $\tilde{K}$. We then proceed iteratively as follows: Given $x_1 \in K'$, we let $(f_{1,n_k})_{k \in \mathbb{N}}$ a subsequence such that $|f_{1,n_k}(x_1) \to f(x_1)| \to 0$ monotonically. Next, given $(f_{i,n_k})_{k \in \mathbb{N}}$ for some $i \in \mathbb{N}$, we choose a subsequence $(f_{i+1,n_k})_{k \in \mathbb{N}}$ thereof in such a way that $|f_{i+1,n_k}(x_{i+1}) - f(x_{i+1})| \to 0$ monotonically. The resulting diagonal sequence, which for convenience we denote by $(f_n)_{n \in \mathbb{N}}$, has the desired property that $|f_n(x) - f(x)| \to 0$ monotonically for every $x \in K'$. Applying [**104**, Corollary 16], we conclude that $f_n \to f$ uniformly on K'.

The generalization to the whole space $\tilde{K}$ is obtained as follows: For any $n \in \mathbb{N}$, we observe that $g_n := |f_n - f|$ is a continuous function on $\tilde{K}$; in particular, its maximum is attained. Since $K' \subset \tilde{K}$ is a dense subset, for any $\varepsilon > 0$ there is $x(n, \varepsilon) \in K'$ such that $\sup_{\tilde{K}} g_n - \varepsilon < g_n(x(n, \varepsilon))$. From this we conclude that, for every $n \in \mathbb{N}$,

$$\sup_{x \in \tilde{K}} |f_n(x) - f(x)| = \sup_{x \in K'} |f_n(x) - f(x)| \;.$$

As a consequence,

$$\lim_{n \to \infty} \sup_{x \in \tilde{K}} |f_n(x) - f(x)| = \lim_{n \to \infty} \sup_{x \in K'} |f_n(x) - f(x)| = 0\,.$$

This proves uniform convergence as desired. Applying [**72**, Theorem 6.8 (c)] implies that $\|f_n - f\|_\infty \to 0$. Employing the convention in [**6**, Section X.4] we deduce that

$$\lim_{n \to \infty} \sup_{x \in K} |f_n(x) - f(x)| = \lim_{n \to \infty} \|f_n - f\|_\infty = 0\,.$$

In particular, we infer that $f_n(x) \to f(x)$ for ρ-almost all $x \in K$. This allows us to apply Lebesgue's dominated convergence theorem [**16**, Theorem 2.8.1], implying that

$$\lim_{n\to\infty} \int_{\mathbb{R}^m} f_n(y)\, d\rho(y) = \int_{\mathbb{R}^m} f(y)\, d\rho(y)\,.$$

This completes the proof. □

LEMMA 5.5.8. *Let $d\rho$ be a positive finite measure on $\mathbb{R}^m$. Then there exists a sequence of measures $(d\rho_n)_{n\in\mathbb{N}}$ on $\mathbb{R}^m$, each of them being absolutely continuous with respect to Lebesgue measure $d\mu$ on $\mathbb{R}^m$, $d\rho_n \ll d\mu$ for all $n \in \mathbb{N}$, such that $(d\rho_n)_{n\in\mathbb{N}}$ converges vaguely (see e.g.* [**9**, §31]*) to $d\rho$ in the sense that*

$$\lim_{n\to\infty} \int_{\mathbb{R}^m} f\, d\rho_n = \int_{\mathbb{R}^m} f\, d\rho \qquad \text{for all } f \in C_c(\mathbb{R}^m)\,.$$

By virtue of the Hahn-Jordan decomposition, this result can be generalized to complex measures on Euclidean space.

PROOF. The idea of the proof is to make use of Dirac sequences (cf. [**5**, §2.13]). To this end, let $\varphi \in C_c^\infty(\mathbb{R}^m)$ be a compactly supported test function with

$$\varphi \geq 0 \qquad \text{and} \qquad \int_{\mathbb{R}^m} \varphi\, d\mu = 1\,.$$

For each $n \in \mathbb{N}$, we let $\varepsilon_n = 1/n$ and

$$\varphi_n(x) := \varepsilon_n^{-m} \varphi\left(\frac{x}{\varepsilon_n}\right)\,.$$

Next, we introduce the functions

$$\eta_n : \mathbb{R}^m \to \mathbb{R}\,, \qquad \eta_n(x) := \int_{\mathbb{R}^m} \varphi_n(x-y)\, d\rho(y)$$

as well as corresponding measures

$$d\rho_n := \eta_n\, d\mu \qquad \text{for all } n \in \mathbb{N}\,.$$

For arbitrary $f \in C_c(\mathbb{R}^m)$, by applying Fubini's theorem we arrive at

$$\begin{aligned}
\lim_{n\to\infty} \int_{\mathbb{R}^m} f(x)\, d\rho_n(x) &= \lim_{n\to\infty} \int_{\mathbb{R}^m} f(x)\, \eta_n(x)\, d\mu(x) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}^m} f(x) \left(\int_{\mathbb{R}^m} \varphi_n(x-y)\, d\rho(y) \right) d\mu(x) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}^m} \underbrace{\left(\int_{\mathbb{R}^m} f(x)\, \varphi_n(x-y)\, d\mu(x) \right)}_{= (\varphi_n * f)(y)} d\rho(y) = \int_{\mathbb{R}^m} f(y)\, d\rho(y)\,,
\end{aligned}$$

where we employed Proposition 5.5.7 for a suitable subsequence of $(\varphi_n)_{n\in\mathbb{N}}$ (which for convenience we again denote by $(\varphi_n)_{n\in\mathbb{N}}$). Since $f \in C_c(\mathbb{R}^m)$ was arbitrary, we obtain vague convergence of measures (for details see [**9**, §31]),

$$d\rho_n = \eta_n\, d\mu \to d\rho \qquad \text{vaguely}\,.$$

Given a complex measure $d\tilde{\rho}$ on $\mathbb{R}^m$, by applying the Hahn-Jordan decomposition to its real and imaginary part, respectively, the above procedure yields a sequence of (complex) measures $(d\tilde{\rho}_n)_{n\in\mathbb{N}}$ on $\mathbb{R}^m$ which converges vaguely to $d\tilde{\rho}$. This completes the proof. □

Similar arguments show that Proposition 5.5.7 and Lemma 5.5.8 also hold in the case that $d\rho$ is a locally finite measure on $\mathbb{R}^m$.

In what follows, let us consider vector-valued L^p-functions (cf. [**117**, Chapter 21]). Given a measurable mapping $f : \mathbb{R}^m \to \mathrm{L}(V)$, for all $1 \leq p < \infty$ we define

$$\|f\|_{L^p(\mathbb{R}^m, \mathrm{L}(V))} := \left(\int_{\mathbb{R}^m} \|f\|^p \, d\mu \right)^{1/p},$$

where $d\mu$ denotes Lebesgue measure on $\mathbb{R}^m$ and $\|.\|$ is an arbitrary norm on $\mathrm{L}(V)$. For general measures $d\rho$ on $\mathbb{R}^m$ we let

$$\|f\|_{L^p(\mathbb{R}^m, \mathrm{L}(V); d\rho)} := \left(\int_{\mathbb{R}^m} \|f\|^p \, d\rho \right)^{1/p}$$

for all $1 \leq p < \infty$. As usual, we introduce (cf. [**117**, §21-5])

$$\|f\|_{L^\infty(\mathbb{R}^m, \mathrm{L}(V))} := \inf \{C \geq 0 : \|f(x)\| \leq C \text{ for } \mu\text{-a.e. } x \in \mathbb{R}^m\},$$
$$\|f\|_{L^\infty(\mathbb{R}^m, \mathrm{L}(V); d\rho)} := \inf \{C \geq 0 : \|f(x)\| \leq C \text{ for } \rho\text{-a.e. } x \in \mathbb{R}^m\}.$$

Accordingly, for all $1 \leq p \leq \infty$ we let

$$L^p(\mathbb{R}^m, \mathrm{L}(V)) := \left\{ f : \mathbb{R}^m \to \mathrm{L}(V) \text{ measurable} : \|f\|_{L^p(\mathbb{R}^m, \mathrm{L}(V))} < \infty \right\},$$
$$L^p(\mathbb{R}^m, \mathrm{L}(V); d\rho) := \left\{ f : \mathbb{R}^m \to \mathrm{L}(V) \text{ measurable} : \|f\|_{L^p(\mathbb{R}^m, \mathrm{L}(V); d\rho)} < \infty \right\}.$$

Since all norms on $\mathrm{L}(V)$ are equivalent and $\mathbb{R} \ni x \mapsto |x|^p$ is a convex function for all $p \geq 1$, for $1 \leq p < \infty$ the elements f of $L^p(\mathbb{R}^m, \mathrm{L}(V))$ and $L^p(\mathbb{R}^m, \mathrm{L}(V); d\rho)$ may also be characterized as measurable functions $f : \mathbb{R}^m \to \mathrm{L}(V)$ with

$$\sum_{i,j=1}^{2n} \int_{\mathbb{R}^m} |f_{i,j}|^p \, d\mu < \infty \qquad \text{and} \qquad \sum_{i,j=1}^{2n} \int_{\mathbb{R}^m} |f_{i,j}|^p \, d\rho < \infty,$$

respectively, where $f_{i,j} = \langle \mathfrak{e}_i \mid f \mathfrak{e}_j \rangle$ for all $i, j = 1, \ldots, 2n$. Choosing $m = 4$, the above definitions clearly apply to $\mathscr{M}$ and $\hat{\mathscr{M}}$.

PROPOSITION 5.5.9. *Let $g \in L^\infty(\mathbb{R}^m, \mathrm{L}(V))$ and assume that $(f_k)_{k \in \mathbb{N}}$ is a sequence in $L^1(\mathbb{R}^m, \mathrm{L}(V))$ which converges to $f \in L^1(\mathbb{R}^m, \mathrm{L}(V))$. Then*

$$\lim_{k \to \infty} \int_{\mathbb{R}^m} g(x) \, f_k(x) \, d\mu(x) = \int_{\mathbb{R}^m} g(x) \, f(x) \, d\mu(x).$$

PROOF. We basically apply [**72**, Theorem 6.8], implying that

$$\lim_{k \to \infty} \left| \int_{\mathbb{R}^m} g(x) \, f_k(x) \, d\mu(x) - \int_{\mathbb{R}^m} g(x) \, f(x) \, d\mu(x) \right| \leq \lim_{k \to \infty} \|g \, (f_k - f)\|_{L^1(\mathbb{R}^n, \mathrm{L}(V))}$$
$$\leq \lim_{k \to \infty} \|g\|_{L^\infty(\mathbb{R}^m, \mathrm{L}(V))} \, \|f_k - f\|_{L^1(\mathbb{R}^m, \mathrm{L}(V))} = 0.$$

This gives the claim. □

Concerning the Fourier transform, it is "useful to remember that the factors of 2π always go with the measure on momentum space" (see [**73**]). Accordingly, following the usual notation in physics [**73**], we define Fourier transform as follows:

DEFINITION 5.5.10. *The Fourier transform of $f \in L^1(\mathscr{M})$, denoted by $\widehat{f} \in C_b(\hat{\mathscr{M}})$, is given by*

$$\widehat{f}(k) := \int_{\mathscr{M}} e^{ik\xi} \, f(\xi) \, d^4\xi \qquad \text{for all } k \in \hat{\mathscr{M}},$$

where $d^4\xi$ is the Haar measure on $\mathscr{M}$. The inverse Fourier transform reads

$$f^\vee(\xi) = \int_{\hat{\mathscr{M}}} e^{-ik\xi}\, f(k)\, \frac{d^4k}{(2\pi)^4} \qquad \text{for all } \xi \in \mathscr{M} \tag{5.5.9}$$

whenever $f \in L^1(\hat{\mathscr{M}})$, where $k\xi = g_{\mu\nu}\, k^\nu\, \xi^\mu = k_\mu\, \xi^\mu$ (see [**73**, Chapter 1]*).*

We point out that most results concerning the Fourier transform carry over to the vector-valued case in a straightforward way.

PROPOSITION 5.5.11. *Assuming that $Q \in L^1(\mathscr{M}, \mathrm{L}(V))$ and $f \in L^\infty(\mathscr{M}, \mathrm{L}(V))$, there is a sequence $(Q_\ell)_{\ell\in\mathbb{N}}$ in $C_c^\infty(\mathscr{M}; \mathrm{L}(V))$ such that*

$$\lim_{\ell\to\infty} \int_{\mathscr{M}} \mathrm{Tr}_V\left(Q_\ell(\xi)\, f(\xi)\right)\, d\mu(\xi) = \int_{\mathscr{M}} \mathrm{Tr}_V\left(Q(\xi)\, f(\xi)\right)\, d\mu(\xi)\,,$$

and, for any $\hat{f} \in L^1(\hat{\mathscr{M}}, \mathrm{L}(V))$,

$$\lim_{\ell\to\infty} \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}_\ell(k)\, \hat{f}(k)\right)\, d\mu(k) = \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}(k)\, \hat{f}(k)\right)\, d\mu(k)\,,$$

where $\hat{Q}, \hat{Q}_\ell$ denotes the Fourier transform of Q, Q_ℓ for every $\ell \in \mathbb{N}$, respectively.

PROOF. Due to [**72**, Proposition 8.17] there is a sequence $(Q_\ell)_{\ell\in\mathbb{N}}$ in $C_c^\infty(\mathscr{M}; \mathrm{L}(V))$ which converges to Q in $L^1(\mathscr{M}, \mathrm{L}(V))$, and $\hat{Q}, \hat{Q}_\ell \in C_b(\hat{\mathscr{M}}, \mathrm{L}(V)) \subset L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))$ for all $\ell \in \mathbb{N}$ due to [**72**, Theorem 8.22]. According to [**72**, Section 8.3], the Fourier transform is a linear mapping from $L^1(\mathscr{M}, \mathrm{L}(V))$ to $L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))$ which is continuous in the sense that

$$\|\hat{Q}_\ell - \hat{Q}\|_{L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))} = \|\widehat{(Q_\ell - Q)}\|_{L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))} \leq \|Q_\ell - Q\|_{L^1(\mathscr{M}, \mathrm{L}(V))} \to 0$$

as $\ell \to \infty$. Applying [**72**, Theorem 6.8] we thus obtain

$$\begin{aligned} \lim_{\ell\to\infty} \left| \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}_\ell \hat{f}\right) d\mu - \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}\, \hat{f}\right) d\mu \right| &\leq \lim_{\ell\to\infty} \|\left(\hat{Q}_\ell - \hat{Q}\right) \hat{f}\|_{L^1(\hat{\mathscr{M}}, \mathrm{L}(V))} \\ &\leq \lim_{\ell\to\infty} \left\|\hat{f}\right\|_{L^1(\hat{\mathscr{M}}, \mathrm{L}(V))} \left\|\hat{Q}_\ell - \hat{Q}\right\|_{L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))} = 0\,. \end{aligned}$$

For $f \in L^\infty(\mathscr{M}, \mathrm{L}(V))$, the first equations follows in the same fashion from

$$\begin{aligned} \lim_{\ell\to\infty} \left| \int_{\mathscr{M}} \mathrm{Tr}_V\left(Q_\ell\, f\right)\, d\mu - \int_{\mathscr{M}} \mathrm{Tr}_V\left(Q\, f\right)\, d\mu \right| &\leq \lim_{\ell\to\infty} \|\left(Q_\ell - Q\right) f\|_{L^1(\mathscr{M}, \mathrm{L}(V))} \\ &\leq \lim_{\ell\to\infty} \left\|f\right\|_{L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))} \left\|Q_\ell - Q\right\|_{L^1(\hat{\mathscr{M}}, \mathrm{L}(V))} = 0\,. \end{aligned}$$

This completes the proof. □

5.5.4. Alternative Representation of the First Variation. The aim of this subsection is to provide an alternative representation of (5.5.8) in momentum space. To begin with, we apply the technical preparations in §5.5.3 in order to clarify integration with respect to operator-valued measures. More precisely, we assume that $d\eta \in \mathfrak{Ovm}$ is an operator-valued measure on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$, and let $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ be a pseudo-orthonormal basis of V. Then

$$d\eta_{i,j} := d \prec \mathfrak{e}_i \mid \eta\, \mathfrak{e}_j \succ \qquad \text{for all } i, j \in \{1, \ldots, 2n\}$$

is a complex measure in $\mathbf{M}_{\mathbb{C}}(\hat{\mathscr{M}})$ (see Definition 5.3.3). Denoting by $d\rho$ the variation measure $d|\eta|$ of $d\eta$ (for details we refer to Definition 5.3.4), we conclude that

$$d\rho := d|\eta|$$

is a non-negative finite measure on $\mathcal{B}(\hat{\mathscr{M}})$. In particular, whenever $\eta(\Omega) = 0$ for some Borel set $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$,

$$\eta_{i,j}(\Omega) = \prec \mathfrak{e}_i \mid \eta(\Omega)\, \mathfrak{e}_j \succ = 0 \qquad \text{for all } i,j \in \{1,\ldots,2n\}\,,$$

implying that each component $d\eta_{i,j}$ is absolutely continuous with respect to $d\rho$. For this reason, the operator-valued measure $d\eta$ is absolutely continuous with respect to $d\rho$ (for details see [**40**, Section VII.2]). Applying the Radon-Nikodým theorem to each component $d\eta_{i,j}$ gives rise to functions $\hat{H}_{i,j} : \hat{\mathscr{M}} \to \mathbb{C}$ with $d\eta_{i,j} = \hat{H}_{i,j}\, d\rho$ for all $i,j \in \{1,\ldots,2n\}$. We thus end up with a mapping

$$\hat{H} : \hat{\mathscr{M}} \to \mathrm{L}(V) \qquad \text{such that} \qquad d\eta = \hat{H}\, d\rho\,.$$

From the fact that $d\eta = \hat{H}\, d\rho$ is bounded we deduce that $\hat{H} \in L^1(\hat{\mathscr{M}}, \mathrm{L}(V); d\rho)$ (see for instance [**72**, Section 3.2]).

Definition 5.5.12. *Let $d\rho \in \mathfrak{Dvm}$ and let $d\rho := d|\eta|$ be its variation measure. Assuming that $\hat{H} \in L^1(\hat{\mathscr{M}}, \mathrm{L}(V); d\rho)$ such that $d\eta = \hat{H}\, d\rho$, for every $f \in L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V); d\rho)$ we define*

$$\int d\eta(k)\, f(k) := \int \hat{H}(k)\, f(k)\, d\rho(k)\,,$$

and for all $u, v \in V$ we let

$$\prec u \mid \left(\int d\eta(k)\, f(k) \right) v \succ := \int \prec u \mid \hat{H}(k)\, f(k)\, v \succ \; d\rho(k)\,,$$

which is well-defined in view of [**72**, Theorem 6.8].

Since $d\rho$ is a finite measure on $\hat{\mathscr{M}} \simeq \mathbb{R}^4$, we may apply Lemma 5.5.8 to obtain a sequence of measures $(d\rho_n)_{n\in\mathbb{N}}$ on $\hat{\mathscr{M}}$, each of them being absolutely continuous with respect to Lebesgue measure $d\mu$ on $\hat{\mathscr{M}}$ such that $d\rho_n \to d\rho$ vaguely. Furthermore, in view of [**139**, Theorem 3.14], we know that $C_c(\hat{\mathscr{M}}; d\rho)$ is dense in $L^1(\hat{\mathscr{M}}; d\rho)$, implying that

$$\hat{u}_m \to \hat{H} \qquad \text{in } L^1(\hat{\mathscr{M}}, \mathrm{L}(V); d\rho)$$

for some sequence $(\hat{u}_m)_{m\in\mathbb{N}}$ in $C_c(\hat{\mathscr{M}}; \mathrm{L}(V))$. For any function $g \in L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V); d\eta)$ we may define

$$\int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \left(g(k)\, d\eta(k) \right) := \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \left(g(k)\, \hat{H}(k) \right) d\rho(k)\,.$$

Making use of Proposition 5.5.9, we then obtain

$$\int_{\hat{\mathscr{M}}} g(k)\, \hat{H}(k)\, d\rho(k) = \lim_{m\to\infty} \int_{\hat{\mathscr{M}}} g(k)\, \hat{u}_m(k)\, d\rho(k) = \lim_{m\to\infty} \lim_{n\to\infty} \int_{\hat{\mathscr{M}}} g(k)\, \hat{u}_m(k)\, d\rho_n(k)\,,$$

and by continuity of the trace we arrive at

$$\int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \left(g(k)\, d\eta(k) \right) = \lim_{m,n\to\infty} \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \left(g(k)\, \hat{u}_m(k) \right) d\rho_n(k)\,.$$

We are now in the position to express the first variation of the causal action (5.5.8) in momentum space.

LEMMA 5.5.13. *Assume that the Lagrangian satisfies Assumption 5.5.4. Moreover, let* $d\omega \in \mathfrak{Dvm}$ *and assume that* $Q[\omega] \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^\infty(\mathscr{M}, \mathrm{L}(V))$ *with* $Q = Q[\omega] : \mathscr{M} \to \mathrm{L}(V)$ *given by* (5.5.7)*. Then* $\hat{Q} \in C_0(\hat{\mathscr{M}}; \mathrm{L}(V)) \cap L^2(\hat{\mathscr{M}}, \mathrm{L}(V))$*, where* $\hat{Q} = \hat{Q}[\omega]$ *denotes the Fourier transform of* Q*, and the first variation of the causal action* (5.5.8) *at* $d\omega$ *in direction* $d\eta \in \mathfrak{Dvm}$ *takes the form*

$$\delta\mathcal{S}(\omega; \eta) = 4 \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \left(\hat{Q}[\omega](q)\, d\eta(q) \right) . \tag{5.5.10}$$

PROOF. Assuming that $Q[\omega] \in L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^1(\mathscr{M}, \mathrm{L}(V))$, by [**72**, Theorem 8.29] we obtain $\hat{Q}[\omega] \in L^2(\hat{\mathscr{M}}, \mathrm{L}(V))$, and $\hat{Q}[\omega] \in C_0(\hat{\mathscr{M}}; \mathrm{L}(V))$ is an immediate consequence of [**72**, Theorem 8.22].

Next, given $d\eta \in \mathfrak{Dvm}$, by $d\rho := d|\eta|$ we denote the variation measure according to Definition 5.3.4. Following §5.5.3, there is a function $\hat{H} \in L^1(\hat{\mathscr{M}}, \mathrm{L}(V); d\rho)$ such that

$$d\eta = \hat{H}\, d\rho ,$$

and in virtue of Lemma 5.5.8 there exists a sequence of measures $(d\rho_j)_{j \in \mathbb{N}}$ on $\hat{\mathscr{M}}$ with the property that $d\rho_j \ll d\mu$ for all $j \in \mathbb{N}$ and $d\rho_j \to d\rho$ vaguely. By $\hat{R}_j \in L^1(\hat{\mathscr{M}}, \mathrm{L}(V))$ we denote the corresponding density, i.e. $d\rho_j = \hat{R}_j\, d\mu$ for each $j \in \mathbb{N}$. On the other hand, according to [**139**, Theorem 8.14], there is a sequence $(\hat{u}_m)_{m \in \mathbb{N}}$ in $C_c(\hat{\mathscr{M}}; \mathrm{L}(V))$ which converges to $\hat{H}$ in $L^1(\hat{\mathscr{M}}, \mathrm{L}(V); d\rho)$. In order to obtain a sequence of Schwartz functions, we employ [**72**, Proposition 8.17] to conclude that, for each $m \in \mathbb{N}$, there exists a sequence $(\hat{u}_{m,n})_{n \in \mathbb{N}}$ in $C_c^\infty(\hat{\mathscr{M}}; \mathrm{L}(V))$ such that $\hat{u}_{m,n} \to \hat{u}_m$ in $L^1(\hat{\mathscr{M}}, \mathrm{L}(V))$ as $n \to \infty$. By the same reasoning, for each $j \in \mathbb{N}$ there exists a sequence $(\hat{\varphi}_{j,i})_{i \in \mathbb{N}}$ in $C_c^\infty(\hat{\mathscr{M}}; \mathrm{L}(V))$ which converges to $\hat{R}_j$ in $L^1(\hat{\mathscr{M}}, \mathrm{L}(V))$. In summary, we thus obtain

$$\begin{aligned} P[\eta](\xi) &= \int_{\hat{\mathscr{M}}} e^{ik\xi}\, d\eta(k) = \int_{\hat{\mathscr{M}}} e^{ik\xi}\, \hat{H}\, d\rho(k) = \lim_{m \to \infty} \int_{\hat{\mathscr{M}}} e^{ik\xi}\, \hat{u}_m(k)\, d\rho(k) \\ &= \lim_{m \to \infty} \lim_{j \to \infty} \int_{\hat{\mathscr{M}}} e^{ik\xi}\, \hat{u}_m(k)\, d\rho_j(k) = \lim_{m \to \infty} \lim_{j \to \infty} \int_{\hat{\mathscr{M}}} e^{ik\xi}\, \hat{u}_m(k)\, R_j(k)\, d^4k \end{aligned}$$

for arbitrary $\xi \in \mathscr{M}$. Since $\hat{u}_m \in L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V))$, by [**72**, Theorem 6.8] we conclude that

$$P[\eta](\xi) = \lim_{m \to \infty} \lim_{j \to \infty} \lim_{i \to \infty} \int_{\hat{\mathscr{M}}} e^{ik\xi}\, \hat{u}_m(k)\, \varphi_{j,i}(k)\, d^4k$$

for any $\xi \in \mathscr{M}$, and applying the same arguments once again, we deduce that

$$P[\eta](\xi) = (2\pi)^4 \lim_{m \to \infty} \lim_{j \to \infty} \lim_{i \to \infty} \lim_{n \to \infty} \int_{\hat{\mathscr{M}}} e^{-ik(-\xi)} \left(\hat{u}_{m,n}\, \hat{\varphi}_{j,i} \right)(k)\, \frac{d^4k}{(2\pi)^4} .$$

In view of (5.5.9), the integral expression is exactly the inverse Fourier transform,

$$\left(\hat{u}_{m,n}\, \hat{\varphi}_{j,i} \right)^\vee(-\xi) = \int_{\hat{\mathscr{M}}} e^{-ik(-\xi)} \left(\hat{u}_{m,n}\, \hat{\varphi}_{j,i} \right)(k)\, \frac{d^4k}{(2\pi)^4} \qquad \text{for all } \xi \in \mathscr{M} .$$

Note that, up to an irrelevant constant factor, the inverse Fourier transform (5.5.9) is precisely of the familiar form of Fourier transforms in mathematical textbooks (see e.g. [**133**]). Consequently, the convolution theorem [**133**, Theorem IX.3] implies that

$$\left(\hat{u}_{m,n}\, \hat{\varphi}_{j,i} \right)^\vee(\xi) = \left((\hat{u}_{m,n})^\vee * (\hat{\varphi}_{j,i})^\vee \right)(\xi) \qquad \text{for all } \xi \in \mathscr{M} ,$$

and according to [**166**, Section VI.1], the convolution of two Schwartz functions is again a Schwartz function. Since $\hat{u}_{m,n}, \hat{\varphi}_{j,i} \in \mathcal{S}(\hat{\mathscr{M}}; \mathrm{L}(V))$, this shows

$$\big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee \in \mathcal{S}(\mathscr{M}; \mathrm{L}(V)) \qquad \text{for all } i,j,m,n \in \mathbb{N}\,.$$

Assuming that $Q \in L^\infty(\mathscr{M}, \mathrm{L}(V))$, by Proposition 5.5.9 we obtain

$$\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q(\xi)\, P[\eta](-\xi)\big)\, d\mu(\xi) = (2\pi)^4 \lim_{m,j,i,n\to\infty} \mathrm{Tr}_V \int_{\mathscr{M}} Q(\xi)\, \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee(-\xi)\, d\mu(\xi)\,.$$

From $\mathcal{S}(\mathscr{M}; \mathrm{L}(V)) \subset C_0(\mathscr{M}; \mathrm{L}(V))$ we conclude that $\big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee \in L^\infty(\mathscr{M}, \mathrm{L}(V))$. As a consequence, by Proposition 5.5.11 there exists a sequence $(Q_\ell)_{\ell\in\mathbb{N}}$ in $C_c^\infty(\mathscr{M}; \mathrm{L}(V))$ such that

$$\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q(\xi)\, P[\eta](-\xi)\big)\, d\mu(\xi) = (2\pi)^4 \lim_{m,j,i,n,\ell\to\infty} \mathrm{Tr}_V \int_{\mathscr{M}} Q_\ell(\xi)\, \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee(-\xi)\, d\mu(\xi)$$

in view of Hille's theorem [**29**, Theorem II.2.6] and continuity of the trace. Since Q_ℓ as well as its Fourier transform, denoted by $\hat{Q}_\ell$, are Schwartz functions for each $\ell \in \mathbb{N}$ (see e.g. [**72**, Corollary 8.23], the Fourier inversion theorem [**133**, Theorem IX.1] yields that

$$\hat{Q}_\ell(k) = (\hat{Q}_\ell^\vee)^\vee(-k) \qquad \text{and} \qquad \Big(\big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee\Big)^\vee(k) = \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)(-k)$$

for all $k \in \hat{\mathscr{M}}$ and $\ell, m, j, i, n \in \mathbb{N}$ (also cf. [**71**, Fourier Inversion Theorem II (4.32)]). Applying Parseval's theorem [**166**, Section VI.2], we thus obtain

$$\begin{aligned}\int_{\mathscr{M}} Q_\ell(\xi)\, \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee(-\xi)\, d\mu(\xi) &= \int_{\hat{\mathscr{M}}} Q_\ell^\vee(-k)\, \Big(\big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)^\vee\Big)^\vee(-k)\, \frac{d^4k}{(2\pi)^4} \\ &= \int_{\hat{\mathscr{M}}} \hat{Q}_\ell(k)\, \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)(k)\, \frac{d^4k}{(2\pi)^4}\end{aligned}$$

for all $\ell, m, j, i, n \in \mathbb{N}$. Again applying Proposition 5.5.11 gives

$$\lim_{\ell\to\infty} \int_{\hat{\mathscr{M}}} \hat{Q}_\ell(k)\, \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)(k)\, \frac{d^4k}{(2\pi)^4} = \int_{\hat{\mathscr{M}}} \hat{Q}(k)\, \big(\hat{u}_{m,n}\, \hat{\varphi}_{j,i}\big)(k)\, \frac{d^4k}{(2\pi)^4}\,,$$

and by virtue of [**72**, Theorem 6.8] we deduce that

$$\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q(\xi)\, P[\eta](-\xi)\big)\, d\mu(\xi) = \lim_{m,j\to\infty} \mathrm{Tr}_V \int_{\hat{\mathscr{M}}} \hat{Q}(k)\, \hat{u}_m(k)\, d\rho_j(k)\,.$$

Making use of $\hat{u}_m \in C_c(\hat{\mathscr{M}}; \mathrm{L}(V))$, by vague convergence of measures we obtain

$$\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q(\xi)\, P[\eta](-\xi)\big)\, d\mu(\xi) = \lim_{m\to\infty} \mathrm{Tr}_V \int_{\hat{\mathscr{M}}} \hat{Q}(k)\, \hat{u}_m(k)\, d\rho(k)\,.$$

Applying Proposition 5.5.9 once again, we finally arrive at

$$\int_{\mathscr{M}} \mathrm{Tr}_V\big(Q(\xi)\, P[\eta](-\xi)\big)\, d\mu(\xi) = \mathrm{Tr}_V \int_{\hat{\mathscr{M}}} \hat{Q}(k)\, \hat{H}(k)\, d\rho(k) = \mathrm{Tr}_V \int_{\hat{\mathscr{M}}} \hat{Q}(k)\, d\eta(k)\,.$$

Combining the last equality with (5.5.8) yields (5.5.10). This completes the proof. □

REMARK 5.5.14. *Given a negative definite measure $d\nu$ and an operator-valued $d\eta$ in such a way that, for some $\delta > 0$, the variations $d\tilde{\nu}_\tau := d\nu + \tau\, d\eta$ are negative definite measures for all $\tau \in [0, \delta)$, we are mainly interested in the one-sided first variation*

$$\delta^+\mathcal{S}(\nu; \eta) := \lim_{\tau\searrow 0} \frac{\mathcal{S}(\nu + \tau\,\eta) - \mathcal{S}(\nu)}{\tau}\Big|_{\tau=0}\,.$$

Whenever $\mathcal{S}$ is Fréchet differentiable, we conclude that $\delta^+\mathcal{S}$ coincides with the Fréchet derivative of $\mathcal{S}$ (cf. [**168**, Section 40.1]*). In this case, the result of Lemma 5.5.13 holds true whenever $\delta\mathcal{S}$ is replaced by $\delta^+\mathcal{S}$ (and similarly for Corollary 5.5.15 below).*

Lemma 5.5.13 allows us to derive the subsequent corollary.

COROLLARY 5.5.15. *Assume that the Lagrangian $\mathcal{L}$ satisfies Assumption 5.5.4, and for $d\omega \in \mathfrak{Ovm}$ suppose that $Q[\omega] \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^\infty(\mathscr{M}, \mathrm{L}(V))$, where the mapping $Q[\omega] : \mathscr{M} \to \mathrm{L}(V)$ is given by* (5.5.7)*. Then for every $p \in \hat{\mathscr{M}}$, each Borel set $U \in \mathcal{B}(\hat{\mathscr{M}})$ and any operator $A' \in \mathrm{L}(V)$ satisfying the constraints*

$$\mathrm{Tr}_V(A') = \mathrm{Tr}_V(\nu(U)) \qquad \text{and} \qquad \mathrm{Tr}_V(SA') = \mathrm{Tr}_V(S\nu(U))\,, \tag{5.5.11}$$

the first variation of the causal action at $d\omega$ in direction $d\eta = A'\, d\delta_p - \chi_U\, d\omega$ reads

$$\delta\mathcal{S}(\omega;\eta) = 4\,\mathrm{Tr}_V\left(\hat{Q}[\omega](p)\,A'\right) - 4\int_U \mathrm{Tr}_V\left(\hat{Q}[\omega](q)\,d\omega(q)\right), \tag{5.5.12}$$

where $\hat{Q} = \hat{Q}[\omega]$ is the Fourier transform of $Q = Q[\omega]$.

PROOF. Starting point for the proof of (5.5.12) is expression (5.5.10),

$$\delta\mathcal{S}(\omega;\eta) = 4\int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}[\omega](q)\,d\eta(q)\right),$$

valid for operator-valued measures $d\omega, d\eta \in \mathfrak{Ovm}$. Next, given $p \in \hat{\mathscr{M}}$ and $U \in \mathcal{B}(\hat{\mathscr{M}})$, for any operator $A' \in \mathrm{L}(V)$ satisfying the constraints

$$\mathrm{Tr}_V(A') = \mathrm{Tr}_V(\nu(U)) \qquad \text{and} \qquad \mathrm{Tr}_V(SA') = \mathrm{Tr}_V(S\nu(U))\,,$$

we define the operator-valued measure $d\eta \in \mathfrak{Ovm}$ by

$$d\eta := A'\, d\delta_p - d\omega\,,$$

where $d\delta_p$ denotes the Dirac measure supported at p. For $\delta > 0$, we then consider the operator-valued measures $(d\tilde{\nu}_\tau)_{\tau\in(-\delta,\delta)}$ in $\mathfrak{Ovm}$, given by

$$d\tilde{\omega}_\tau := d\omega + \tau\, d\eta = (1-\tau)\, d\omega + \tau\, A'\, d\delta_p \qquad \text{for all } \tau \in (-\delta, \delta)\,.$$

Furthermore, the first variation of the kernel of the fermionic projector is given by

$$\delta P[\omega;\eta] = \frac{d}{d\tau} P[\omega + \tau\,\eta]|_{\tau=0} = P[\eta]\,.$$

Combining the last equality in the proof of Lemma 5.5.13 with (5.5.8), we conclude that

$$\delta\mathcal{S}(\omega;\eta) = 4\int_{\mathscr{M}} \mathrm{Tr}_V\left(Q[\omega](\xi)\,P[\eta](-\xi)\right) d\mu(\xi) = 4\int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}[\omega](q)\,d\eta(q)\right).$$

Inserting the the specific form of $d\eta$, the first variation of the causal action reads

$$\delta\mathcal{S}(\omega;\eta) = 4\,\mathrm{Tr}_V\left(\hat{Q}[\omega](p)\,A'\right) - 4\int_{\hat{\mathscr{M}}} \mathrm{Tr}_V\left(\hat{Q}[\omega](q)\,d\omega(q)\right),$$

thus proving (5.5.12). □

Having accomplished to express the first variation of the causal action (5.5.10) in momentum space, we may now focus on deriving the corresponding EL equations. To this end, we assume that $d\nu$ is a minimizer of the causal action $\mathcal{S}$ in the class $\mathfrak{Ndm}$ with respect to the side conditions

$$\mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = c \qquad \text{and} \qquad \mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f$$

for positive constants $c, f > 0$. Restricting attention to the subset

$$N := \left\{ d\nu \in \mathfrak{Ndm} : \mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = c \text{ and } \mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f \right\} ,$$

the negative definite measure $d\nu$ is a minimizer in the sense of Definition 5.4.9 (in the terminology of [**168**], the negative definite measure $d\nu$ is in particular a *bound local minimum* in the sense of [**168**, Definition 43.4]).

In what follows, we consider variations of $d\nu$ in the spirit of Corollary 5.5.15. More precisely, let $p \in \hat{\mathscr{M}}$ and $U \in \mathcal{B}(\hat{\mathscr{M}})$. For any negative definite operator $A' \in \mathrm{L}(V)$ satisfying the constraints

$$\mathrm{Tr}_V(A') = \mathrm{Tr}_V(\nu(U)) \qquad \text{and} \qquad \mathrm{Tr}_V(SA') = \mathrm{Tr}_V(S\nu(U)) ,$$

we introduce the operator-valued measure $d\eta$ by

$$d\eta := A' \, d\delta_p - d\nu ,$$

where $d\delta_p$ denotes the Dirac measure supported at p. Given $\delta \in (0,1)$, we then define the variations $(d\tilde{\nu}_\tau)_{\tau \in [0,\delta)}$ by

$$d\tilde{\nu}_\tau := d\nu + \tau \, d\eta = (1-\tau) \, d\nu + \tau \, A' \, d\delta_p \qquad \text{for all } \tau \in [0,\delta) .$$

Since A' is negative definite, we conclude that $d\tilde{\nu}_\tau \in \mathfrak{Ndm}$ for all $\tau \in [0,\delta)$. Moreover, we also have $d\tilde{\nu}_\tau \in N$ for all $\tau \in [0,\delta)$ with $N \subset \mathfrak{Ndm}$ according to Definition 5.4.8. In view of Definition 5.4.9, the minimizing property of $d\nu$ implies that

$$\mathcal{S}(\tilde{\nu}_\tau) - \mathcal{S}(\nu) \geq 0 \qquad \text{for all } \tau \in [0,\delta) .$$

Restricting attention to the right-handed limit in order to vary in $\mathfrak{Ndm}$, we obtain

$$\delta^+ \mathcal{S}(\nu; \eta) = \lim_{\tau \searrow 0} \frac{\mathcal{S}(\nu + \tau\eta) - \mathcal{S}(\nu)}{\tau} |_{\tau=0} \geq 0 \tag{5.5.13}$$

(for an illustration in a different context see [**12**, Figure 2]).

PROPOSITION 5.5.16. *Assume that the Lagrangian $\mathcal{L}$ is continuously differentiable, and let the negative definite measure $d\nu \in \mathfrak{Ndm}$ be a minimizer of the causal variational principle*

$$\textit{minimize } \mathcal{S}(\nu) \textit{ in the class } \mathfrak{Ndm}$$

with respect to the constraints

$$\mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = c \qquad \textit{and} \qquad \mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f .$$

Moreover, assume that $Q[\nu] \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^\infty(\mathscr{M}, \mathrm{L}(V))$, where the mapping $Q[\nu] : \mathscr{M} \to \mathrm{L}(V)$ is given by (5.5.7)*. Then for every Borel set $U \in \mathcal{B}(\hat{\mathscr{M}})$ and any negative definite matrix $A' \in \mathrm{L}(V)$ satisfying the constraints*

$$\mathrm{Tr}_V(A') = \mathrm{Tr}_V(\nu(U)) \qquad \textit{and} \qquad \mathrm{Tr}_V(SA') = \mathrm{Tr}_V(S\nu(U)) ,$$

the inequality

$$\mathrm{Tr}_V\left(\hat{Q}(p)\, A'\right) \geq \int_U \mathrm{Tr}_V\left(\hat{Q}(q)\, d\nu(q)\right) \tag{5.5.14}$$

holds for each $p \in \hat{\mathscr{M}}$, every Borel set $U \in \mathcal{B}(\hat{\mathscr{M}})$ and any negative definite matrix A' in $\mathrm{L}(V)$ satisfying (5.5.11)*. In particular, for all $U \in \mathcal{B}(\hat{\mathscr{M}})$ and $p \in \hat{\mathscr{M}}$ we have*

$$\mathrm{Tr}_V\left(\hat{Q}(p)\, \nu(U)\right) \geq \int_U \mathrm{Tr}_V\left(\hat{Q}(q)\, d\nu(q)\right) . \tag{5.5.15}$$

PROOF. In analogy to the proof of Corollary 5.5.15, we introduce $d\eta$ in $\mathfrak{Ovm}$ by

$$d\eta := A' \, d\delta_p - d\nu \,.$$

Given $\delta \in (0,1)$, we then define variations $(d\tilde{\nu}_\tau)_{\tau \in [0,\delta)}$ in $\mathfrak{Ndm}$ by setting

$$d\tilde{\nu}_\tau := d\nu + \tau \, d\eta = (1-\tau)\, d\nu + \tau \, A' \, d\delta_p \qquad \text{for all } \tau \in [0,\delta) \,.$$

Making use of the fact that $d\nu$ is a minimizer (see Definition 5.4.9), from

$$0 \leq \lim_{\tau \searrow 0} \frac{1}{\tau} \big(\mathcal{S}(\tilde{\nu}_\tau) - \mathcal{S}(\nu) \big)|_{\tau=0} = \delta^+ \mathcal{S}(\nu; \eta)$$

(see inequality (5.5.13)) together with (5.5.12) we infer the inequality (5.5.14),

$$\mathrm{Tr}_V \left(\hat{Q}[\nu](p)\, A' \right) \geq \int_U \mathrm{Tr}_V \left(\hat{Q}[\nu](q)\, d\nu(q) \right) .$$

Choosing $A' = \nu(U)$ immediately yields (5.5.15). □

Proposition 5.5.16 yields a lower bound for the expression on the left-hand side of (5.5.14). In order for deriving EL equations in momentum space, this motivates to consider auxiliary variational principles (Section 5.6); afterwards we deduce the desired EL equations under appropriate assumptions (Section 5.7).

5.6. Auxiliary Variational Principles

Throughout this section, we let $d\nu$ be negative definite measure and consider the mapping $Q = Q[\nu] : \mathscr{M} \to \mathrm{L}(V)$ given by (5.5.7). Its Fourier transform shall be denoted by $\hat{Q} = \hat{Q}[\nu]$. The goal of this section is to clarify the structure of $\hat{Q}$. To this end, we first derive some useful properties of the Fourier transform $\hat{Q}$ as introduced in Lemma 5.5.13 (§5.6.1). Next, we consider appropriate auxiliary variational principles which will be treated by Lagrange multipliers (§5.6.2). This allows us to represent the operator $\hat{Q}$ in a special form (§5.6.3). The main result of this section (Theorem 5.6.8) is the starting point for the derivation of Euler-Lagrange equations in momentum space in Section 5.7 below.

5.6.1. Preliminaries. This subsection is devoted to the proof that, for any $k \in \hat{\mathscr{M}}$, the operator $\hat{Q}(k)$ is self-adjoint (with respect to $\prec . \,|\, . \succ$).

PROPOSITION 5.6.1. *For any $\xi \in \mathscr{M}$, the operator $Q(\xi)$ is symmetric (with respect to the spin scalar product $\prec . \,|\, . \succ$) in the sense that $Q(\xi)^* = Q(-\xi)$.*

PROOF. For any $\xi \in \mathscr{M}$, in view of (5.3.6), (5.5.6) and Proposition 5.5.6 we are given

$$\begin{aligned} Q(\xi)^* &= \frac{1}{2} \Big(R(\xi) + \frac{1}{2} P(\xi)\, \mathcal{M}(-\xi) \Big)^* \\ &= \frac{1}{2} \Big(R(\xi)^* + \frac{1}{2} \mathcal{M}(-\xi)\, P(-\xi) \Big) = \frac{1}{2} \left(R(-\xi) + R(\xi)^* \right) = Q(-\xi) \,. \end{aligned}$$

This proves the claim. □

Under the assumptions of Lemma 5.5.13 we then obtain symmetry of $\hat{Q}$.

COROLLARY 5.6.2. *Assuming that $Q \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V))$, for any $k \in \hat{\mathscr{M}}$ the operator $\hat{Q}(k)$ is symmetric (with respect to $\prec . \,|\, . \succ$), i.e. $\hat{Q}(k)^* = \hat{Q}(k)$.*

PROOF. Assuming that $Q \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V))$, by employing linearity, Hille's theorem (see [**29**, Theorem II.2.6]) as well as Parseval's formula [**166**, Chapter VI, Section 2] we obtain

$$\begin{aligned} 0 &= \int_{\mathscr{M}} (\prec u \mid Q(\xi)^* v \succ - \prec u \mid Q(\xi)\, v \succ)\, g(\xi)\, d\mu(\xi) \\ &= \int_{\hat{\mathscr{M}}} \Big(\prec u \mid \hat{Q}(k)^* v \succ - \prec u \mid \hat{Q}(k)\, v \succ\Big)\, \hat{g}(-k)\, d^4k \end{aligned}$$

for all Schwartz functions $g \in \mathcal{S}(\mathscr{M})$ and thus for all $\hat{g} \in \mathcal{S}(\hat{\mathscr{M}})$, where in the first step we applied Proposition 5.6.1. In view of $C_c^\infty(\hat{\mathscr{M}}; \mathrm{L}(V)) \subset \mathcal{S}(\hat{\mathscr{M}}; \mathrm{L}(V))$, the fundamental lemma of calculus of variations (see for instance [**24**, Theorem 6.3-2]) implies that

$$\prec u \mid \hat{Q}(k)^* v \succ = \prec u \mid \hat{Q}(k)\, v \succ \qquad \text{for all } k \in \hat{\mathscr{M}}\,,$$

where we made use of continuity of $\hat{Q} \in C_0(\hat{\mathscr{M}}; \mathrm{L}(V))$ in virtue of Lemma 5.5.13. Hence non-degeneracy of the spin scalar product yields $\hat{Q}(k)^* = \hat{Q}(k)$ for all $k \in \hat{\mathscr{M}}$. □

5.6.2. Auxiliary Variational Principles. In order to obtain a useful representation of the operator $\hat{Q}$, in this subsection we focus on auxiliary variational principles based on inequality (5.5.14). More precisely, for given constants $c, f > 0$ we assume that $d\nu \in \mathfrak{Mm}$ is a minimizer (in the sense of Definition 5.4.9) of the causal variational principle (5.4.8) with respect to the side conditions

$$\mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = c \qquad \text{and} \qquad \mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f\,. \tag{5.6.1}$$

Considering arbitrary $p \in \hat{\mathscr{M}}$ and $U \in \mathcal{B}(\hat{\mathscr{M}})$, by Proposition 5.5.16 the inequality

$$\mathrm{Tr}_V\left(\hat{Q}[\nu](p)\, A'\right) \geq \int_U \mathrm{Tr}_V\left(\hat{Q}[\nu](q)\, d\nu(q)\right)$$

holds true for any negative definite matrix $A' \in \mathrm{L}(V)$ satisfying (5.5.11), i.e.

$$\mathrm{Tr}_V(A') = \mathrm{Tr}\,(\nu(U)) =: c_U \qquad \text{and} \qquad \mathrm{Tr}_V(-SA') = \mathrm{Tr}\,(-S\nu(U)) =: f_U\,. \tag{5.6.2}$$

The main ideas of this section can be summarized as follows: On the one hand, for every $U \in \mathcal{B}(\hat{\mathscr{M}})$ we regard (5.6.2) as side conditions for a variational principle under variations of A'. On the other hand, we restrict attention to the Euclidean inner product $\langle\, .\,|\,.\,\rangle$ on V instead of the indefinite inner product $\prec .\,|\,.\succ$. More precisely, for arbitrary $k \in \hat{\mathscr{M}}$, the matrix

$$X(k) := -\hat{Q}(k) S$$

is symmetric with respect to the scalar product[4]

$$\langle\, .\,|\,.\,\rangle = \prec .\,|\,S.\succ$$

on V as given by (5.2.1); indeed, for all $u, v \in V$,

$$\langle u \mid X(k)\, v\rangle = -\prec u \mid S\hat{Q}(k) S v \succ = -\prec \hat{Q}(k) S\, u \mid Sv \succ = \langle X(k)\, u \mid v\rangle\,,$$

where we used $\hat{Q}(k)^* = \hat{Q}(k)$ according to Corollary 5.6.2 (with respect to $\prec .\,|\,.\succ$). Next, the operator $A \in \mathrm{L}(V)$ defined by $A := -SA'$ satisfies the conditions

$$\begin{aligned} \mathrm{Tr}_V(A) &= \mathrm{Tr}_V(-SA') = \mathrm{Tr}_V(-S\nu(U)) = f_U\,, \\ \mathrm{Tr}_V(SA) &= \mathrm{Tr}_V(-A') = \mathrm{Tr}_V(-\nu(U)) = -c_U\,. \end{aligned} \tag{5.6.3}$$

[4]For similar considerations see [**59**, Exercise 1.15].

Since V is finite-dimensional, the operator $A \in \mathrm{L}(V)$ is compact. Furthermore, it is self-adjoint and positive with respect to $\langle\,.\,|\,.\,\rangle$ in view of

$$\langle u \,|\, Av\rangle = \prec u \,|\, -SSA'v \succ = \prec -A'u \,|\, v \succ = \prec -SA'u \,|\, Sv \succ = \langle Au \,|\, v\rangle$$

(where we made use of $S^2 = \mathbb{1}$ and the fact that A' is negative definite) and

$$\langle u \,|\, Au\rangle = \prec -A'u \,|\, u \succ \geq 0$$

for arbitrary $u, v \in V$. Applying [**157**, Satz VI.3.4], we conclude that there is a unique positive self-adjoint (with respect to $\langle\,.\,|\,.\,\rangle$) compact operator $B \in \mathrm{L}(V)$ such that

$$A = B^2 = BB = BB^* = B^*B\,.$$

In order satisfy the constraints (5.6.3), the operator B has to fulfill the conditions

$$\mathrm{Tr}_V\,(BB^*) = f_U\,, \qquad \mathrm{Tr}_V\,(SBB^*) = -c_U\,. \tag{5.6.4}$$

For given $k \in \hat{\mathscr{M}}$ and $U \in \mathfrak{B}(\hat{\mathscr{M}})$, we introduce the functional $F = F(k, U)$ by

$$F : \mathrm{L}(V) \to \mathbb{R}\,, \qquad B \mapsto \mathrm{Tr}_V(X(k)\,BB^*)\,.$$

We now focus on the (auxiliary) variational principle

$$\text{minimize } F \text{ under the constraints (5.6.4)}\,. \tag{5.6.5}$$

Denoting by $\|\cdot\|_{\mathrm{HS}}$ the Hilbert-Schmidt norm with respect to the Hilbert-Schmidt inner product $\langle .,. \rangle_{\mathrm{HS}}$ on Hilbert-Schmidt operators,[5] given by

$$\langle A, B\rangle_{\mathrm{HS}} = \mathrm{Tr}_V(A^*B) = \mathrm{Tr}_V(BA^*)$$

for all $A, B \in \mathrm{L}(V)$, the side conditions (5.6.4) imply that

$$\|B\|_{\mathrm{HS}} = \sqrt{\langle B, B\rangle_{\mathrm{HS}}} = \sqrt{\mathrm{Tr}_V(BB^*)} = \sqrt{f_U} \tag{5.6.6}$$

for admissible variations B (that is, operators $B \in \mathrm{L}(V)$ satisfying condition (5.6.4)). From this we read off that admissible variations of the variational principle (5.6.5) are contained within a compact subset of $\mathrm{L}(V)$. For this reason, there exist minimizers of the above variational principle (5.6.5). (Considering a minimizing sequence, compactness readily yields the existence of a convergent subsequence in such a way that its limit is the desired minimizer.)

In what follows, we apply Lagrange multipliers to the variational principle (5.6.5) with respect to the side conditions (5.6.4) in order to obtain more information about the matrix $X(k) = -\hat{Q}(k)S$. To this end, we first point out that $\mathrm{L}(V)$ is a real locally convex (Banach) space. Introducing the functionals

$$\begin{aligned} F &: \mathrm{L}(V) \to \mathbb{R}\,, & F(B) &:= \mathrm{Tr}_V\,(XBB^*) \\ G_1 &: \mathrm{L}(V) \to \mathbb{R}\,, & G_1(B) &:= \mathrm{Tr}_V\,(BB^*) - f_U \\ G_2 &: \mathrm{L}(V) \to \mathbb{R}\,, & G_2(B) &:= \mathrm{Tr}_V\,(SBB^*) + c_U \end{aligned}$$

(where $X = X(k)$ for given $k \in \hat{\mathscr{M}}$), the variational principle (5.6.5), (5.6.4) reads

$$\begin{aligned} F(B) &= \min! \\ G_i(B) &= 0 \qquad \text{for } i = 1, 2\,. \end{aligned}$$

For convenience, let $G = (G_1, G_2) : \mathrm{L}(V) \to Y = \mathbb{R}^2$, and in order to treat the side conditions we introduce $N = \{B \in \mathrm{L}(V) : G(B) = 0\}$ (see [**168**, Definition 43.4]).

[5] Since V is finite-dimensional, the Hilbert-Schmidt norm coincides with the Frobenius norm.

In order to make use of Lagrange multipliers, we require the functional F to be Fréchet differentiable and G to be a submersion (for details see [**168**, Section 40.1 and Definition 43.15]).[6] For $B, H \in \mathrm{L}(V)$ the Gâteaux derivative $F'(B) : \mathrm{L}(V) \to \mathbb{R}$ is given by the continuous linear functional

$$\langle F'(B), H\rangle = \lim_{\tau\to 0} \frac{F(B+\tau H) - F(B)}{\tau} = \mathrm{Tr}_V\left(XBH^* + XHB^*\right) .$$

Since $\mathrm{L}(V)$ is a normed space and $F'(B)$ is continuous at B, the Gâteaux derivative coincides with the Fréchet derivative of F (see [**168**, Section 40.1]). Next, we focus on the functional G. Similar calculations yield

$$\begin{aligned}\langle G_1'(B), H\rangle &= \mathrm{Tr}_V(BH^* + HB^*) ,\\ \langle G_2'(B), H\rangle &= \mathrm{Tr}_V(SBH^* + SHB^*)\end{aligned}$$

for $B, H \in \mathrm{L}(V)$, proving that G is continuously Fréchet differentiable. As the null space $\mathcal{N}(G'(B))$ is finite-dimensional for every $B \in \mathrm{L}(V)$, it splits $\mathrm{L}(V)$ in the sense that there exists a continuous projection operator P of $\mathrm{L}(V)$ to $\mathcal{N}(G'(B))$ in virtue of [**168**, Example 43.16]. Thus in order for G to be a submersion (for details we refer to [**168**, Definition 43.15]), it remains to show that the mapping $G'(B) : \mathrm{L}(V) \to \mathbb{R}$ is surjective, i.e. $\mathcal{R}(G'(B)) = \mathbb{R}^2$. Whenever this condition holds and the operator B is a minimizer of the auxiliary variational principle (5.6.5), (5.6.4) (and thus a bound local minimum in the sense of [**168**, Definition 43.4]), we may apply [**168**, Theorem 43.D] to conclude the existence of $\Lambda \in Y^* = \mathbb{R}^2$ such that

$$F'(B)\,H - \Lambda(G'(B)\,H) = 0 \qquad \text{for all } H \in \mathrm{L}(V) ,$$

where Λ is called a *Lagrange multiplier*. In this case, the operator B is a critical point with respect to the side conditions N according to [**168**, Proposition 43.21].

5.6.3. Introducing Lagrange Multipliers. Making use of the auxiliary variational principle (5.6.5) with respect to the side conditions (5.6.4) as introduced in §5.6.2, this subsection is devoted to derive a useful representation of the operator $\hat{Q}$ in terms of Lagrange multipliers. Let us begin with the following lemma.

LEMMA 5.6.3. *Let $k \in \hat{\mathscr{M}}$ and $U \in \mathcal{B}(\hat{\mathscr{M}})$, and assume that the operator $B = B(k)$ is a minimizer of the variational principle* (5.6.5), (5.6.4) *such that $G'(B) : \mathrm{L}(V) \to \mathbb{R}^2$ is surjective. Then there exists a basis such that the matrix $X = X(k) = -\hat{Q}(k)S$ as defined above is given by*

$$X = -\beta\,\mathbb{1} - \alpha\,S + \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix}$$

for some matrix $Y \in \mathbb{C}^{(2n-p)\times(2n-p)}$ and real parameters $\alpha = \alpha(k,U)$, $\beta = \beta(k,U)$, where p denotes the number of non-trivial eigenvalues of $A = B^2$. Moreover, if B is invertible, $Y = 0$, and $\mathrm{Tr}_V(XBB^*) = \alpha\, f_U - \beta\, c_U$ (where the parameters f_U and c_U are given by (5.6.4)*).*

PROOF. Since $G'(B)$ is surjective, it is a submersion in view of [**168**, Definition 43.15]. For this reason, the fact that B is a minimizer (in the sense of [**168**, Definition 43.4])

[6] As $\mathrm{L}(V)$ is finite-dimensional, the Fréchet and Gâteaux derivative clearly coincide with the total and directional derivative, respectively.

implies that there exists a Lagrange multiplier $\Lambda = \Lambda(k, U) \in \mathbb{R}^2$ such that

$$F'(B)\,H - \Lambda(G'(B)\,H) = 0 \qquad \text{for all } H \in \mathrm{L}(V)\,.$$

Making use of the fact that the trace is cyclic, for any $H \in \mathrm{L}(V)$ we arrive at

$$\begin{aligned} 0 &= \mathrm{Tr}\,(XBH^* + XHB^*) - \lambda_1\ \mathrm{Tr}_V(BH^* + HB^*) - \lambda_2\ \mathrm{Tr}_V(SBH^* + SHB^*) \\ &= \mathrm{Tr}_V\left(\big(XB - \lambda_1\,B - \lambda_2\,SB\big)\,H^*\right) + \mathrm{Tr}_V\left(H\left(B^*X - \lambda_1\,B^* - \lambda_2\,B^*S\right)\right) \end{aligned}$$

with $\Lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2$, where $\lambda_{1/2} = \lambda_{1/2}(k, U)$. Considering variations $H = Z\,e^{i\varphi}$ in $\mathrm{L}(V)$ with $\varphi \in \mathbb{R}$ and arbitrary $Z \in \mathrm{L}(V)$, the last condition reads

$$0 = \mathrm{Tr}_V\left(\big(XB - \lambda_1\,B - \lambda_2\,SB\big)\,Z^*\,e^{-i\varphi}\right) + \mathrm{Tr}_V\left(Z\left(B^*X - \lambda_1\,B^* - \lambda_2\,B^*S\right)e^{i\varphi}\right)\,.$$

Since $Z \in \mathrm{L}(V)$ and $\varphi \in \mathbb{R}$ are arbitrary, we conclude that

$$\big(X - \lambda_1\,\mathbb{1} - \lambda_2\,S\big)\,B = 0 \qquad \text{as well as} \qquad B^*\big(X - \lambda_1\,\mathbb{1} - \lambda_2\,S\big) = 0\,.$$

We proceed by distinguishing the following two cases.

(i) If B is invertible, we obtain $X - \lambda_1\,\mathbb{1} - \lambda_2\,S = 0$, or equivalently,

$$X = \lambda_1\,\mathbb{1} + \lambda_2\,S\,.$$

As a consequence, for the matrix $A = BB^*$ we arrive at

$$\mathrm{Tr}_V(XA) = \lambda_1\ \mathrm{Tr}_V(BB^*) + \lambda_2\ \mathrm{Tr}_V(SBB^*) \overset{(5.6.4)}{=} \lambda_1\,f_U - \lambda_2\,c_U\,.$$

(ii) If B is not invertible, we note that the restriction to its image vanishes,

$$\big(X - \lambda_1\,\mathbb{1} - \lambda_2\,S\big)|_{\mathrm{im}(B)} = 0\,.$$

Moreover, for the matrix $A = BB^*$ we are given $\mathrm{im}(A) = \mathrm{im}(B)$, implying that

$$\big(X - \lambda_1\,\mathbb{1} - \lambda_2\,S\big)|_{\mathrm{im}(A)} = 0\,. \tag{5.6.7}$$

Since $A = BB^*$ is a positive operator with respect to $\langle\,.\,|\,.\,\rangle$, it is self-adjoint with respect to $\langle\,.\,|\,.\,\rangle$. According to [**114**, Theorem 4, p. 106], the operator A has real eigenvalues and the corresponding eigenvectors form an orthonormal basis of V. Consequently, the operator A is diagonalizable, i.e. there exists a basis of V such that A is diagonal (cf. [**108**, Theorem VIII.5.3 and Theorem VIII.5.4]). We thus can arrange that

$$A = \mathrm{diag}\,\big(\kappa_1, \ldots, \kappa_p, 0, \ldots, 0\big) \tag{5.6.8}$$

with $\kappa_1, \ldots, \kappa_p > 0$ due to positivity of A. In view of (5.6.7), we thus arrive at

$$X - \lambda_1\,\mathbb{1} - \lambda_2\,S = \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix}$$

for some $(2n-p) \times (2n-p)$-matrix Y.

Setting $\alpha = -\lambda_2$ and $\beta = -\lambda_1$ completes the proof. □

The next proposition shows that the matrix Y is positive.

PROPOSITION 5.6.4. *Under the assumptions of Lemma 5.6.3, the matrix Y given by*

$$X = -\beta\,\mathbb{1} - \alpha\,S + \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix}$$

is positive with respect to $\langle\,.\,|\,.\,\rangle$ in the sense that $\langle u \,|\, Yu\rangle \geq 0$ for all $u \in \mathbb{C}^{2n-p}$.

PROOF. It suffices to consider the case $Y \neq 0$. In view of (5.6.8), the matrix

$$B_0 := \operatorname{diag}\big(\sqrt{\kappa_1}, \ldots, \sqrt{\kappa_p}, 0, \ldots, 0\big)$$

satisfies $B_0^2 = A$. Given $\delta > 0$, for arbitrary matrices $D \in \mathbb{C}^{(2n-p)\times p}$ we now consider variations of B_0 of the form[7]

$$\tilde{B}_0(\tau) := B_0 + \tau \underbrace{\begin{pmatrix} 0 & D^* \\ D & 0 \end{pmatrix}}_{=:\,\mathcal{D}} \in \mathbb{C}^{2n\times 2n} \qquad \text{for all } \tau \in (-\delta, \delta)\,.$$

Multiplying out, for all $\tau \in (-\delta, \delta)$ we arrive at

$$\tilde{B}_0(\tau)^*\, \tilde{B}_0(\tau) = A + \tau \big(B_0^* \mathcal{D} + \mathcal{D}^* B_0\big) + \tau^2\, \mathcal{D}^*\mathcal{D} \in \mathbb{C}^{2n\times 2n}\,,$$

thus giving rise to variations of A. In order to restrict attention to variations being admissible with respect to the side conditions (5.6.3), we add an additional matrix, thus being led to consider variations of the form

$$\tilde{A}(\tau) := A + \tau\, B_0^* \mathcal{D} + \tau\, \mathcal{D}^* B_0 + \tau^2\, \mathcal{D}^*\mathcal{D} + \tau^2 \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$$

for all $\tau \in (-\delta, \delta)$. More precisely, in order for $\tilde{A}(\tau)$ to satisfy the constraints (5.6.3), we choose the matrix $E \in \mathbb{C}^{p\times p}$ such that $\operatorname{Tr}(\tilde{A}(\tau)) = f_U$ and $\operatorname{Tr}(S\tilde{A}(\tau)) = -c_U$ for all $\tau \in (-\delta, \delta)$. Multiplying out, we obtain

$$\tilde{A}(\tau) = A + \tau \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix} + \tau \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix} + \tau^2 \begin{pmatrix} D^*D & 0 \\ 0 & DD^* \end{pmatrix} + \tau^2 \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}$$

for all $\tau \in (-\delta, \delta)$, where the expressions

$$\tau \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix} + \tau \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix}$$

do not enter the constraints (5.6.3). As a consequence, the matrix E is determined by

$$\operatorname{Tr}_V \left\{ \begin{pmatrix} D^*D & 0 \\ 0 & DD^* \end{pmatrix} + \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \right\} \stackrel{!}{=} 0\,, \quad \operatorname{Tr}_V \left\{ S \begin{pmatrix} D^*D & 0 \\ 0 & DD^* \end{pmatrix} + S \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \right\} \stackrel{!}{=} 0\,.$$

As B is a minimizer of the variational principle (5.6.5), (5.6.4), the matrix $A = BB^*$ is a minimizer of the corresponding variational principle

$$\text{minimize } F(A) = \operatorname{Tr}_V(XA)$$

by varying A in $\mathrm{L}(V)$ with respect to the side conditions (5.6.3). In view of $\tilde{A}(0) = A$ and the fact that the expression

$$\operatorname{Tr}_V \big((\beta\, \mathbb{1} + \alpha\, S)\, \tilde{A}(\tau)\big) = \beta\, \operatorname{Tr}_V \big(\tilde{A}(\tau)\big) + \alpha\, \operatorname{Tr}_V \big(S\tilde{A}(\tau)\big) = \beta\, f_U - \alpha\, c_U$$

is constant for all $\tau \in (-\delta, \delta)$, we conclude that $\operatorname{Tr}_V \big((X + \beta\, \mathbb{1} + \alpha\, S)\, \tilde{A}(\tau)\big)$ is minimal for $\tau = 0$. Making use of (5.6.7), i.e. $(X + \beta\, \mathbb{1} + \alpha\, S)|_{\mathrm{im}(A)} = 0$, we infer the inequality

$$\operatorname{Tr}_V \big((X + \beta\, \mathbb{1} + \alpha\, S)\, \tilde{A}(\tau)\big) \geq 0 \qquad \text{for all } \tau \in (-\delta, \delta)\,.$$

[7]Note that $D^* \in \mathbb{C}^{p\times(2n-p)}$ if and only if $D \in \mathbb{C}^{(2n-p)\times p}$. As a consequence, neither D nor D^* will in general be self-adjoint, and $D^*D \in \mathbb{C}^{p\times p}$. Moreover, $(D^*)^* = D$ and

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}^* = \begin{pmatrix} B_{11}^* & B_{21}^* \\ B_{12}^* & B_{22}^* \end{pmatrix}$$

for block matrices B_{rs} with $r, s \in \{1, 2\}$.

In view of $A = \operatorname{diag}\big(\kappa_1, \ldots, \kappa_p, 0, \ldots, 0\big)$, for all $\tau \in (-\delta, \delta)$ this gives rise to

$$\begin{aligned} 0 &\leq \operatorname{Tr}_V \left((X + \beta\, \mathbb{1} + \alpha\, S)\, \tilde{A}(\tau) \right) \\ &= \operatorname{Tr}_V \left\{ \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} \left(A + \tau \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix} + \tau^2 \begin{pmatrix} D^* D & 0 \\ 0 & D D^* \end{pmatrix} \right) + \tau^2 \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \right\} \\ &= \operatorname{Tr}_V \left\{ \tau^2 \begin{pmatrix} 0 & 0 \\ 0 & Y D D^* \end{pmatrix} \right\} . \end{aligned}$$

As a result, we obtain $\operatorname{Tr}_{\mathbb{C}^{2n-p}} (Y D D^*) \geq 0$ for arbitrary matrices $D \in \mathbb{C}^{(2n-p) \times p}$. Moreover, from the fact that the matrices X, $\mathbb{1}$ and S are self-adjoint with respect to the Euclidean scalar product $\langle\, .\, |\, .\, \rangle_{\mathbb{C}^{2n}}$, we deduce that the matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} = X + \beta\, \mathbb{1} + \alpha\, S \in \mathbb{C}^{2n \times 2n}$$

is self-adjoint with respect to $\langle\, .\, |\, .\, \rangle_{\mathbb{C}^{2n}}$, too. Hence in order for Y to be positive, it remains to show that $\langle Y x \mid x \rangle_{\mathbb{C}^{2n-p}} \geq 0$ for all $x \in \mathbb{C}^{2n-p}$ (see e.g. [**157**, Section VI.3]). In view of

$$\operatorname{Tr}_{\mathbb{C}^{2n-p}}(XW) = \operatorname{Tr}_{\mathbb{C}^p}(WX)\,,$$

whenever $X \in \mathbb{C}^{(2n-p) \times p}$ and $W \in \mathbb{C}^{p \times (2n-p)}$, we deduce that

$$0 \leq \operatorname{Tr}_{\mathbb{C}^p} (D^* Y D) = \sum_{i=1}^{p} \langle e_i \mid D^* Y D\, e_i \rangle_{\mathbb{C}^p} = \sum_{i=1}^{p} \langle D\, e_i \mid Y D\, e_i \rangle_{\mathbb{C}^{2n-p}}\,,$$

where $(e_i)_{i=1,\ldots,p}$ is the standard orthonormal basis of $\mathbb{C}^p$. Choosing $D \in \mathbb{C}^{(2n-p) \times p}$ for arbitrary $x \in \mathbb{C}^{2n-p}$ according to

$$D = D(x) = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ x_{2n-p} & 0 & \cdots & 0 \end{pmatrix},$$

we obtain $D\, e_1 = x$ and $D\, e_i = 0$ for $i = 2, \ldots, p$. In particular,

$$0 \leq \langle x \mid Y x \rangle_{\mathbb{C}^{2n-p}}$$

for all $x \in \mathbb{C}^{2n-p}$, proving that $Y \geq 0$ as desired. □

In order for Lemma 5.6.3 and Proposition 5.6.4 to hold, it is essential that the Fréchet derivative $G'(B) = (G_1'(B), G_2'(B)) : \mathrm{L}(V) \to \mathbb{R}^2$ at $B \in \mathrm{L}(V)$ is regular in the sense that it is a surjective mapping from $\mathrm{L}(V)$ to $\mathbb{R}^2$. For this reason, our task is to work out conditions under which the mapping

$$\langle G_1'(B), \cdot \rangle\, b_1 + \langle G_2'(B), \cdot \rangle\, b_2 : \mathrm{L}(V) \to \mathbb{R}^2$$

is surjective, where $(b_i)_{i=1,2}$ is a basis of $\mathbb{R}^2$ and

$$\begin{aligned} \langle G_1'(B), H \rangle &= \operatorname{Tr}_V (B H^* + H B^*)\,, \\ \langle G_2'(B), H \rangle &= \operatorname{Tr}_V (S B H^* + S H B^*) \end{aligned}$$

for $B, H \in \mathrm{L}(V)$. Making use of the fact that the trace is linear and cyclic,

$$\operatorname{Tr}_V (\{B, H\}) = \operatorname{Tr}_V (BH + HB) = 2 \operatorname{Tr}_V (BH)$$

for all $H \in \mathrm{L}(V)$. Thus, assuming that $H = H^*$, the above mapping reads

$$\langle G'(B), H\rangle = \begin{pmatrix} \mathrm{Tr}_V(\{B,H\}) \\ \mathrm{Tr}_V(S\{B,H\}) \end{pmatrix} = \begin{pmatrix} 2\,\mathrm{Tr}_V(BH) \\ \mathrm{Tr}_V(\{S,B\}H) \end{pmatrix} ,$$

where $\{.,.\}$ denotes the anti-commutator. This leads us to the following result.

LEMMA 5.6.5. *Given $B \in \mathrm{L}(V)$, the mapping*

$$G'(B) : \mathrm{L}(V) \to \mathbb{R}^2 , \qquad H \mapsto \langle G'(B), H\rangle = \begin{pmatrix} 2\,\mathrm{Tr}_V(BH) \\ \mathrm{Tr}_V(\{S,B\}H) \end{pmatrix}$$

is surjective unless

$$B = \begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{or} \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} \tag{5.6.9}$$

for any block matrix $b \in \mathbb{C}^{n\times n}$.

PROOF. Let us first point out that the matrices B and $\{S,B\}$ are symmetric (with respect to $\langle\,.\,|\,.\,\rangle$). If B is of the form (5.6.9), then

$$\{S,B\} = 2B \qquad \text{or} \qquad \{S,B\} = -2B ,$$

i.e. B and $\{S,B\}$ are linearly dependent, implying that $G'(B)$ is not surjective. On the other hand, if B is not of the form (5.6.9), the operators B and $\{S,B\}$ are linearly independent within the set of all symmetric matrices on V. Namely,

$$\begin{aligned} \alpha\{S,B\} + \beta B &= \alpha \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} * & * \\ * & * \end{pmatrix} + \alpha \begin{pmatrix} * & * \\ * & * \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} + \beta \begin{pmatrix} * & * \\ * & * \end{pmatrix} \\ &= 2\alpha \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} + \beta \begin{pmatrix} * & * \\ * & * \end{pmatrix} \stackrel{!}{=} 0 \end{aligned}$$

if and only if $\alpha = 0 = \beta$. Note that the set of symmetric matrices on V forms a real vector space; therefore, it is a real Hilbert space endowed with the Hilbert-Schmidt scalar product

$$\langle M, M'\rangle_{\mathrm{HS}} := \mathrm{Tr}_V(MM') .$$

In particular, if B is not of the form (5.6.9), $B \neq 0$. As a consequence, there exist symmetric matrices M_1, M_2 such that

$$\begin{aligned} \mathrm{Tr}_V(BM_1) = 1 , \qquad & \mathrm{Tr}_V(BM_2) = 0 , \\ \mathrm{Tr}_V(\{S,B\}M_1) = 0 , \qquad & \mathrm{Tr}_V(\{S,B\}M_2) = 1 \end{aligned}$$

(i.e. $B \perp_{\mathrm{HS}} M_2$ and $\{S,B\} \perp_{\mathrm{HS}} M_1$ with respect to $\langle\,.\,|\,.\,\rangle_{\mathrm{HS}}$). Choosing the symmetric operator $H = \alpha/2\,M_1 + \beta\,M_2$, we obtain

$$\langle G'(B), H\rangle = \langle G_1'(B), H\rangle\, b_1 + \langle G_2'(B), H\rangle\, b_2 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} .$$

Since $\alpha, \beta \in \mathbb{R}$ were arbitrary, we conclude that $G'(B)$ is indeed surjective. □

COROLLARY 5.6.6. *Let $k \in \hat{\mathscr{M}}$ and $U \in \mathcal{B}(\hat{\mathscr{M}})$, and assume that the operator $B = B(k)$ is a minimizer of the variational principle* (5.6.5), (5.6.4). *If the operator B is of the form* (5.6.9) *and $f_U \neq 0$ in* (5.6.2), *then there exists a basis such that the matrix $X = X(k) = -\hat{Q}(k)S$ as defined above is given by*

$$X = \lambda\,\mathbb{1} + \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix}$$

for some positive $(2n-p)\times(2n-p)$*-matrix* Y *and some real parameter* $\lambda=\lambda(k,U)$*, where* p *denotes the number of non-trivial eigenvalues of* $A=B^2$*. Furthermore, if* B *is invertible,* $Y=0$*, and* $\mathrm{Tr}_V(XBB^*)=\lambda\, f_U$*.*

PROOF. If B is of the form (5.6.9), we know that A is of the form

$$A=\begin{pmatrix} b^2 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{or} \qquad A=\begin{pmatrix} 0 & 0 \\ 0 & b^2 \end{pmatrix}.$$

In the first case, $\mathrm{Tr}_V(A)=\mathrm{Tr}_V(SA)$, leaving us with only one constraint. In the second case, $\mathrm{Tr}_V(A)=-\mathrm{Tr}_V(SA)$, giving rise to $f_U=\mathrm{Tr}_V(A)=c_U$. Thus we also end up with only one constraint, which is implemented by considering the functional

$$\tilde{G}:\mathrm{L}(V)\to\mathbb{R}\,, \qquad B\mapsto\tilde{G}(B):=\mathrm{Tr}_V(BB^*)-f_U\,.$$

As a consequence, the Fréchet derivative $\tilde{G}'(B)$ is surjective if $\mathrm{Tr}_V(B)\neq 0$; namely, considering $H=\varphi\mathbb{1}$ with $\varphi\in\mathbb{R}$, the Fréchet derivative is given by

$$\langle\tilde{G}'(B),H\rangle=\mathrm{Tr}_V\left(BH^*+HB^*\right)=2\varphi\,\mathrm{Tr}_V(B)\,.$$

Since $\mathrm{Tr}_V(A)=\mathrm{Tr}_V(B^2)=f_U$ and A is positive (with respect to $\langle\,.\,|\,.\,\rangle$), we obtain the equivalence $\mathrm{Tr}_V(B)=0$ if and only if $B=0$ if and only if $f_U=0$. Thus $f_U\neq 0$ in (5.6.2) yields surjectivity of the Fréchet derivative. As a consequence, there exists a Lagrange multiplier $\lambda=\lambda(k,U)$ according to [**168**, Theorem 43.D] such that

$$\langle F'(B),H\rangle-\lambda\,\langle\tilde{G}'(B),H\rangle=0$$

for all $H\in\mathrm{L}(V)$. Making use of the fact that the trace is cyclic, we obtain

$$\begin{aligned}0&=\mathrm{Tr}_V(XBH^*+XHB^*)-\lambda\,\mathrm{Tr}_V(BH^*+HB^*)\\&=\mathrm{Tr}_V\left((XB-\lambda B)\,H^*\right)+\mathrm{Tr}_V\left(H\,(B^*X-\lambda B^*)\right).\end{aligned}$$

Considering variations $H=Ze^{i\varphi}$ with $\varphi\in\mathbb{R}$ and arbitrary $Z\in\mathrm{L}(V)$ in analogy to the proof of Lemma 5.6.3, the latter condition reads

$$0=\mathrm{Tr}_V\left((XB-\lambda B)\,Z\,e^{-i\varphi}\right)+\mathrm{Tr}_V\left(Z\,(B^*X-\lambda B^*)\,e^{i\varphi}\right).$$

Since $Z\in\mathrm{L}(V)$ and $\varphi\in\mathbb{R}$ are arbitrary, we conclude that

$$\big(X-\lambda\mathbb{1}\big)\,B=0 \qquad \text{as well as} \qquad B^*\big(X-\lambda\mathbb{1}\big)=0\,.$$

We again distinguish the following two cases.

(i) If B is invertible, we obtain

$$X=\lambda\mathbb{1}\,,$$

giving rise to

$$\mathrm{Tr}_V(XA)=\lambda\,\mathrm{Tr}_V(BB^*)=\lambda\,f_U\,.$$

(ii) If B is not invertible, proceeding in analogy to the proof of Lemma 5.6.3 yields

$$X-\lambda\mathbb{1}=\begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix}$$

for some $(2n-p)\times(2n-p)$-matrix Y, where p denotes the number of non-trivial eigenvalues of $A=BB^*$. Proceeding similarly to the proof of Proposition 5.6.4, we conclude that $Y\geq 0$.

This completes the proof. □

Making use of $X(k) = -\hat{Q}(k)S$ and setting $\beta = -\lambda$, the operator $\hat{Q}$ has the form

$$\hat{Q} = -XS = -\left(\lambda \mathbb{1} + \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix}\right) S = \beta S - \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} S$$

with $Y \geq 0$ and $Y = 0$ if B is invertible. Introducing the matrix $R(k)$ by

$$R(k) = -\begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} S\,,$$

the following proposition shows that $R(k)$ is negative definite with respect to the indefinite inner product $\prec . \,|\, . \succ$.

PROPOSITION 5.6.7. *Let $k \in \hat{\mathscr{M}}$ and $U \in \mathcal{B}(\hat{\mathscr{M}})$, and assume that the operator $B = B(k)$ is a minimizer of the variational principle* (5.6.5), (5.6.4). *Then the operator*

$$R(k) := \hat{Q}(k) - \beta(k)\, S - \alpha(k)\, \mathbb{1} = -\begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} S \tag{5.6.10}$$

is negative definite with respect to $\prec . \,|\, . \succ$ (where $\alpha = 0$ if B is of the form (5.6.9)*). Moreover, $R(k)\, A' = 0 = A'\, R(k)$, where $A' = -SA$ satisfies* (5.6.2) *and $A = B^2$.*

PROOF. In view of $Y \geq 0$ (see Proposition 5.6.4 and Corollary 5.6.6), for all $u \in V$ holds

$$\prec u \,|\, -R(k)\, u \succ = \langle u \,|\, -SR(k)\, u \rangle = \langle u \,|\, S \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} Su \rangle = \langle Su \,|\, \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} Su \rangle \geq 0\,.$$

In view of (5.6.8) one immediately obtains

$$R(k)\, A' = -\begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} S(-S)\, A = \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix} = 0$$

as well as

$$A'\, R(k) = (-S)A \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} (-S) = 0\,,$$

which completes the proof. □

Taken the previous results together, we deduce that, for real parameters $\alpha = \alpha(k)$ and $\beta = \beta(k)$, the operator $X = X(k)$ is of the general form

$$X = -\beta\, \mathbb{1} - \alpha\, S + \begin{pmatrix} 0 & 0 \\ 0 & Y \end{pmatrix} \qquad \text{with } Y \geq 0\,.$$

In view of $X(k) = -\hat{Q}(k)S$ and (5.6.10), the operator $\hat{Q} = \hat{Q}(k)$ can be written as

$$\hat{Q}(k) = -X(k)S = \beta(k)\, S + \alpha(k)\, \mathbb{1} + R(k)$$

with $Y = 0$ if B is invertible, and $\alpha = 0$ if B is of the form (5.6.9) and $f_U \neq 0$ in (5.6.2). In the standard pseudo-orthonormal basis, the signature matrix S takes the familiar form (5.2.2).

For convenience, let us summarize the above results in the following theorem.

THEOREM 5.6.8. *Let $d\nu \in \mathfrak{Ndm}$ be a minimizer of the causal variational principle* (5.4.8) *with respect to the side conditions* (5.4.7). *Then for any $k \in \hat{\mathscr{M}}$ and $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$ there is an operator $A' = A'(k, \Omega)$ in $\mathrm{L}(V)$ being a minimizer of the variational principle*

$$\textit{minimize} \qquad \mathrm{Tr}_V \left(\hat{Q}(k)\, \tilde{A}'\right) \qquad \textit{under the constraints } (5.6.2)$$

in the class of negative definite matrices on $\mathrm{L}(V)$*. Moreover, we have* $R(k) \circ A' = 0$ *as well as* $[A', R(k)] = 0$*, where the matrix* $R(k) = R(k, \Omega)$ *is defined by* (5.6.10)*, that is*

$$R(k) = \hat{Q}(k) - \beta(k)\, S - \alpha(k)\, \mathbb{1}$$

with real parameters $\alpha(k) = \alpha(k, \Omega)$ *and* $\beta(k) = \beta(k, \Omega)$*.*

5.7. Euler-Lagrange Equations in Momentum Space

The aim of this section is to derive Euler-Lagrange equations in momentum space. Throughout this section we shall assume that $d\nu$ is a negative definite measure on momentum space $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ which is a minimizer (see Definition 5.4.9) of the causal variational principle (5.4.8) under the side conditions (5.6.1). Unless stated otherwise, by $(\mathfrak{e}_i)_{i=1,\ldots,2n}$ we denote the standard pseudo-orthonormal basis of V in such a way that the signature matrix S takes the convenient form (5.2.2).

Essentially making use of the matrices A' obtained in Theorem 5.6.8, our method is to proceed in the spirit of [**12**]. More precisely, we first outline some preliminary considerations in order to clarify the connection to Theorem 5.6.8 (§5.7.1). Afterwards, we introduce strictly negative definite measures in the strong sense (§5.7.2). This allows us to state our main result (Theorem 5.7.10). For the proof of this theorem we proceed similarly to [**12**, Section 3] by first deriving an equality on the support of $d\nu$ (§5.7.3) and then deducing an inequality on whole momentum space $\hat{\mathscr{M}}$ (§5.7.4).

5.7.1. Preliminary Considerations. Throughout this section we shall assume that (5.5.2) holds and that the Lagrangian is a continuously differentiable mapping in the sense of Assumption 5.5.4.

Following the explanations in §5.5.4, we denote the variation measure of $d\nu$ by

$$d\zeta := d|\nu| \,.$$

Henceforth, the negative definite measure $d\nu$ is absolutely continuous with respect to $d\zeta$ (in symbols, $d\nu \ll d\zeta$), and there is $\Upsilon \in L^1(\hat{\mathscr{M}}, \mathrm{L}(V); d\zeta)$ such that

$$d\nu = \Upsilon\, d\zeta$$

(in view of the Radon-Nikodým theorem). According to Definition 5.3.7, the support of $d\nu$ is given by

$$\operatorname{supp} d\nu = \operatorname{supp} d\zeta \,,$$

implying that

$$\int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \big(\Upsilon(k)\big)\, d\zeta(k) = \int_{\operatorname{supp} d\nu} \mathrm{Tr}_V \big(\Upsilon(k)\big)\, d\zeta(k) \,.$$

Starting point in what follows is the first variation $\delta\mathcal{S}$ as worked out in §5.5.4. More precisely, let $Q = Q[\nu]$ be given by (5.5.7) and denote by $\hat{Q}$ its Fourier transform. Whenever $Q \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^\infty(\mathscr{M}, \mathrm{L}(V))$, the first variation of $\mathcal{S}$ in direction $d\eta \in \mathfrak{Dvm}$ takes the form

$$\delta\mathcal{S}(\nu; \eta) = 4 \int_{\hat{\mathscr{M}}} \mathrm{Tr}_V \big(\hat{Q}[\nu](q)\, d\eta(q)\big) \tag{5.7.1}$$

according to Lemma 5.5.13. Note that the first variation (5.7.1) depends crucially on the operator-valued measure $d\eta \in \mathfrak{Dvm}$. For this reason, in the following we always assume that $Q \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^\infty(\mathscr{M}, \mathrm{L}(V))$. In this case, by Lemma 5.5.13

we obtain $\hat{Q} \in C_0(\hat{\mathscr{M}}; \mathrm{L}(V)) \cap L^2(\hat{\mathscr{M}}, \mathrm{L}(V))$, and from $C_0(\hat{\mathscr{M}}; \mathrm{L}(V)) \subset C_b(\hat{\mathscr{M}}; \mathrm{L}(V))$ we conclude that

$$\hat{Q} \in L^2(\hat{\mathscr{M}}, \mathrm{L}(V); d\zeta) \,.$$

Furthermore, we let $A' = A'(\cdot, \hat{\mathscr{M}})$ be the minimizer given by Theorem 5.6.8. Due to the constructions in Section 5.6, for every $k \in \hat{\mathscr{M}}$ there is $B(k) \in \mathrm{L}(V)$ such that

$$A'(k) = -SB(k)^2 \,,$$

where $B(k)$ is bounded according to (5.6.6). As a consequence,

$$A' \in L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V); d\zeta) \cap L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V)) \,.$$

Next, we consider the operator R as given by (5.6.10),

$$R(k) = \hat{Q}(k) - \alpha(k)\, \mathbb{1} - \beta(k)\, S \qquad \text{for all } k \in \hat{\mathscr{M}} \,,$$

where $\alpha(k) = \alpha(k, \hat{\mathscr{M}})$ and $\beta(k) = \beta(k, \hat{\mathscr{M}})$ for every $k \in \hat{\mathscr{M}}$ denote the Lagrange multipliers corresponding to the auxiliary variational principle in §5.6.2. In virtue of Theorem 5.6.8 we know that $R(k) \circ A'(k) = 0$ for each $k \in \mathbb{N}$. Moreover, for technical simplicity in what follows *we assume that $A'(k)$ depends continuously on $k \in \hat{\mathscr{M}}$*. In this case, the mapping

$$\hat{\mathscr{M}} \ni k \mapsto \hat{\mathfrak{q}}(k) := \mathrm{Tr}_V \left(\hat{Q}(k)\, A'(k) \right) \tag{5.7.2}$$

is continuous. Due to Theorem 5.6.8, the mapping (5.7.2) equivalently reads

$$\hat{\mathscr{M}} \ni k \mapsto \hat{\mathfrak{q}}(k) = \mathrm{Tr}_V \left(\alpha(k)\, A'(k) + \beta(k)\, S\, A'(k) \right) .$$

5.7.2. Strictly Negative Definite Measures. The subtle point for the subsequent constructions to work is to prove that the side conditions are satisfied. In particular, we need to vary exclusively within the class of negative definite measures $\mathfrak{Ndm}$. For a convenient terminology, let us first state the following definition:

Definition 5.7.1. *Let $d\nu \in \mathfrak{Ndm}$ be a negative definite measure. Given $d\eta \in \mathfrak{Dvm}$ with*

$$\mathrm{Tr}_V(\eta(\hat{\mathscr{M}})) = 0 \qquad \textit{and} \qquad \mathrm{Tr}_V(-S\eta(\hat{\mathscr{M}})) = 0 \,, \tag{5.7.3}$$

*the variations $d\tilde{\nu}_\tau := d\nu + \tau\, d\eta$ are said to be **admissible** if there are $a \leq 0$ and $b > 0$ such that $d\tilde{\nu}_\tau \in \mathfrak{Ndm}$ for all $\tau \in (a, b)$*

Remark 5.7.2. *Whenever $d\nu \in \mathfrak{Ndm}$ satisfies the constraints*

$$\mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = c \qquad \textit{and} \qquad \mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f \,, \tag{5.7.4}$$

and $d\tilde{\nu}_\tau$ is an admissible variation, the calculations

$$\mathrm{Tr}_V(\tilde{\nu}_\tau(\hat{\mathscr{M}})) = c + \tau\, \mathrm{Tr}_V(\eta(\hat{\mathscr{M}})) = c \,,$$
$$\mathrm{Tr}_V(-S\tilde{\nu}_\tau(\hat{\mathscr{M}})) = f + \tau\, \mathrm{Tr}_V(-S\eta(\hat{\mathscr{M}})) = f$$

show that admissible variations satisfy the constraints (5.7.4).

In order for deriving an equality on the support of $d\nu$ as well as an inequality on whole momentum space, we need to ensure that sufficiently small variations of $d\nu$ in certain directions still give rise to negative definite measures. More precisely, in order to

derive an equality we assume that $x_0, y_0 \in \operatorname{supp} d\nu$ and $g \in C_c(\hat{\mathscr{M}})$ are given; then for sufficiently small $\varepsilon > 0$ and $\delta > 0$ we require the measures

$$d\tilde{\nu}_\tau := d\nu + \tau\, d\eta \qquad \text{for all } \tau \in (-\delta, \delta)$$

to be negative definite, where $d\eta = d\eta(x_0, y_0, \varepsilon, g)$ is given by

$$d\eta := \left(\int_{\hat{\mathscr{M}}} g(z)\, d\zeta(z) \right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta - \left(\int_{\hat{\mathscr{M}}} g(z)\, d\zeta(z) \right) A'(y_0)\, \eta_{y_0,\varepsilon}\, d\zeta\,, \tag{5.7.5}$$

and $\eta_{x_0,\varepsilon} \in C_c^+(B_\varepsilon(x_0))$, $\eta_{y_0,\varepsilon} \in C_c^+(B_\varepsilon(y_0))$ are compactly supported functions with

$$\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta = 1\,, \qquad \int_{B_\varepsilon(y_0)} \eta_{y_0,\varepsilon}\, d\zeta = 1\,.$$

On the other hand, for deriving an inequality we let $x_0 \in \operatorname{supp} d\nu$ and $g \in C_c^+(\hat{\mathscr{M}})$ be given; in this case we need to guarantee that, for sufficiently small $\varepsilon > 0$ and $\delta > 0$, the measures

$$d\tilde{\nu}_\tau := d\nu + \tau\, d\eta \qquad \text{for all } \tau \in (0, \delta)$$

are negative definite, where $d\eta_g = d\eta_g(x_0, \varepsilon)$ is given by

$$d\eta_g := -\left(\int_{\hat{\mathscr{M}}} g(k)\, d^4k \right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta + d \int_{\hat{\mathscr{M}}} g(k)\, A'(k)\, \delta_k(\cdot)\, d^4k \tag{5.7.6}$$

with $\eta_{x_0,\varepsilon} \in C_c^+(B_\varepsilon(x_0))$ being a compactly supported function such that

$$\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta = 1\,,$$

and $d\delta_k$ being the Dirac measure supported at $k \in \hat{\mathscr{M}}$.

For simplicity, we only focus on the first case (5.7.5). Then our task is to show that there is $\delta > 0$ in such a way that for all $u \in V \setminus \{0\}$ and $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$,

$$\prec u \,|\, -\tilde{\nu}_\tau(\Omega)\, u \succ \geq 0 \qquad \text{for all } \tau \in (-\delta, \delta)\,. \tag{5.7.7}$$

Without loss of generality it suffices to restrict attention to $u \in \mathbb{S}$, where

$$\mathbb{S} := \{u \in V : \|u\| = 1\}\,.$$

Our strategy for proving (5.7.7) is to show that there is a constant $c > 0$, independent of $u \in \mathbb{S}$ and $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$, with the property that

$$\inf_{u \in \mathbb{S}} \int_\Omega \prec u \,|\, -\Upsilon(k)\, u \succ\; d\zeta(k) > c\, \zeta(\Omega) \qquad \text{whenever } \zeta(\Omega) > 0\,. \tag{5.7.8}$$

In this case we may prove (5.7.7) by contradiction: Assuming conversely that (5.7.7) is not true, then for any $\delta > 0$ there is $\tau_0 \in (-\delta, \delta)$, $u_0 \in \mathbb{S}$ and $\Omega_0 \in \mathcal{B}(\hat{\mathscr{M}})$ such that

$$\prec u_0 \,|\, -\tilde{\nu}_{\tau_0}(\Omega_0)\, u_0 \succ < 0\,.$$

However, choosing

$$\delta < \|A'\|^{-1}_{L^\infty(\hat{\mathscr{M}}, \mathrm{L}(V); d\zeta)}\, \|g\|^{-1}_{L^\infty(\hat{\mathscr{M}})}\, \|\eta_{x_0,\varepsilon}\|^{-1}_{L^\infty(\hat{\mathscr{M}})}\, c$$

we obtain the contradiction

$$\begin{aligned}
0 < c\,\zeta(\Omega_0) &\leq \prec u_0 \mid -\nu(\Omega_0)\,u_0 \succ \\
&\leq \prec u_0 \mid -\nu(\Omega_0)\,u_0 \succ + \tau \prec u_0 \mid -\int_{\Omega_0} \left(\int_{\hat{\mathscr{M}}} g(k)\, d^4k \right) A'(y_0)\, \eta_{y_0,\varepsilon}\, d\zeta\, u_0 \succ \\
&< \tau \prec u_0 \mid -\int_{\Omega_0} \left(\int_{\hat{\mathscr{M}}} g(k)\, d^4k \right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta\, u_0 \succ < c\,\zeta(\Omega_0)\,.
\end{aligned}$$

This yields (5.7.7), provided the existence of a constant $c > 0$ satisfying (5.7.8).

In order for condition (5.7.8) to hold, let us introduce *strictly negative definite measures in the strong sense* as follows:

DEFINITION 5.7.3. *A negative definite measure $d\nu = \Upsilon\, d\zeta$ is called* ***strictly negative definite in the strong sense*** *if and only if Υ is continuous and the mapping*

$$Y_u : \hat{\mathscr{M}} \to \mathbb{R}\,, \qquad x \mapsto Y_u(x) := \prec u \mid -\Upsilon(x)\,u \succ$$

is strictly positive $d\zeta$-almost everywhere for all $u \in V \setminus \{0\}$, where $d\zeta = d|\nu|$.

Let $d\nu$ be strictly negative definite in the strong sense and assume that $\nu(\Omega) \neq 0$ for some Borel set $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$. Then $\Omega \cap \operatorname{supp} d\nu \neq \varnothing$, implying that $\zeta(\Omega) > 0$. Whenever Ω is relatively compact, by continuity of Y_u we have

$$\prec u \mid -\nu(\Omega) u \succ = \int_\Omega Y_u(z)\, d\zeta(z) \geq \inf_{z \in \Omega \setminus E} Y_u(z)\, \zeta(\Omega) > 0$$

for any $u \in V \setminus \{0\}$. Whenever $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$ is a general Borel set with $\Omega \cap \operatorname{supp} d\nu \neq \varnothing$ we may restrict attention to relatively compact subsets thereof. Applying the same arguments as before yields $\prec u \mid -\nu(\Omega) u \succ$ for any $u \in V \setminus \{0\}$. Accordingly, we conclude that every strictly negative definite measure in the strong sense is *strictly negative definite* in the following sense:

DEFINITION 5.7.4. *A negative definite measure $d\nu$ is called* ***strictly negative definite*** *if the following condition is satisfied: Whenever $\Omega \in \mathcal{B}(\hat{\mathscr{M}})$ such that $\nu(\Omega) \neq 0$,*

$$\prec u \mid -\nu(\Omega)\, u \succ > 0 \qquad \text{for all } u \in V \setminus \{0\}\,.$$

For technical simplicity, in the following we shall only restrict attention to strictly negative definite measures in the strong sense. Applying a rescaling argument, it suffices to restrict attention to $\mathbb{S}$ instead of $V \setminus \{0\}$. According to Definition 5.7.3, for every $u \in \mathbb{S}$ we then introduce the mapping Y_u by

$$Y_u : \hat{\mathscr{M}} \to \mathbb{R}\,, \qquad x \mapsto Y_u(x) = \prec u \mid -\Upsilon(x)\, u \succ$$

such that $Y_u(x) > 0$ for all $x \in \hat{\mathscr{M}} \setminus E$, where E is a ζ-zero set. Whenever $B \subset \hat{\mathscr{M}}$ is relatively compact and $B \cap \operatorname{supp} d\zeta \neq \varnothing$, we infer that the quantity $\gamma = \gamma(B)$ defined by

$$\gamma := \inf_{u \in \mathbb{S}} \inf_{x \in B \setminus E} Y_u(x)$$

is strictly positive. Introducing the constant $c := \gamma/2$ yields the following result:

LEMMA 5.7.5. *Assume that $d\nu$ is a strictly negative definite measure in the strong sense. Then for every relatively compact subset $B \subset \hat{\mathscr{M}}$ there is a constant $c = c(B) > 0$ with the property* (5.7.8)*, that is, for any Borel set $\Omega \subset B$ holds*

$$\inf_{u \in \mathbb{S}} \int_\Omega \prec u \mid -\Upsilon(k)\, u \succ \; d\zeta(k) > c\,\zeta(\Omega) \qquad \text{whenever } \zeta(\Omega) > 0\,.$$

PROOF. Whenever $\zeta(\Omega) > 0$, this is a consequence of

$$\inf_{u \in \mathbb{S}} \int_\Omega \prec u \,|\, -\Upsilon(k)\, u \succ \, d\zeta(k) \geq \inf_{u \in \mathbb{S}} \inf_{x \in \Omega \setminus E} Y_u(x)\, \zeta(\Omega \setminus E) > \gamma/2\, \zeta(\Omega)\,,$$

which completes the proof. □

Lemma 5.7.5 asserts that property (5.7.8) is satisfied for some constant $c > 0$. As a consequence, there is some $\delta > 0$ in such a way that the variations $(d\tilde{\nu}_\tau)_{\tau \in (-\delta, \delta)}$ are admissible (5.7.7). Similar arguments can be applied for the variations (5.7.6).

5.7.3. First Variations with Fixed Support. We are now in the position to derive an equality on the support of $d\zeta = d|\nu|$ whenever $d\nu$ is strictly negative definite in the strong sense (see Definition 5.7.3). To this end, we shall apply the following result, which can be considered as a generalization of the fundamental lemma of calculus of variations (see e.g. [**93**, Lemma 1.1.1]) for continuous functions on locally compact topological Hausdorff spaces.

LEMMA 5.7.6. *Let X be a locally compact topological Hausdorff space, and let $d\mu$ be a positive Radon measure on $\mathcal{B}(X)$. Suppose that $f \in C(X)$ satisfies*

$$\int_X f(x)\, \eta(x)\, d\mu(x) \geq 0 \qquad \text{for all } \eta \in C_c(X) \text{ with } \eta \geq 0 \tag{5.7.9}$$

or

$$\int_X f(x)\, \eta(x)\, d\mu(x) = 0 \qquad \text{for all } \eta \in C_c(X)\,.$$

Then $f(x) \geq 0$ or $f(x) = 0$ for all $x \in X$, respectively.

PROOF. First of all, the measure $d\mu$ has support in view of [**43**, §2.2.5]. Next, assume conversely that there is $x_0 \in \operatorname{supp} d\mu$ such that $f(x_0) < 0$. Since $f \in C(X)$, there is an open set $U \subset X$ containing x_0 such that $f(x) < f(x_0)/2$ for all $x \in U$. Moreover, let $\eta \in C_c(X)^+$ such that $\operatorname{supp} \eta \subset V$ and $\eta(x_0) > 0$. Choosing $W \subset \operatorname{supp} \eta$ open with $x_0 \in W$ and $\eta(x) > \eta(x_0)/2$ for all $x \in W$, according to $\mu(W) > 0$ we obtain

$$\int_V (-f)\, \eta \, d\mu \geq \int_W (-f)\, \eta \, d\mu \geq -\frac{1}{4} f(x_0)\, \eta(x_0)\, \mu(W) > 0\,,$$

giving rise to the contradiction

$$0 \overset{(5.7.9)}{\leq} \int_X f\, \eta \, d\mu = \int_V f\, \eta \, d\mu < 0\,.$$

Hence $f(x) \geq 0$ for all $x \in X$. The second assertion of the lemma follows readily from the first one. □

This allows us to state the main result of this subsection:

LEMMA 5.7.7. *Let $Q \in L^1(\mathscr{M}, \mathrm{L}(V)) \cap L^2(\mathscr{M}, \mathrm{L}(V)) \cap L^\infty(\mathscr{M}, \mathrm{L}(V))$, and let $d\nu \in \mathfrak{Ndm}$ be minimizer of the causal variational principle* (5.4.8), (5.4.7). *Whenever $d\nu$ is* strictly *negative definite* in the strong sense, *for any $x_0, y_0 \in \operatorname{supp} d\nu$ we are given*

$$\operatorname{Tr}_V \left(\hat{Q}(x_0)\, A'(x_0) \right) = \operatorname{Tr}_V \left(\hat{Q}(y_0)\, A'(y_0) \right);$$

in other words, there is a constant $c(A')$ such that $\operatorname{Tr}_V \left(\hat{Q}\, A' \right)|_{\operatorname{supp} d\nu} = c(A')$.

PROOF. Given arbitrary $x_0, y_0 \in \operatorname{supp} d\nu$ and $\varepsilon > 0$, we consider compactly supported functions $\eta_{x_0,\varepsilon} \in C_c^+(B_\varepsilon(x_0))$ and $\eta_{y_0,\varepsilon} \in C_c^+(B_\varepsilon(y_0))$ with the property that

$$\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta = 1 \qquad \text{and} \qquad \int_{B_\varepsilon(y_0)} \eta_{y_0,\varepsilon}\, d\zeta = 1\,.$$

For arbitrary $g \in C_c(\hat{\mathscr{M}})$, we then consider variations of $d\nu$ in direction $d\eta = d\eta_g$ with

$$d\eta := \left(\int_{\hat{\mathscr{M}}} g(z)\, d\zeta(z)\right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta - \left(\int_{\hat{\mathscr{M}}} g(z)\, d\zeta(z)\right) A'(y_0)\, \eta_{y_0,\varepsilon}\, d\zeta$$

(cf. (5.7.5)). These variations clearly leave the support of $d\nu$ unaffected. Since $d\nu$ is supposed to be strictly negative definite in the strong sense, there is $\delta = \delta(\eta)$ such that

$$d\tilde{\nu}_\tau := d\nu + \tau\, d\eta \in \mathfrak{Ndm}$$

is negative definite for all $\tau \in (-\delta, \delta)$, and the side conditions (5.7.3) reduce to

$$\operatorname{Tr}_V\big(A'(x_0)\big) = \operatorname{Tr}_V\big(A'(y_0)\big) \qquad \text{and} \qquad \operatorname{Tr}_V\big(-SA'(x_0)\big) = \operatorname{Tr}_V\big(-SA'(y_0)\big)\,,$$

which is immediately satisfied in view of (5.6.2). On the other hand, from (5.5.13) we get

$$\mathcal{S}(\tilde{\nu}_\tau) \geq \mathcal{S}(\nu) \qquad \text{for all } \tau \in (-\delta, \delta)\,,$$

implying that

$$\delta\mathcal{S}(\nu;\eta) = \frac{d}{d\tau}\mathcal{S}(\tilde{\nu}_\tau)|_{\tau=0} = 0\,.$$

Making use of Fubini's theorem, we arrive at

$$0 = \delta\mathcal{S}(\nu;\eta) \overset{(5.7.1)}{=} 4 \int_{\hat{\mathscr{M}}} \operatorname{Tr}_V\big(\hat{Q}[\nu](q)\, d\eta(q)\big) = \int_{\hat{\mathscr{M}}} g(z)\ \operatorname{Tr}_V\left(h_\varepsilon(x_0,y_0)\right)\ d\zeta(z)\,,$$

where $h_\varepsilon(x_0,y_0)$ is the constant expression defined by

$$h_\varepsilon(x_0,y_0) := \int_{B_\varepsilon(x_0)} \hat{Q}(q)\, A'(x_0)\, \eta_{x_0,\varepsilon}(q)\, d\zeta(q) - \int_{B_\varepsilon(y_0)} \hat{Q}(q)\, A'(y_0)\, \eta_{y_0,\varepsilon}(q)\, d\zeta(q)\,.$$

Since $g \in C_c(\hat{\mathscr{M}})$ was arbitrary, by applying Lemma 5.7.6 we conclude that

$$\int_{B_\varepsilon(x_0)} \hat{Q}(q)\, A'(x_0) \eta_{x_0,\varepsilon}(q)\, d\zeta(q) = \int_{B_\varepsilon(y_0)} \hat{Q}(q)\, A'(y_0)\, \eta_{y_0,\varepsilon}(q)\, d\zeta(q)\,.$$

Applying the mean value theorem for integrals [**87**, §27], in the limit $\varepsilon \searrow 0$ we obtain

$$\operatorname{Tr}_V\big(\hat{Q}(x_0)\, A'(x_0)\big) = \operatorname{Tr}_V\big(\hat{Q}(y_0)\, A'(y_0)\big)\,.$$

Since $x_0, y_0 \in \operatorname{supp} d\nu$ are arbitrary, the claim follows. □

5.7.4. First Variations with Varying Support. In this subsection we finally derive the desired Euler-Lagrange equations in momentum space. To this end, we essentially employ the fundamental lemma of calculus of variations, which for convenience we briefly recall (see e.g. [**78**, Chapter 1, Lemma 3]).

LEMMA 5.7.8. *Let $\Omega \subset \mathbb{R}^n$ open. Suppose that f is of class $L^1(\Omega)$ and satisfies*

$$\int_\Omega f(x)\, \eta(x)\, dx \geq 0 \qquad \textit{for all } \eta \in C_c^\infty(\Omega) \textit{ with } \eta \geq 0$$

or

$$\int_\Omega f(x)\, \eta(x)\, dx = 0 \qquad \textit{for all } \eta \in C_c^\infty(\Omega)\,.$$

Then we obtain $f(x) \geq 0$ or $f(x) = 0$, respectively a.e. on Ω.

In what follows, we assume that $d\nu \in \mathfrak{Ndm}$ is a minimizer of the causal variational principle (5.4.8), (5.4.7). In analogy to [**12**, §3.3], we now focus on variations which *modify* the support of $d\nu$ (by contrast to (5.7.5)). More explicitly, for arbitrary $x_0 \in \operatorname{supp} d\zeta$ and $\varepsilon > 0$ we let $\eta_{x_0,\varepsilon} \in C_c^+(B_\varepsilon(x_0))$ such that

$$\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta = 1\,.$$

For arbitrary $g \in C_c^+(\hat{\mathscr{M}})$ we then consider variations of $d\nu$ in direction (cf. (5.7.6))

$$d\eta_g := -\left(\int_{\hat{\mathscr{M}}} g(k)\, d^4k\right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta + d\int_{\hat{\mathscr{M}}} g(k)\, A'(k)\, \delta_k(\cdot)\, d^4k \in \mathfrak{Odm}\,.$$

Then for sufficiently small $\delta > 0$ and all $\tau \in [0, \delta)$, the variations $d\tilde{\nu}_\tau := d\nu + \tau\, d\eta_g$ are negative definite measures which for every $\tau \in [0, \delta)$ are given by

$$d\tilde{\nu}_\tau := d\nu - \tau\left(\int_{\hat{\mathscr{M}}} g(k)\, d^4k\right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta + \tau\, d\int_{\hat{\mathscr{M}}} g(k)\, A'(k)\, \delta_k(\cdot)\, d^4k \in \mathfrak{Ndm}\,.$$

Concerning the side conditions (5.7.4), for any $g \in C_c(\hat{\mathscr{M}})$ we need to ensure that (5.7.3) is satisfied, i.e.

$$\mathrm{Tr}_V(\eta_g(\hat{\mathscr{M}})) = 0 \qquad \text{and} \qquad \mathrm{Tr}_V(-S\eta_g(\hat{\mathscr{M}})) = 0\,.$$

These conditions clearly are equivalent to

$$\mathrm{Tr}_V\left[-\left(\int_{\hat{\mathscr{M}}} g(k)\, A'(x_0)\, d^4k\right)\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta + \int_{\hat{\mathscr{M}}} g(k)\, A'(k)\, d^4k\right] = 0\,,$$

$$\mathrm{Tr}_V\left[S\int_{\hat{\mathscr{M}}} g(k)\, A'(x_0)\, d^4k\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta - S\int_{\hat{\mathscr{M}}} g(k)\, A'(k)\, d^4k\right] = 0\,.$$

Making use of $\mathrm{Tr}_V(A'(x_0)) = \mathrm{Tr}_V(A'(k))$ as well as $\mathrm{Tr}_V(-SA'(x_0)) = \mathrm{Tr}_V(-SA'(k))$ for all $k \in \hat{\mathscr{M}}$, it is easy to see that these conditions are satisfied for any $g \in C_c(\hat{\mathscr{M}})$, implying that variations in direction $d\eta_g$ are admissible for any $g \in C_c^+(\hat{\mathscr{M}})$ (in the sense of Definition 5.7.1).

LEMMA 5.7.9. *Let $d\nu \in \mathfrak{Ndm}$ be a minimizer of the causal variational principle* (5.4.8), (5.4.7). *Whenever $d\nu$ is strictly negative definite in the strong sense,*

$$\mathrm{Tr}_V\left(\hat{Q}(k)\, A'(k)\right) \geq c(A') \qquad \text{for all } k \in \hat{\mathscr{M}}\,,$$

where the constant $c(A')$ is given by Lemma 5.7.7.

PROOF. For arbitrary $g \in C_c^+(\hat{\mathscr{M}})$, $x_0 \in \operatorname{supp} d\nu$ and $\varepsilon > 0$ we introduce $d\eta_g$ by (5.7.6),

$$d\eta_g := -\left(\int_{\hat{\mathscr{M}}} g(k)\, d^4k\right) A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta + d\int_{\hat{\mathscr{M}}} g(k)\, A'(k)\, \delta_k(\cdot)\, d^4k\,,$$

where $\eta_{x_0,\varepsilon} \in C_c^+(B_\varepsilon(x_0))$ is a compactly supported function such that

$$\int_{B_\varepsilon(x_0)} \eta_{x_0,\varepsilon}\, d\zeta = 1\,.$$

Following the arguments in §5.7.2, for every $g \in C_c^+(\hat{\mathscr{M}})$ there is $\delta > 0$ such that

$$d\tilde{\nu}_\tau = d\nu + \tau\, d\eta_g \in \mathfrak{Ndm} \qquad \text{for all } \tau \in [0, \delta)\,.$$

Varying exclusively in the class of negative definite measures, by (5.5.13) and (5.7.1) we arrive at

$$0 \leq \delta^+ \mathcal{S}(\nu; \eta_g) = 4 \int_{\hat{\mathscr{M}}} g(k)\, \mathrm{Tr}_V \left[\hat{Q}(k)\, A'(k) - \int_{B_\varepsilon(x_0)} \hat{Q}(q)\, A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta(q) \right] d^4k\,.$$

Since $g \in C_c^+(\hat{\mathscr{M}}) \supset C_c^\infty(\hat{\mathscr{M}})^+$ is arbitrary and $\hat{\mathscr{M}} \ni k \mapsto \mathrm{Tr}_V\left(\hat{Q}(k)\, A'(k)\right)$ is a continuous mapping, the fundamental lemma of calculus of variations implies that

$$\mathrm{Tr}_V\left(\hat{Q}(k)\, A'(k)\right) \geq \mathrm{Tr}_V \left(\int_{B_\varepsilon(x_0)} \hat{Q}(q)\, A'(x_0)\, \eta_{x_0,\varepsilon}\, d\zeta(q) \right) \qquad \text{for all } k \in \hat{\mathscr{M}}\,.$$

Applying the mean value theorem for integrals [**87**, §27], in the limit $\varepsilon \searrow 0$ we get

$$\mathrm{Tr}_V\left(\hat{Q}(k)\, A'(k)\right) \geq \mathrm{Tr}_V\left(\hat{Q}(x_0)\, A'(x_0)\right) \qquad \text{for all } k \in \hat{\mathscr{M}}\,.$$

Moreover, according to Lemma 5.7.7, there is a constant $c(A')$ such that

$$\mathrm{Tr}_V\left(\hat{Q}(p)\, A'(p)\right) = c(A') \qquad \text{for all } p \in \operatorname{supp} d\nu\,.$$

Since $x_0 \in \operatorname{supp} d\nu$ is arbitrary, the previous two formulas give the claim. □

Combining Lemma 5.7.7 and Lemma 5.7.9, we obtain the main result of this chapter:

THEOREM 5.7.10. *Assume that $d\nu \in \mathfrak{Ndm}$ is a minimizer of the causal variational principle* (5.4.8) *with respect to the side conditions* (5.4.7). *Whenever Assumption 5.5.4 is satisfied and the minimizer $d\nu$ is a strictly negative definite measure in the strong sense (see Definition 5.7.3), the corresponding* ***Euler-Lagrange equations in momentum space*** *read*

$$\boxed{\hat{\mathfrak{q}}|_{\operatorname{supp} d\nu} \equiv \inf_{k \in \hat{\mathscr{M}}} \hat{\mathfrak{q}}(k)\,,} \tag{5.7.10}$$

where the mapping $\hat{\mathfrak{q}} : \hat{\mathscr{M}} \to \mathbb{R}$ is given by (5.7.2),

$$k \mapsto \hat{\mathfrak{q}}(k) = \mathrm{Tr}_V\left(\alpha(k)\, A'(k) + \beta(k)\, S\, A'(k)\right)\,,$$

and the mapping $A'(k)$ is supposed to depend continuously on $k \in \hat{\mathscr{M}}$.

PROOF. We essentially employ the fact that, according to Theorem 5.6.8,

$$\hat{Q}(k)\, A'(k) = \alpha(k)\, A'(k) + \beta(k)\, S A'(k) \qquad \text{for all } k \in \hat{\mathscr{M}}\,.$$

Introducing the mapping $\hat{\mathfrak{q}} : \hat{\mathscr{M}} \to \mathbb{R}$ by (5.7.2), i.e.

$$k \mapsto \hat{\mathfrak{q}}(k) = \mathrm{Tr}_V\left(\alpha(k)\, A'(k) + \beta(k)\, A'(k)\right) \qquad \text{for every } k \in \hat{\mathscr{M}}\,,$$

and making use of Lemma 5.7.7, for all $p \in \operatorname{supp} d\nu$ we obtain

$$\hat{\mathfrak{q}}(p) = \mathrm{Tr}_V\left(\hat{Q}(p)\, A'(p)\right) = c(A')$$

for some constant $c(A')$. Hence applying Lemma 5.7.9, for all $k \in \hat{\mathscr{M}}$ we are given

$$\hat{\mathfrak{q}}(k) = \mathrm{Tr}_V\left(\hat{Q}(k)\, A'(k)\right) \geq c(A') = \hat{\mathfrak{q}}|_{\operatorname{supp} d\nu}\,.$$

In other words, the EL equations take the form (5.7.10),

$$\hat{\mathfrak{q}}|_{\text{supp}\, d\nu} \equiv \inf_{k \in \hat{\mathscr{M}}} \hat{\mathfrak{q}}(k)\,.$$

This completes the proof. □

REMARK 5.7.11. *We refer to the functions α, β in $\hat{\mathfrak{q}}$ as "Lagrange parameters" because they take into account the side conditions* (5.6.2)*. We also point out the analogy between the EL equations in Theorem 5.7.10 and the equality in* [**53**, Definition 5.6.2 (iii)]*. Let us finally note that the EL equations in momentum space according to Theorem 5.7.10 are of a similar form as the EL equations in position space (see* [**66**, Theorem 4.2] *and* [**110**, Theorem 5.7] *or Theorem 3.4.2 and Theorem 4.6.6).*

5.8. Homogeneous Causal Fermion Systems

The objective of this last section is to clarify the connection to homogeneous causal fermion systems as considered in [**109**]. In [**109**, Section 3] it is shown that, starting from an indefinite inner product space $(V, \prec . \mid . \succ)$ together with a negative definite measure $d\nu$ on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$, one may construct a corresponding causal fermion system. For convenience, we shall briefly recall the main construction steps (§5.8.1). Afterwards we put the results obtained above into the physical context (§5.8.2). For clarity we mention that, following [**109**], a causal fermion system is called *homogeneous* whenever the kernel of the fermionic projector merely depends on the difference of two spacetime points.

5.8.1. Reconstruction of Causal Fermion Systems. In this subsection, we briefly outline the main result of [**109**, Chapter 3]. In short, starting with an indefinite inner product space $(V, \prec . \mid . \succ)$ of dimension $2n$ and signature (n, n) for some $n \in \mathbb{N}$ and a negative definite measure $d\nu$ on momentum space $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$, our goal is to construct a causal fermion system $(\mathcal{H}, \mathcal{F}, d\rho)$. To this end, we shall proceed in several steps.

To begin with, we introduce *wave functions* ψ as mappings from position space $\mathscr{M}$ to V. The set of all wave functions $\psi : \mathscr{M} \to V$ shall be denoted by $\mathcal{K}$. Given a basis of V as well as a norm $\|.\|$ on V, a wave function $\psi : \mathscr{M} \to V$ is said to be *continuous* if and only if $x \to y$ in $\mathscr{M}$ implies that $\|\psi(x) - \psi(y)\| \to 0$. (Since all norms on V are equivalent, this definition is independent of the norm.) Let us denote the vector space of continuous wave functions $\psi : \mathscr{M} \to V$ with compact support by $C_c(\mathscr{M}; V)$. Restricting attention to smooth wave functions in $C_c(\mathscr{M}; V)$ and applying a density argument gives rise to the Schwartz space $\mathcal{K}_0 := \mathcal{S}(\mathscr{M}; V)$.

Next, we introduce the kernel of the fermionic projector in analogy to §5.3.3 by

$$P(x,y) : V \to V\,, \qquad P(x,y) := \int_{\hat{\mathscr{M}}} e^{ik(y-x)}\, d\nu(k) \qquad \text{for all } x, y \in \mathscr{M}\,.$$

Accordingly, the fermionic projector is a well-defined mapping

$$P : \mathcal{K}_0 \to \mathcal{K}\,, \qquad (P\phi)(x) := \int_{\mathscr{M}} P(x,y)\, \phi(y)\, d\mu(x) \qquad \text{for all } x \in \mathscr{M}\,.$$

We observe that the fermionic projector P is symmetric with respect to the inner product $< . \mid . > : \mathcal{K}_0 \times \mathcal{K}_0 \to \mathbb{C}$, defined by

$$< \psi \mid \phi > := \int_{\mathscr{M}} \prec \psi(x) \mid \phi(x) \succ\, d\mu(x) \qquad \text{for all } \psi, \phi \in \mathcal{K}_0\,.$$

We then let $\mathcal{H}_0 := P(\mathcal{K}_0)$, endowed with the inner product

$$\langle P\psi \mid P\phi \rangle_{\mathcal{H}_0} := < \psi \mid (-P)\phi > .$$

Dividing out the null space $\mathcal{N}$ and setting $\mathcal{I} := \mathcal{H}_0/\mathcal{N}$, the restriction

$$\langle\, .\,|\,.\,\rangle_{\mathcal{I}} := \langle\, .\,|\,.\,\rangle_{\mathcal{H}_0}|_{\mathcal{I}\times\mathcal{I}}$$

defines a scalar product on $\mathcal{I}$. Considering the closure with respect to $\langle\, .\,|\,.\,\rangle_{\mathcal{I}}$,

$$\mathcal{H} := \overline{\mathcal{I}}^{\langle\, .\,|\,.\,\rangle_{\mathcal{I}}} ,$$

endowed with the induced scalar product $\langle\, .\,|\,.\,\rangle_{\mathcal{H}}$, gives rise to a Hilbert space $\mathcal{H}$.

In order to establish the connection to the setting of causal fermion systems, we proceed in analogy to [**59**, §1.2.2] by considering the mappings $F_{\mathcal{I}} : \mathscr{M} \to \mathrm{L}(\mathcal{I})$,

$$\langle \psi \mid F_{\mathcal{I}}(x)\, \phi \rangle_{\mathcal{H}} := \prec \psi(x) \mid \phi(x) \succ \qquad \text{for all } \psi, \phi \in \mathcal{I} .$$

The operators $F_{\mathcal{I}}(x)$ are linear, bounded and symmetric with respect to $\langle\, .\,|\,.\,\rangle_{\mathcal{H}}$ for every $x \in \mathscr{M}$. Accordingly, the Friedrichs extension [**157**, Satz VII.2.11] allows us to extend the operators $F_{\mathcal{I}}(x)$ canonically to self-adjoint operators $F(x) \in \mathrm{L}(\mathcal{H})$ with at most n positive and at most n negative eigenvalues for every $x \in \mathscr{M}$. Denoting by

$$\mathcal{F} \subset \mathrm{L}(\mathcal{H})$$

the set of self-adjoint (linear) operators on $\mathcal{H}$ with at most n positive and at most n negative eigenvalues, we thus obtain *local correlation operators* $F : \mathscr{M} \to \mathcal{F}$. Finally introducing the universal measure $d\rho$ as the push-forward measure

$$d\rho := F_*\, d\mu$$

gives rise to a causal fermion system $(\mathcal{H}, \mathcal{F}, d\rho)$.

5.8.2. Interpretation and Outlook. We finally are in the position to interpret the results obtained before and to put them into the physical context. Starting point for the development of causal fermion systems is the idea to take the concept of a Dirac sea literally (cf. §5.2.2). As a consequence, the basic object is the kernel of the fermionic projector (5.2.5),

$$P(x,y) = \int_{\hat{\mathscr{M}}} \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k^0)\, e^{-ik(x-y)}$$

for all $x, y \in \mathscr{M}$ (for details we refer to [**53**] and [**59**]). Assuming that $P(x,y)$ is homogeneous (see [**56**, Section 4] and Section 5.3), the kernel of the fermionic projector can be generalized to (5.3.4) in terms of negative definite measures. The construction outlined in §5.8.1 shows that each negative definite measure on $\hat{\mathscr{M}}$ with values in $\mathrm{L}(V)$ gives rise to a corresponding causal fermion system. Now, following the explanations in [**59**, Chapter 1], the translation to causal fermion systems amounts to introducing a "regularization." Furthermore, the objects appearing in causal fermion systems are considered as being regularized, thereby describing the unknown microstructure of spacetime. Thus for any negative definite measure $d\nu$, the fermionic projector (5.3.5) can be thought of as a regularization of (5.2.5). Following our basic strategy, by making use of the existence result [**56**, Theorem 4.2] together with the EL equations obtained in Theorem 5.7.10 it seems promising to construct a minimizer $d\nu$ in momentum space in the fashion of [**66**]. Reconstructing the corresponding causal fermion system according to §5.8.1, the fermionic projector (5.3.4) corresponding to the minimizer $d\nu$ may be interpreted as the optimal

regularization of (5.2.5) in the homogeneous setting. Moreover, the support of the corresponding universal measure $d\rho$ is considered as describing spacetime. The support of the measure $d\nu$, on the other hand, which is a subset of momentum space $\hat{\mathscr{M}}$, is expected to have the following properties: For small momenta compared to the Planck energy, the support of $d\nu$ should resemble the hyperboloid of the lower mass shell, whereas for large momenta it is supposed to have some different, yet unknown shape (for an illustration see [**53**, Figure 4.1]). In case that this imagination holds true, it is conceivable that the measure $d\nu$ comes with some natural cutoff in momentum space, thus giving a possible explanation why the renormalization program in relativistic quantum field theory yields rather accurate results. Further investigations shall be postponed to future projects, including a thorough analysis of the corresponding construction process of minimizers in momentum space.

APPENDIX A

Supplementary Results

A.1. Non-Triviality of the Constructed Measure

The following results show that the measure ρ given by (3.4.7) is non-zero in the case that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is either of compact range or is continuous and decays in entropy.

LEMMA A.1.1. *Assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is of compact range. Then the measure ρ obtained in* (3.4.7) *is non-zero. The total volume $\rho(\mathcal{F})$ is possibly infinite.*

PROOF. Let $(K_n)_{n\in\mathbb{N}}$ be the compact exhaustion of $\mathcal{F}$ with $K_1 \neq \varnothing$ and $K_n \subset K_{n+1}^\circ$ for all $n \in \mathbb{N}$. By construction of $\rho^{(n)}$ we are given $\operatorname{supp} \rho^{(n)} \subset K_n$ for every $n \in \mathbb{N}$. Assuming that $\mathcal{L}$ is of compact range, for every $x \in \mathcal{F}$ there is $K_x \subset \mathcal{F}$ compact such that $\mathcal{L}(x, y) = 0$ for all $y \notin K_x$, and $K_x \subset K_n^\circ$ for sufficiently large $n \in \mathbb{N}$. As a consequence,

$$1 \leq \int_{K_x} \mathcal{L}(x, y)\, d\rho^{(n)}(y) \leq \sup_{y\in K_x} \mathcal{L}(x, y)\, \rho^{(n)}(K_x)\,,$$

showing that $\rho^{(n)}(K_x) \geq c_x$ for some constant $c_x > 0$ for sufficiently large $n \in \mathbb{N}$. Moreover, choosing $f \in C_c(\mathcal{F}; [0, 1])$ with $f|_{K_x} \equiv 1$, applying vague convergence (3.4.9) yields

$$\int_{\mathcal{F}} f\, d\rho \overset{(3.4.9)}{=} \lim_{n\to\infty} \int_{\mathcal{F}} f\, d\rho^{(n)} \geq \lim_{n\to\infty} \rho^{(n)}(K_x) \geq c_x > 0\,.$$

We conclude that the measure is non-zero.

Whenever $(K_{x_j})_{j\in\mathbb{N}}$ is a disjoint sequence of such compact sets, we may choose open sets $U_j \supset K_{x_j}$ such that $U_i \cap U_j = \varnothing$ for all $i \neq j$. Choosing $f_j \in C_c(U_j; [0, 1])$ with $f_j|_{K_{x_j}} \equiv 1$ for every $j \in \mathbb{N}$, we obtain

$$\rho(\mathcal{F}) = \int_{\mathcal{F}} d\rho \geq \sum_{j\in\mathbb{N}} \int_{U_j} f_j\, d\rho = \sum_{j\in\mathbb{N}} \lim_{k\to\infty} \int_{U_j} f_j\, d\rho^{(k)} \geq \sum_{j\in\mathbb{N}} \lim_{k\to\infty} \rho^{(k)}(K_{x_j}) \geq \sum_{j\in\mathbb{N}} c_{x_j}\,.$$

Hence the total volume $\rho(\mathcal{F})$ is infinite if $\sum_{j\in\mathbb{N}} c_{x_j}$ diverges. □

LEMMA A.1.2. *Assume that the Lagrangian $\mathcal{L} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ is continuous and decays in entropy. Then the measure ρ obtained in* (3.4.7) *is non-zero, and the total volume $\rho(\mathcal{F})$ is possibly infinite.*

PROOF. For arbitrary $x \in \mathcal{F}$ and $0 < \varepsilon < 1$ let

$$\tilde{\mathcal{L}}_{x,\varepsilon} : \mathcal{F} \to \mathbb{R}_0^+\,, \qquad \tilde{\mathcal{L}}_{x,\varepsilon}(y) := \mathcal{L}(x, y)\, \chi_{K_{x,\varepsilon}}$$

with $K_{x,\varepsilon} \subset \mathcal{F}$ compact given by (3.5.6). In particular, $\tilde{\mathcal{L}}_{x,\varepsilon}(y) = 0$ for all $y \notin K_{x,\varepsilon}$. As a consequence, $\rho^{(n)}(K_{x,\varepsilon}) \geq c_{x,\varepsilon}$ for some constant $c_{x,\varepsilon} > 0$ for sufficiently large $n \in \mathbb{N}$,

as the following argument shows. Assuming conversely that $\rho^{(n)}(K_{x,\varepsilon}) \to 0$ as $n \to \infty$, in view of (3.5.6) we arrive at the contradiction

$$1 \leq \int_{\mathcal{F} \setminus K_{x,\varepsilon}} \mathcal{L}(x,y)\, d\rho^{(n)}(y) + \int_{K_{x,\varepsilon}} \tilde{\mathcal{L}}_{x,\varepsilon}(y)\, d\rho^{(n)}(y) \leq \varepsilon + \sup_{y \in K_{x,\varepsilon}} \mathcal{L}(x,y)\, \rho^{(n)}(K_{x,\varepsilon}) < 1$$

for sufficiently large $n \in \mathbb{N}$. Choosing $f \in C_c(\mathcal{F}; [0,1])$ with $f|_{K_{x,\varepsilon}} \equiv 1$, according to vague convergence (3.4.9) we obtain

$$\int_{\mathcal{F}} f\, d\rho \overset{(3.4.9)}{=} \lim_{n \to \infty} \int_{\mathcal{F}} f\, d\rho^{(n)} \geq \lim_{n \to \infty} \rho^{(n)}(K_{x,\varepsilon}) \geq c_{x,\varepsilon} > 0\,.$$

Thus the measure ρ is non-zero. The last assertion can be proven analogously to the proof of Lemma A.1.1. □

A.2. Topological Properties of Causal Fermion Systems

The goal of this appendix is to prove the following result:

THEOREM A.2.1. *Let $(\mathcal{H}, \mathcal{F}, \rho)$ be a causal fermion system. Then $\mathcal{F}$ is a Polish space.*[1]

Throughout this section we assume that $(\mathcal{H}, \mathcal{F}, \rho)$ is a causal fermion system of spin dimension $s \in \mathbb{N}$. More precisely, we consider a (possibly infinite-dimensional) separable complex Hilbert space $\mathcal{H}$ endowed with a scalar product $\langle\,.\,|\,.\,\rangle_{\mathcal{H}}$. Denoting the set of all bounded linear operators on $\mathcal{H}$ by $\mathrm{L}(\mathcal{H})$, we let $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ be the subset consisting of those operators $A \in \mathrm{L}(\mathcal{H})$ which are self-adjoint with respect to the scalar product $\langle\,.\,|\,.\,\rangle_{\mathcal{H}}$ on $\mathcal{H}$ and have at most s positive and at most s negative eigenvalues (see [**59**, §1.1.1]). The proof of Theorem A.2.1 is split up in two parts: We first point out that $\mathcal{F}$ is separable (§A.2.1). Afterwards, we prove that $\mathcal{F}$ is completely metrizable (§A.2.2). The result can be immediately generalized to the case of operators which have at most p positive and at most q negative eigenvalues.

A.2.1. Separability. Concerning separability of $\mathcal{F}$, we argue as follows: Since $\mathcal{H}$ is an infinite-dimensional, separable complex Hilbert space, the set of linear operators on $\mathcal{H}$, denoted by $\mathrm{L}(\mathcal{H})$, is a non-separable Banach space; the class of compact operators $\mathcal{K}(H) \subset \mathrm{L}(H)$, however, is a separable Banach space (cf. [**99**, §3.A and §12.E]). Since each $A \in \mathcal{F}$ has finite rank and thus is compact, we conclude that $\mathcal{F} \subset \mathcal{K}(\mathcal{H})$ (for details see [**119**, Chapter 15] or [**157**, Section II.3]). Since $\mathrm{L}(\mathcal{H})$ is metrizable by the Fréchet metric induced by the operator norm on $\mathrm{L}(\mathcal{H})$, the set $\mathcal{K}(\mathcal{H})$ is metrizable, and hence $\mathcal{F}$ is separable in view of [**4**, Corollary 3.5].

A.2.2. Completeness. The aim of this subsection is to show that $\mathcal{F}$ is completely metrizable with respect to the Fréchet metric induced by the operator norm on $\mathrm{L}(\mathcal{H})$. To this end, we proceed as follows. Given a sequence of operators $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{F}$, our task is to prove that its limit $A \in \mathcal{K}(\mathcal{H})$ is self-adjoint (with respect to the scalar product $\langle\,.\,|\,.\,\rangle_{\mathcal{H}}$ on $\mathcal{H}$) and has at most n positive and at most n negative eigenvalues.

[1] More precisely, endowed with the (Fréchet) metric d induced by the operator norm on $\mathrm{L}(\mathcal{H})$, the space $(\mathcal{F}, d)$ is a separable, complete metric space.

A.2.2.1. *Self-Adjointness.* We start by proving that A is self-adjoint in the case of a general Hilbert space H.

LEMMA A.2.2. *Let $(H, \langle . \,|\, . \rangle_H)$ be a Hilbert space, and let $(A_n)_{n \in \mathbb{N}}$ be a sequence of self-adjoint operators in $\mathrm{L}(H)$ converging in norm to some $A \in \mathrm{L}(H)$. Then A is self-adjoint.*

PROOF. For any $u, v \in H$, applying the Cauchy-Schwarz inequality and making use of the fact that A_n is self-adjoint for every $n \in \mathbb{N}$ yields

$$\begin{aligned} |\langle u \,|\, A^* \, v \rangle_H - \langle u \,|\, A \, v \rangle_H| &= |\langle A \, u \,|\, v \rangle_H - \langle u \,|\, A \, v \rangle_H| \\ &= |\langle A \, u \,|\, v \rangle_H - \langle A_n \, u \,|\, v \rangle_H + \langle A_n \, u \,|\, v \rangle_H - \langle u \,|\, A \, v \rangle_H| \\ &\leq 2 \, \|A - A_n\|_{\mathrm{L}(H)} \, \|u\| \, \|v\| \to_{n \to \infty} 0 \,. \end{aligned}$$

This completes the proof. □

A.2.2.2. *Operators, Resolvents and Spectra.* The remainder of this section is dedicated to the proof that the limit $A \in \mathrm{L}(\mathcal{H})$ of a sequence $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{F}$ (with respect to the operator norm) has at most s positive and at most s negative eigenvalues. To this end, we will essentially make use of results in [**97**], which we now briefly recall.

For Banach spaces X and Y, by $\mathcal{B}(X, Y)$ and $\mathcal{C}(X, Y)$ we denote the set of all bounded and closed operators from X to Y, respectively. Then $\mathcal{B}(X, Y)$ is a Banach space, and we let $\mathcal{B}(X) := \mathcal{B}(X, X)$ and $\mathcal{C}(X) := \mathcal{C}(X, X)$. We denote the domain of an operator T from X to Y by $D(T)$, and its graph $\mathsf{G}(T)$ is by definition the subset of $X \times Y$ consisting of all elements of the form (u, Tu) with $u \in D(T)$. Note that $\mathsf{G}(T)$ is a closed linear subspace of $X \times Y$ if and only if $T \in \mathcal{C}(X, Y)$ (see [**97**, III-§5.2]).

In what follows, let X, Y be complex Banach spaces, and let H be a complex Hilbert space. For $\zeta \in \mathbb{C}$ and $T \in \mathcal{C}(X)$, we introduce the operator

$$T_\zeta := T - \zeta \mathbb{1} \,.$$

Then the *resolvent set* $\rho(T)$ is defined to consist of all $\xi \in \mathbb{C}$ for which T_ζ has an inverse, denoted by

$$R(\zeta) = R(\zeta, T) := (T - \zeta)^{-1} \,.$$

We call $R(\zeta, T)$ the *resolvent* of T (see [**97**, III-§6] and [**134**, Definition 8.38]). The *spectrum* $\sigma(T)$ of T is given by the complementary set of the resolvent set in the complex plane, $\sigma(T) := \mathbb{C} \setminus \rho(T)$. Note that the spectrum of a compact operator T in a Banach space X has a simple structure analogous to that of an operator in a finite-dimensional space. Namely, for compact operators, each non-zero eigenvalue is of finite multiplicity:

THEOREM A.2.3. *Let $T \in \mathcal{B}(X)$ be compact. Then $\sigma(T)$ is a countable set with no accumulation point different from zero, and each nonzero $\lambda \in \sigma(T)$ is an eigenvalue of T with finite multiplicity.*

PROOF. See [**97**, Theorem III-6.26]. □

Moreover, the spectrum $\sigma(T)$ of a self-adjoint operator T in H is a subset of the real axis.

An isolated point of the spectrum is referred to as *isolated eigenvalue* [**97**, III-§6.5]. Concerning compact operators T on X, we may state the following remark.

REMARK A.2.4. *Let T be a compact operator. Then every complex number $\lambda \neq 0$ belongs to $\rho(T)$ or is an isolated eigenvalue with finite multiplicity. See* [**97**, Remark III-6.27].

A.2.2.3. *Projection and Decomposition.* Let X, Y be Banach spaces, and let $M \subset X$ be a linear subspace (or "manifold" in the terminology of [**97**]). As usual, an idempotent operator $P \in \mathcal{B}(X)$ ($P^2 = P$) is called a *projection*, giving rise to the decomposition

$$X = M \oplus N\,, \tag{A.2.1}$$

where $M = PX$ and $N = (1-P)X$ are *closed* linear subspaces of X which are referred to as *complementary* (see [**97**, III-§3.4]). Then each $x \in X$ can be uniquely expressed in the form $u = u' + u''$ with $u' \in M$ and $u'' \in N$. The vector u' is called the *projection of u on M along N*, and P is called the *projection operator* (or simply the *projection*) *on M along N*. Accordingly, the operator $1 - P$ is the projection on N along M. The range of P is M and the null space of P is N. For convenience we often write $\dim P$ for $\dim M = \dim \mathcal{R}(P)$, where $\mathcal{R}(P)$ denotes the range of P. Since $Pu \in M$ for every $u \in X$, we have $PPu = Pu$, implying that P is *idempotent*: $P^2 = P$.

Next, a linear subspace M is said to be *invariant* under an operator $T \in \mathcal{B}(X)$ if $TM \subset M$. In this case, T *induces* a linear operator T_M on M to M, defined by $T_M u = Tu$ for $u \in M$. the operator T_M is called the *part of T in M*. If there are two invariant linear subspaces M, N for T such that $X = M \oplus N$, the operator T is said to be *decomposed* (or *reduced*) by the pair M, N.

The notion of the *decomposition* of T by a pair M, N of complementary subspaces (see (A.2.1), [**97**, III-(3.14)]) can be extended in the following way. An operator T is said to be decomposed according to $X = M \oplus N$ if

$$PD(T) \subset D(T)\,, \qquad TM \subset M\,, \qquad TN \subset N\,, \tag{A.2.2}$$

where P is the projection on M along N. When T is decomposed as above, the *parts* T_M, T_N of T in M, N, respectively, can be defined. Then T_M is an operator in the Banach space M with $D(T_M) = D(T) \cap M$ such that $T_M u = Tu \in M$, and T_N is defined similarly.

A.2.2.4. *Generalized Convergence.* Let us briefly recall the definition of convergence in the generalized sense:

Definition A.2.5. *Let $T, T_n \in \mathcal{B}(X, Y)$ for all $n \in \mathbb{N}$.*

(i) *The convergence of $(T_n)_{n\in\mathbb{N}}$ to T in the sense of $\|T_n - T\| \to 0$ is called* ***uniform convergence*** *or* ***convergence in norm****.*

(ii) *Given closed operators $T, S \in \mathcal{C}(X, Y)$, their graphs $\mathsf{G}(T)$, $\mathsf{G}(S)$ are closed linear subspaces of the product space $X \times Y$. For two closed linear subspaces M, N of a Banach space Z we let $\delta(M, N)$ be the smallest number δ such that*

$$\operatorname{dist}(u, N) \le \delta\,\|u\| \qquad \textit{for all } u \in M.$$

We call $\hat{\delta}(T, S) := \delta(\mathsf{G}(T), \mathsf{G}(S))$ the ***gap*** *between T and S. If $\hat{\delta}(T_n, T) \to 0$, we shall also say that the operator T_n converges to T (or $T_n \to T$)* in the generalized sense.

See [**97**, Chapter III, §3.1] *and* [**97**, Chapter IV, §2.1].

The following theorem establishes a connection between convergence in the generalized sense and uniform convergence.

Theorem A.2.6. *Let $T, T_n \in \mathcal{C}(X, Y)$ for all $n \in \mathbb{N}$. If $T \in \mathcal{B}(X, Y)$, then $T_n \to T$ in the generalized sense if and only if $T_n \in \mathcal{B}(X, Y)$ for sufficiently large n and $\|T_n - T\| \to 0$.*

Proof. See [**97**, Theorem IV-2.23]. □

A.2.2.5. *Separation of the Spectrum.* Sometimes it happens that the spectrum $\sigma(T)$ of a closed operator T contains a bounded part σ' separated from the rest σ'' in such a way that a rectifiable, simple closed curve Γ (or, more generally, a finite number of such curves) can be drawn so as to enclose an open set containing σ' in its interior and σ'' in its exterior. Under such a circumstance, the following *decomposition theorem* holds.

THEOREM A.2.7. *Let $\sigma(T)$ be separated into two parts σ', σ'' in the way described above. Then we have a decomposition of T according to a decomposition $X = M' \oplus M''$ of the space (in the sense of* (A.2.2)*, cf.* [**97**, III-§5.6]*) in such a way that the spectra of the parts $T_{M'}$, $T_{M''}$ coincide with σ', σ'' respectively and $T_{M'} \in \mathcal{B}(M')$.*

PROOF. See [**97**, Theorem III-6.17]. □

The proof of Theorem A.2.7 makes use of the so-called *eigenprojection*

$$P[T] = -\frac{1}{2\pi i} \int_\Gamma R(\zeta, T)\, d\zeta \in \mathcal{B}(X)\,, \tag{A.2.3}$$

which is a projection on $M' = P[T]X$ along $M'' = (1 - P[T])X$.

REMARK A.2.8 (**Finite system of eigenvalues**). *Suppose that the spectrum $\sigma(T)$ of $T \in \mathcal{C}(X)$ has an* isolated point *λ. Obviously $\sigma(T)$ is divided into two separate parts σ' and σ'', where σ' consists of the single point λ; any closed curve enclosing λ but no other point of $\sigma(T)$ may be chosen as Γ. Then the spectrum of the operator $T_{M'}$ described in* [**97**, Theorem III-6.17] *(see Theorem A.2.7) consists of the single point λ.*

If M' is finite-dimensional, λ is an eigenvalue *of T. In fact, since λ belongs to the spectrum of the finite-dimensional operator $T_{M'}$, it must be an eigenvalue of $T_{M'}$ and hence of T. In this case,* $\dim M'$ *is called the* (algebraic) multiplicity *of the eigenvalue λ of T.*

For brevity, a finite collection $\lambda_1, \ldots, \lambda_s$ of eigenvalues with finite multiplicities will be called a finite system of eigenvalues*.*

A.2.2.6. *Continuity of the Spectrum.* This paragraph is devoted to the proof that the spectrum of a sequence of operators converging in the generalized sense behaves continuously.

THEOREM A.2.9. *Let $T \in \mathcal{C}(X)$ and let $\sigma(T)$ be separated into two parts $\sigma'(T)$, $\sigma''(T)$ by a closed curve Γ as in §A.2.2.5 (cf.* [**97**, III-§6.4]*). Let $X = M'(T) \oplus M''(T)$ be the associated decomposition of X. Then there exists $\delta > 0$, depending on T and Γ, with the following properties. Any $S \in \mathcal{C}(X)$ with $\hat{\delta}(S, T) < \delta$ has spectrum $\sigma(S)$ likewise separated by Γ into two parts $\sigma'(S)$, $\sigma''(S)$ (Γ itself running in $\rho(S)$). In the associated decomposition $X = M'(S) \oplus M''(S)$, $M'(S)$ and $M''(S)$ are isomorphic with $M'(T)$ and $M''(T)$, respectively. In particular,* $\dim M'(S) = \dim M'(T)$, $\dim M''(S) = \dim M''(T)$ *and both $\sigma'(S)$ and $\sigma''(S)$ are nonempty if this is true for T. The decomposition $X = M'(S) \oplus M''(S)$ is continuous in S in the sense that the projection $P[S]$ of X onto $M'(S)$ along $M''(S)$ tends to $P[T]$ in norm as $\hat{\delta}(S, T) \to 0$.*

PROOF. See [**97**, Theorem IV-3.16]. □

LEMMA A.2.10. *Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of compact operators in $\mathcal{K}(X)$, and suppose that $T_n \to T$ in norm for some operator $T \in \mathcal{K}(X)$. Moreover, let $\sigma'(T)$ be a finite system of m eigenvalues, separated from the rest $\sigma''(T)$ of $\sigma(T)$ by a closed curve Γ in the manner of* [**97**, III-§6.4]*. Then $\sigma(T_n)$ is separated by Γ into $\sigma'(T_n)$, $\sigma''(T_n)$ such that each $\sigma'(T_n)$ also consists of m eigenvalues of T_n, provided that n is sufficiently large.*

PROOF. Since T is a compact operator, Remark A.2.4 states that each non-zero eigenvalue of T is isolated (see [**97**, Remark III-6.27]). For this reason, each non-zero eigenvalue of T as well as all positive or all non-zero eigenvalues of T can be enclosed by a closed curve Γ running in $\rho(T)$ as described in §A.2.2.5 (cf. [**97**, III-§6.4]). We denote the set of eigenvalues lying within Γ by $\sigma'(T)$, and the set of eigenvalues without Γ by $\sigma''(T)$; thus Γ encloses a finite system of eigenvalues. By virtue of Theorem A.2.7, we have a decomposition of T according to $X = M' \oplus M''$ such that the spectra of the parts $T_{M'}$, $T_{M''}$ coincide with $\sigma'(T)$, $\sigma''(T)$. Since $T_n, T \in \mathcal{B}(X)$ for all $n \in \mathbb{N}$, Theorem A.2.6 ensures that $T_n \to T$ in the generalized sense, i.e. $\hat{\delta}(T_n, T) \to 0$ in the limit $n \to \infty$. Hence we can apply Theorem A.2.9, giving rise to some $N \in \mathbb{N}$ such that Γ separates the spectra $\sigma(T_n)$ into two parts $\sigma'(T_n)$, $\sigma''(T_n)$ for all $n \geq N$. Considering the corresponding decomposition $X = M'(T_n) \oplus M''(T_n)$ for all $n \geq N$, by virtue of Theorem A.2.9 there exist isomorphisms $M'(T_n) \simeq M'(T)$ for all $n \geq N$. In particular, we obtain $\dim M'(T_n) = \dim M'(T)$ for all $n \geq N$, and the decomposition $X = M'(T_n) \oplus M''(T_n)$ is continuous in n in the sense that the projection $P[T_n]$ (see (A.2.3)) of X onto $M'(T_n)$ along $M''(T_n)$ tends to $P[T]$ in norm as $\hat{\delta}(T_n, T) \to 0$ which holds true in view of Theorem A.2.6. A fortiori, the algebraic multiplicity of eigenvalues of T within Γ coincides with the algebraic multiplicity of eigenvalues of T_n within Γ for all $n \geq N$. □

Making use of continuity of the spectrum according to Lemma A.2.10, we may derive the following result.

LEMMA A.2.11. *Let H be a Hilbert space, and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of self-adjoint compact operators in $\mathcal{K}(H)$ such that each operator A_n has at most s positive and at most s negative eigenvalues. If $(A_n)_{n\in\mathbb{N}}$ converges in norm to some $A \in \mathrm{L}(H)$, then A is also a self-adjoint compact operator which has at most s positive and at most s negative eigenvalues.*

PROOF. Given a sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{K}(H)$ with $A_n \to A$ in $\mathrm{L}(H)$, the fact that $\mathcal{K}(H)$ is a closed subspace of $\mathrm{L}(H)$ implies that the limit A is a compact operator. Thus in view of Theorem A.2.3 and Remark A.2.4, each non-zero eigenvalue of A is isolated and has finite multiplicity, and according to Remark A.2.8 the non-zero eigenvalues of A form a finite system of isolated eigenvalues. In particular, there is a closed curve Γ as described in Paragraph A.2.2.5 (cf. [**97**, III-§6.4]) which encloses all positive (negative) eigenvalues of A. Now assume that A has $m > s$ positive (negative) eigenvalues. Then Lemma A.2.10 yields the existence of some $N \in \mathbb{N}$ such that the spectrum $\sigma(A_n)$ is separated by Γ into two parts $\sigma'(A_n)$ within Γ and $\sigma''(A_n)$ without Γ for all $n \geq N$. As a consequence, $\sigma'(A_n)$ consists of $m > s$ positive (negative) eigenvalues for all $n \geq N$ in contradiction to the fact that A_n has at most s positive and at most s negative eigenvalues for all $n \in \mathbb{N}$. Hence $A \in \mathcal{K}(H)$ is a self-adjoint operator which has at most s positive and at most s negative eigenvalues. This concludes the proof. □

A.2.2.7. *Application to Causal Fermion Systems.* After these preparations, we are finally in the position to prove Theorem A.2.1.

PROOF OF THEOREM A.2.1. Let $(\mathcal{H}, \mathcal{F}, \rho)$ be a causal fermion system. Separability of $\mathcal{F}$ follows from §A.2.1. By virtue of Lemma A.2.11, we conclude that $\mathcal{F} \subset \mathrm{L}(\mathcal{H})$ is closed. Since $\mathrm{L}(\mathcal{H})$ is a complete metric space with respect to the Fréchet metric induced by the operator norm, we conclude that $\mathcal{F}$ is completely metrizable. Taken together, $\mathcal{F}$ is a separable, completely metrizable space, and thus Polish [**99**, Definition (3.1)]. □

More precisely, the space $(\mathcal{F}, d)$ is a complete metric space, where d is the Fréchet metric induced by the operator norm on $\mathrm{L}(\mathcal{H})$ (cf. [**5**] or [**21**]).

A.3. Support of Locally Finite Measures on Polish Spaces

In this section we derive useful topological properties concerning the support of locally finite measures (or Borel measures in the sense of [**76**]) on Polish spaces (see Lemma A.3.2 below). To begin with, let us recall the following preparatory result.

PROPOSITION A.3.1. *Let X be Polish and μ a finite measure on $\mathcal{B}(X)$. Then $A \subset X$ is μ-measurable if and only if there exists a σ-compact set $F \subset A$ with $\mu(A \setminus F) = 0$.*

PROOF. See [**99**, Theorem (17.11)]. □

Moreover, based on [**19**, Chapter IX], a Borel measure (in the sense of [**76**]) on a topological Hausdorff space X is said to be *moderated* if X is the union of countably many open subsets of finite μ-measure (see [**40**, Chapter VIII]). (Since open sets are measurable, every moderated measure is σ-finite.) We point out that, due to Ulam's theorem [**40**, Satz VIII.1.16], every Borel measure on a Polish space is regular and moderated. (Due to Meyer's theorem, the same is true for Borel measures on Souslin spaces, see [**40**, Satz VIII.1.17].) As a consequence, we may derive useful properties of the support of Borel measures on Polish spaces, as the following lemma shows.

LEMMA A.3.2. *Let X be a Polish space, and assume that μ is a Borel measure on $\mathcal{B}(X)$. Then* $\operatorname{supp}\mu \subset X$ *is a σ-compact topological space. Moreover, there exists a dense subset $F \subset \operatorname{supp}\mu$ such that each $x \in F$ has a compact neighborhood in* $\operatorname{supp}\mu$*. Whenever* $\operatorname{supp}\mu$ *is hemicompact, one can arrange that* $\operatorname{supp}\mu$ *is a Polish space which has the Heine-Borel property. The corresponding Heine-Borel metric can be chosen locally identical to a complete metric on X (in the relative topology of* $\operatorname{supp}\mu$*).*

PROOF. We make essentially use of the fact that the measure μ is moderated in view of Ulam's theorem. As a consequence, there is a sequence of sets $(U_n)_{n\in\mathbb{N}}$ with $U_n \subset X$ open and $\mu(U_n) < \infty$ for all $n \in \mathbb{N}$ such that $\bigcup_{n\in\mathbb{N}} U_n = X$. From the fact that open subsets of Polish spaces are Polish (in the relative topology, see [**9**, §26] or [**99**, Theorem (3.11)]), we conclude that each set $U_n \subset X$ is Polish. Thus each $\mu|_{U_n}$ is a finite Borel measure on a Polish space. Due to [**16**, Proposition 7.2.9], every (finite) Borel measure on a separable metric space has support, implying that

$$\mu|_{U_n}(U_n \setminus \operatorname{supp}\mu|_{U_n}) = 0\,.$$

In view of Proposition A.3.1, we conclude that $\operatorname{supp}\mu|_{U_n}$ is contained in a σ-compact set $F_n \subset U_n$, and thus $\operatorname{supp}\mu|_{U_n}$ is σ-compact for every $n \in \mathbb{N}$. From this we deduce that $\bigcup_{n\in\mathbb{N}} \operatorname{supp}\mu|_{U_n}$ is σ-compact. Making use of the fact that subsets of σ-compact spaces are σ-compact, we conclude that

$$\operatorname{supp}\mu \subset \bigcup_{n\in\mathbb{N}} \operatorname{supp}\mu|_{U_n} \qquad \text{is } \sigma\text{-compact}$$

(where μ has support in view of [**40**, Lemma VIII.2.15]). Since $\operatorname{supp}\mu \subset X$ is closed (see [**40**, §VIII.2.5]), the support $\operatorname{supp}\mu$ is Polish and thus Baire (cf. §4.3.2). From [**161**, §25B] we conclude that there exists a dense subset $F \subset \operatorname{supp}\mu$ such that each $x \in F$ has a compact neighborhood in $\operatorname{supp}\mu$.

Assuming that $\operatorname{supp}\mu$ is hemicompact (see for instance [**161**, §17I]), then $\operatorname{supp}\mu$ is locally compact in view of [**41**, Exercise 3.4.E] and thus a complete σ-locally compact

space (in the sense of [**150**]). In this case, due to [**162**, Theorem 2'] and the explanations in [**66**, Section 3], the space $\operatorname{supp}\mu$ is metrizable by a Heine-Borel metric which is (Cauchy) locally identical to a complete metric on X; endowed with such a metric, the space $\operatorname{supp}\mu$ has the Heine-Borel property, i.e. each closed bounded subset (with respect to the Heine-Borel metric) is compact. □

A.4. Justifying the Side Conditions

This appendix is devoted to justify and explain the side conditions (5.3.9)–(5.3.11). Apart from excluding trivial minimizers in a quite simple way, the following reasoning provides a strong argument for imposing condition (5.3.9). Given a causal fermion system $(\mathcal{H}, \mathcal{F}, d\rho)$, the so-called local trace $\mathrm{tr}(x)$ defined by

$$\mathrm{tr}(x) = \mathrm{Tr}_{S_x}\big(P(x,x)\big) \qquad \text{for all } x \in \operatorname{supp} d\rho$$

is constant on $\operatorname{supp} d\rho$ whenever the measure $d\rho$ is a minimizer of the causal action principle (for details see [**59**, §1.1.3, Proposition 1.4.1 and Section 2.5]). Considering homogeneous causal fermion systems, this suggests to impose that

$$\mathrm{Tr}_V(\nu(\hat{\mathscr{M}})) = \mathrm{Tr}_V\left(\int_{\hat{\mathscr{M}}} d\nu(k)\right) = \mathrm{Tr}_V(P(0)) = \mathrm{Tr}_V(P(x,x)) = c \qquad \text{for all } x \in \mathscr{M}\,,$$

thus motivating the side condition (5.3.9). Following the arguments in [**59**, §1.4.1], we shall always assume that $c \neq 0$, thereby excluding trivial minimizers. Let us briefly explain why the quantity $\mathrm{Tr}_V(P(0)) = \mathrm{Tr}_V(\nu(\hat{\mathscr{M}}))$ in (5.3.9) is also referred to as *mass density*.[2] In order to see that $\mathrm{Tr}_V(P(0))$ can indeed be regarded as a density, let us assume that $(\mathcal{H}, \mathcal{F}, d\rho)$ is a causal fermion system. Whenever $P^\varepsilon(x,y)$ is a regularization of the kernel of the fermionic projector of the vacuum $P(x,y)$ with regularization length ε (where $P(x,y)$ coincides with (5.2.5), cf. [**59**, eq. (1.2.23)]), its trace is given by (see [**59**, eq. (2.5.1)])

$$\mathrm{Tr}_{S_x}\big(P^\varepsilon(x,x)\big) \sim \frac{m}{\varepsilon^2} \qquad \text{for all } x \in \operatorname{supp} d\rho\,.$$

Making use of the fact that the unit of mass equals one over length, we conclude that the quantity $\mathrm{Tr}_{S_x}\big(P^\varepsilon(x,x)\big)$ is a density, which apparently is proportional to the mass m. Carrying these observations over to $\mathrm{Tr}_V(P(0))$ in the homogeneous case justifies the terminology of "mass density."

A possible explanation for introducing the constraint (5.3.11) is that a similar side condition for the closed chain is imposed in the existence theorem [**54**, Theorem 6.1]. Since the fermionic projector $P(0) = \nu(\hat{\mathscr{M}})$ can be diagonalized (up to an arbitrarily small error term) according to [**56**, Lemma 4.4], in order to develop the existence theory of minimizers in the homogeneous setting it seems promising to demand that constraint (5.3.11) is satisfied. On the other hand, following the original ideas in [**53**] and its modifications in [**56**], it is natural to impose a boundedness constraint (5.4.5). The arguments in Section 5.4 show that (5.4.5) already implies condition (5.3.11).

Let us finally discuss the remaining side condition (5.3.10). Since working with the spectral weight as appearing in the constraint (5.3.11) may be awkward, it might seem preferable to work with a similar condition which is more easy to handle. Bearing in mind that the operator $\nu(\hat{\mathscr{M}})$ may be diagonalized (up to an arbitrarily small error term) in virtue of [**56**, Lemma 4.4] in such a way that its diagonal entries are ordered according

[2]Note that the quantity $\mathrm{Tr}_V(P(0)) = \mathrm{Tr}_V(\nu(\hat{\mathscr{M}}))$ coincides with the *local particle density* f_{loc} as introduced in [**56**, eq. (4.4)]. In order to avoid confusion, this notion will not be used in what follows.

to [**56**, eq. (2.6)], the specific form of the signature matrix S (see (5.2.2)) suggests to replace condition (5.3.11) by (5.3.10),

$$\mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}})) = f\,.$$

The same arguments as before illustrate that $\mathrm{Tr}_V(-S\nu(\hat{\mathscr{M}}))$ is a density; we refer to this quantity as *particle density*.

APPENDIX B

Selected Mathematical Definitions and Results

B.1. Topology

DEFINITION B.1.1. *Let X be a set, let $\mathcal{P}(X)$ be the power set of X and let $\mathcal{O} \subset \mathcal{P}(X)$. Then $(X, \mathcal{O})$ is called* ***topological space*** *if the following conditions hold:*

(a) $\varnothing \in \mathcal{O}$ *and* $X \in \mathcal{O}$.
(b) *If* $U_i \in \mathcal{O}$ *for all* $i \in I$ *for any index set* I, *then* $\bigcup_{i\in I} U_i \in \mathcal{O}$.
(c) *If* $U_i \in \mathcal{O}$ *for all* $i \in I$ *for every finite index set* I, *then* $\bigcap_{i\in I} U_i \in \mathcal{O}$.

If the conditions (a)–(c) *hold, the set $\mathcal{O}$ is called the* ***topology*** *of X. The elements of X are called* ***points****, and the elements of $\mathcal{O}$ are referred to as* ***open sets*** *in X. Moreover, a set $A \subset X$ is* ***closed*** *if its complement $A^c = X \setminus A$ is open in X. For any $Y \subset X$, the open sets in the* ***relative topology*** *of Y are defined by $Y \cap U$ with $U \in \mathcal{O}$. See* [**41, 161**].

If the topology $\mathcal{O}$ of a topological space $(X, \mathcal{O})$ is understood, we often refer to X as a "topological space."

DEFINITION B.1.2. *Let $(X, \mathcal{O})$ be a topological space. A family $\mathcal{B} \subset \mathcal{O}$ is called a* ***base*** *for $(X, \mathcal{O})$ if every non-empty subset of X can be represented as the union of a subfamily of $\mathcal{B}$. A topological space $(X, \mathcal{O})$ is called* ***second-countable*** *if $(X, \mathcal{O})$ has a countable base. If for some $x \in X$ and some open $U \subset X$ we have $x \in U$, we say that U is a* ***neighborhood*** *of x. A topological space X is called* ***Hausdorff*** *if for every $x, y \in X$ with $x \neq y$ there exist open sets $U, V \subset X$ such that $x \in U$, $y \in V$ and $U \cap V = \varnothing$. See* [**41**].

DEFINITION B.1.3. *A topological space X is called* ***compact*** *if X is a Hausdorff space and every open cover of X has a finite subcover. A topological space X is referred to as* ***locally compact*** *if every point of X has a neighborhood whose closure is compact. A topological space X is said to be* ***σ-compact*** *if X can be written as the union of countably many compact subsets. Moreover, the space X is called* ***hemicompact*** *if there is a sequence $(K_n)_{n\in\mathbb{N}}$ of compact subsets of X such that if K is any compact subsets of X, then $K \subset K_n$ for some $n \in \mathbb{N}$. The* ***closure*** *of a topological space A is defined as the intersection of all closed sets containing A. A topological space A is called* ***relatively compact*** *if its closure $\overline{A}$ is compact. The* ***interior*** *of $A \subset X$ is defined as the union of all open subsets of A. See for instance* [**40, 41, 161**].

DEFINITION B.1.4. *A* ***metric space*** *is a pair (X, d) consisting of a set X and a metric $d : X \times X \to \mathbb{R}$. A* ***Cauchy sequence*** *in a metric space (X, d) is a sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \in X$ for every $n \in \mathbb{N}$ such that $d(x_m, x_n) \to 0$ as $m, n \to \infty$. A metric space (X, d) is called* ***complete*** *if every Cauchy sequence has a limit in X. A topological space X is* ***metrizable*** *if there exists a metric d on X such that the topology induced by the metric d coincides with the original topology on X. A topological space X is* ***completely metrizable*** *if it admits a compatible metric d such that (X, d) is complete. A separable completely metrizable space is called* ***Polish****. See* [**99, 41**].

B.2. Measure Theory

Following [**76**] and [**143**], Borel and Radon measures on topological Hausdorff spaces are defined in [**40**] as follows (see [**40**, Definition VIII.1.1 and Definition VIII.1.7]):

DEFINITION B.2.1 (**Borel measure**). *Let X be a topological Hausdorff space, and denote the Borel σ-algebra on X by $\mathcal{B} = \mathcal{B}(X)$. Moreover, let $\mathcal{A} \supset \mathcal{B}$ be a σ-algebra and let $\mu : \mathcal{A} \to [0, \infty]$ be a measure.*

(a) *The measure μ is said to be **locally finite** if for every $x \in X$ there is an open neighborhood U of x such that $\mu(U) < \infty$. A locally finite measure $\mu : \mathcal{B} \to [0, \infty]$ is referred to as **Borel measure**.*

(b) *A set $A \in \mathcal{A}$ is called **inner regular** if and only if*

$$\mu(A) = \sup \{\mu(K) : K \subset A \text{ compact}\} .$$

The measure μ is called inner regular *if each $A \in \mathcal{A}$ is inner regular.*

(c) *A set $A \in \mathcal{A}$ is called **outer regular** if and only if*

$$\mu(A) = \inf \{\mu(U) : U \supset A \text{ open}\} .$$

The measure μ is called outer regular *if each $A \in \mathcal{A}$ is outer regular.*

(d) *A set $A \in \mathcal{A}$ is called **regular** if and only if it is inner and outer regular. The measure μ is called* regular *if each $A \in \mathcal{A}$ is regular.*

(e) *An inner regular Borel measure is referred to as **Radon measure**.*

(f) *A Borel measure μ is said to be **moderated** if X is the union of a sequence of* open *sets of finite μ-measure.*

THEOREM B.2.2 (**Ulam's theorem**). *Any Borel measure on a Polish space is regular and moderated. See* [**40**, Satz VIII.1.16].

In what follows, by $C_c(X)$ we denote the set of continuous functions on a topological space X with compact support.

THEOREM B.2.3 (**Riesz representation theorem**). *Let X be a second-countable locally compact Hausdorff space and let $I : C_c(X) \to \mathbb{K}$ be linear and positive in the sense that $I(f) \geq 0$ whenever $f \geq 0$ (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$). Then there is a uniquely determined Borel measure $\mu : \mathcal{B}(X) \to [0, \infty]$ such that (see* [**9**, Corollar 29.13]*)*

$$I(f) = \int_X f \, d\mu \qquad \text{for all } C_c(X) .$$

DEFINITION B.2.4. *A sequence $(\mu_n)_{n\in\mathbb{N}}$ of Radon measures on a topological space X is said to converge **vaguely** to a Radon measure μ if (see e.g.* [**9**, Definition 30.1]*)*

$$\lim_{n\to\infty} \int_X f \, d\mu_n = \int_X f \, d\mu \qquad \text{for all } f \in C_c(X) .$$

In what follows, by $C_b(X)$ we denote the set of bounded continuous functions on a topological space X. Moreover, by $\mathbf{M}^+(X)$ we denote the set of positive finite measures on $\mathcal{B}(X)$.

DEFINITION B.2.5. *A sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures in $\mathbf{M}^+(X)$ is said to converge **weakly** to a measure $\mu \in \mathbf{M}^+(X)$ if (see e.g.* [**9**, Definition 30.7]*)*

$$\lim_{n\to\infty} \int_X f \, d\mu_n = \int_X f \, d\mu \qquad \text{for all } f \in C_b(X) .$$

DEFINITION B.2.6. *Let (X,d) be a metric space. A set $\mathcal{M} \subset \mathbf{M}^+(X)$ is called **relatively compact** if each sequence in $\mathcal{M}$ has a weakly convergent subsequence. A set $\mathcal{M} \subset \mathbf{M}^+(X)$ is called **bounded** if $\{\|\mu\| : \mu \in \mathcal{M}\}$ is bounded, where $\|\mu\|$ is the total variation of a measure μ. Moreover, a set $\mathcal{M} \subset \mathbf{M}^+(X)$ is called **uniformly tight** if for every $\varepsilon > 0$ there exists a compact set $K \subset X$ such that $\mu(K^c) < \varepsilon$ for all $\mu \in \mathcal{M}$. See for instance* [**40, 16**].

THEOREM B.2.7 (**Prohorov's theorem**). *If X is a Polish space, then $\mathcal{M} \subset \mathbf{M}^+(X)$ is relatively compact if and only if $\mathcal{M}$ is uniformly tight and bounded. See e.g.* [**40, 16**].

B.3. Functional Analysis

DEFINITION B.3.1. *A **normed space** is a $\mathbb{K}$-vector space X, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, endowed with a norm $\|\cdot\|$. A **Banach space** is a normed vector space that is complete with respect to the metric induced by the norm. The (topological) dual space X^* is the set of bounded linear functionals on X with values in $\mathbb{K}$. The **weak* topology** is the simply the topology of pointwise convergence, i.e. $f_\alpha \to f$ in X^* if and only if $f_\alpha(x) \to f(x)$ for all $x \in X$. For details see* [**72**].

THEOREM B.3.2 (**Banach-Alaoglu**). *If X is a normed vector space, the closed unit ball $\{f \in X^* : \|f\| \leq 1\}$ in X^* is compact in the weak* topology. See* [**72**, Theorem 5.18].

THEOREM B.3.3 (**Fréchet-Riesz**). *Let $\ell : H \to \mathbb{C}$ be a bounded linear functional on a Hilbert space $(H, \langle .,. \rangle)$ such that $|\ell(x)| \leq C\|x\|$ for some constant $C > 0$. Then ℓ is of the form*

$$\ell(x) = \langle x, y \rangle$$

for some uniquely determined $y \in H$. See [**113**, Section 6.3].

DEFINITION B.3.4. *Let $U \subset \mathbb{R}^m$ be a nonempty open set, and let $C_c^\infty(U)$ be the set of compactly supported smooth functions on U. The set of all test functions on U is denoted by $\mathcal{D}(U) = C_c^\infty(U)$. Then a **distribution** on U is a continuous linear functional on $C_c^\infty(U)$. The set of all distributions on U is denoted by $\mathcal{D}'(U)$; it is endowed with the weak* topology, that is, the topology of pointwise convergence on $C_c^\infty(U)$. A measurable function $f : \mathbb{R}^m \to \mathbb{C}$ is called **locally integrable** (with respect to Lebesgue measure $d\mu$) if $\int_K |f| \, d\mu < \infty$ for every bounded measurable set $K \subset \mathbb{R}^m$. The set of locally integrable functions $L^1_{loc}(U)$ can be regarded as a subspace of $\mathcal{D}'(U)$; its elements are referred to as **regular distributions**. Moreover, a **bi-distribution** is a continuous linear functional on $C_c^\infty(U) \times C_c^\infty(U)$. In the case $U = \mathbb{R}^m$, the space $\mathcal{D}(\mathbb{R}^m)$ is dense in the space of complex-valued Schwartz functions $\mathcal{S}(\mathbb{R}^m)$. A **tempered distribution** is a continuous linear mapping from $\mathcal{S}(\mathbb{R}^m)$ to $\mathbb{C}$. Similarly, a **tempered bi-distribution** is a continuous linear mapping from $\mathcal{S}(\mathbb{R}^m) \times \mathcal{S}(\mathbb{R}^m)$ to $\mathbb{C}$. See* [**72, 134, 140**].

B.4. Differential Geometry

DEFINITION B.4.1 (**Banach manifold**). *Let M be a topological space. A **chart** (U, φ) in M is a pair where U is open in M and $\varphi : U \to U_\varphi$ is a homeomorphism onto an open subset U_φ of a Banach space X over $\mathbb{K}$ (with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The Banach space X_φ is called **chart space** and φ is called a **chart map**. Given $k \in \mathbb{N}_0 \cup \{\infty\}$, two charts (U, φ) and (V, ψ) are called C^k-**compatible** if and only if $U \cap V = \varnothing$ or $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ are C^k. Moreover, a C^k-**atlas** for M is a collection of charts $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ for some index set A which satisfies the following conditions:*

(a) *The sets* $(U_\alpha)_{\alpha\in A}$ *cover* M.
(b) *Any two charts are* C^k*-compatible.*
(c) *All chart spaces* X_α *are Banach spaces over* $\mathbb{K}$.

The topological space M *is said to be a* C^k***-Banach manifold*** *if and only if there exists a* C^k*-atlas for* M. *See* [**169**, Section 73.2].

Definition B.4.2. *Given an integer* $m \geq 1$, *a* ***locally Euclidean space*** M ***of dimension*** m *is a Hausdorff topological space* M *for which each point has a neighborhood homeomorphic to an open subset of Euclidean space* $\mathbb{R}^m$ *An* m***-dimensional differentiable manifold of class*** C^k *is an* m*-dimensional, second-countable, locally Euclidean space* M *together with a* C^k*-atlas* $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$ *for some index set* A *such that* $\varphi_\alpha : U_\alpha \to \mathbb{R}^m$ *for all* $\alpha \in A$. *See* [**155**].

Definition B.4.3 (**vector field**). *Let* M *be a manifold, and denote by* $\mathfrak{F}(M)$ *the set of all smooth real-valued functions on* M. *Whenever* $p \in M$, *a* ***tangent vector to*** M ***at*** p *is a real-valued function* $v : \mathfrak{F}(M) \to \mathbb{R}$ *such that*

(a) $v(af + bg) = av(f) + bv(g)$ *and*
(b) $v(fg) = v(f)g(p) + f(p)v(g)$ *for all* $a, b \in \mathbb{R}$ *and* $f, g \in \mathfrak{F}(M)$.

The set T_pM *of all tangent vectors to* M *at* p *is called the* ***tangent space to*** M ***at*** p. *A* ***vector field*** V *on a manifold* M *is a function that assigns to each point* $p \in M$ *a tangent vector* V_p *to* M *at* p. *If* V *is a vector field on* M *and* $f \in \mathfrak{F}(M)$ *then* Vf *denotes the real-valued function on* M *given by*

$$(Vf)(p) = V_p(f) \qquad \text{for all } p \in M\,.$$

Then V *is* ***smooth*** *provided that* Vf *is smooth for all* $f \in \mathfrak{F}(M)$. *See* [**125**].

Definition B.4.4 (**submersion**). *Let* X, Y *be Banach spaces over* $\mathbb{K}$ *with* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. *A mapping* $G : D(G) \subset X \to Y$ *is called a* ***submersion*** *at* u_0 *if and only if the following conditions hold (see* [**168**, Definition 43.15]*):*

(a) G *is continuously Fréchet differentiable in a neighborhood of* u_0.
(b) $G'(u_0) : X \to Y$ *is surjective, i.e.* $\mathcal{R}(G'(u_0)) = Y$.
(c) *The null space* $\mathcal{N}(G'(u_0))$ *splits* X, *that is, there exists a continuous projection operator* P *of* X *on* $\mathcal{N}(G'(u_0))$ *such that*

$$X = \mathcal{N}(G'(u_0)) \oplus (\mathbb{1} - P)(X)\,.$$

Acknowledgments

At this point I finally would like to express my gratefulness to all those whose contributions led to the success of this thesis.

First and foremost, I would like to thank Felix Finster for supervising the present work and for patiently answering my questions concerning his theory. I am grateful for numerous discussions on miscellaneous subjects and a plenitude of enriching insights.

Next, I want to thank Marco Oppio for many enlightening discussions on physics, the theory of causal fermion systems and a multitude of other topics.

Moreover, I am thankful for various discussions with Mustafa Akman, Maximilian Jokel, Magdalena Lottner, Andreas Platzer, Christian Röken, Johannes Wurm, Saeed Zafari as well as Philipp Zimmermann.

Furthermore, I highly appreciate the competent work of Eva Rütz, the secretary of the chair. In addition, I am obliged to the "Studienstiftung des deutschen Volkes" for providing financial support and for granting multiple opportunities for instructive and valuable stays abroad. I also acknowledge funding by the DFG Graduate School 1693.

Last but definitely not least I would like to thank my parents for all their support.

Bibliography

1. *Link to web platform on causal fermion systems: www.causal-fermion-system.com*
2. I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics: A Practical Introduction: From Relativistic Quantum Mechanics to QED*, fourth ed., vol. 1, CRC Press, 2012.
3. ———, *Gauge Theories in Particle Physics: A practical introduction, Volume 2: Non-Abelian Gauge theories: QCD and the electroweak theory*, fourth ed., vol. 2, CRC Press, 2012.
4. C.D. Aliprantis and K.C. Border, *Infinite Dimensional Analysis: A hitchhiker's guide*, third ed., Springer, Berlin, 2006.
5. H.W. Alt, *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*, Fünfte, überarbeitete Auflage, Springer-Verlag, 2006.
6. H. Amann and J. Escher, *Analysis III*, Grundstudium Mathematik, Birkhäuser Verlag, Basel, 2001.
7. A.V. Arkhangel'skii and V.V. Fedorchuk, *General Topology I: Basic Concepts and Constructions Dimension Theory*, vol. 17, Springer Science & Business Media, 2012.
8. C. Bär and N. Ginoux and F. Pfäffle, *Wave Equations on Lorentzian Manifolds and Quantization*, vol. 3, European Mathematical Society, 2007.
9. H. Bauer, *Measure and Integration Theory*, De Gruyter Studies in Mathematics, vol. 26, Walter de Gruyter & Co., Berlin, 2001, Translated from the German by Robert B. Burckel.
10. J.K. Beem and P.E. Ehrlich and K.L. Easley, *Global Lorentzian Geometry*, second ed., Monographs and Textbooks in Pure and Applied Mathematics, vol. 202, Marcel Dekker, Inc., New York, 1996.
11. A. Beiser, *Concepts of Modern Physics*, sixth ed., McGraw-Hill, Boston, 2003.
12. Y. Bernard and F. Finster, *On the structure of minimizers of causal variational principles in the non-compact and equivariant settings*, Advances in Calculus of Variations, 7, 1, 27–57, 2014, De Gruyter.
13. P. Billingsley, *Convergence of Probability Measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.
14. J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill Book Co., New York-Toronto-London, 1964.
15. ———, *Relativistic Quantum Fields*, McGraw-Hill Book Co., New York-Toronto-London-Sydney, 1965.
16. V.I. Bogachev, *Measure Theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.
17. J. Bognár, J., *Indefinite Inner Product Spaces*, Springer-Verlag, New York-Heidelberg, 1974, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78.
18. M. Born and W. Biem, *Die Relativitätstheorie Einsteins*, vol. 1, Springer, 1969.
19. N. Bourbaki, *Integration. II. Chapters 7–9*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 2004, Translated from the 1963 and 1969 French originals by Sterling K. Berberian.
20. T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Graduate Texts in Mathematics, vol. 98, Springer-Verlag, New York, 1985.
21. N.L. Carothers, *A Short Course on Banach Space Theory*, London Mathematical Society Student Texts, vol. 64, Cambridge University Press, Cambridge, 2005.
22. S.M. Carroll, *Spacetime and Geometry*, Cambridge University Press, 2019.
23. W. Cheney, *Analysis for Applied Mathematics*, Graduate Texts in Mathematics, vol. 208, Springer-Verlag, New York, 2001.
24. P.G. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.
25. M. Creutz, *Quarks, Gluons and Lattices*, vol. 8, Cambridge University Press, Cambridge, 1983.
26. F. de Felice and C.J.S. Clarke, *Relativity on Curved Manifolds*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1990.

27. J. Dever, *Local Hausdorff measure*, arXiv preprint arXiv:1610.00078 (2016).
28. J. Diestel, *Sequences and Series in Banach Spaces*, Graduate Texts in Mathematics, vol. 92, Springer-Verlag, New York, 1984.
29. J. Diestel and J.J. Uhl, Jr., *Vector Measures*, American Mathematical Society, Providence, R.I., 1977, With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.
30. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York-London, 1969, Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.
31. J. Dimock, *Quantum Mechanics and Quantum Field Theory*, Cambridge University Press, Cambridge, 2011, A mathematical primer.
32. P.A.M. Dirac, *The quantum theory of the electron*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **117** (1928), no. 778, 610–624.
33. ———, *The Principles of Quantum Mechanics*, first ed., Clarendon Press, Oxford, 1930.
34. ———, *A theory of electrons and protons*, Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character **126** (1930), no. 801, 360–365.
35. ———, *Discussion of the infinite distribution of electrons in the theory of the positron*, Mathematical Proceedings of the Cambridge Philosophical Society, vol. 30, Cambridge University Press, 1934, pp. 150–163.
36. R.E. Edwards, *A theory of Radon measures on locally compact spaces*, Acta Math. **89** (1953), 133–164.
37. A. Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik **4** (1905).
38. ———, *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, Sitzungsberichte der Preußischen Akad. d. Wissenschaften (Berlin) (1917), 142–152.
39. R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*, second ed., John Wiley & Sons, Inc., New York, 1985.
40. J. Elstrodt, *Maß- und Integrationstheorie*, fourth ed., Springer-Lehrbuch, Springer-Verlag, Berlin, 2005, Grundwissen Mathematik.
41. R. Engelking, *General Topology*, second ed., Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989, Translated from the Polish by the author.
42. ———, *Theory of Dimensions Finite and Infinite*, Sigma Series in Pure Mathematics, vol. 10, Heldermann Verlag, Lemgo, 1995.
43. H. Federer, *Geometric Measure Theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.
44. E.A. Feinberg, P.O. Kasyanov, and N.V. Zadoianchuk, *Fatou's lemma for weakly converging probabilities*, arXiv:1206.4073 [math.PR], Theory Probab. Appl. **58** (2014), no. 4, 683–689.
45. R.P. Feynman and R.B. Leighton and M. Sands, *The Feynman Lectures on Physics. Vol. 1: Mainly mechanics, radiation, and heat*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1963.
46. ———, *The Feynman Lectures on Physics. Vol. 2: Mainly electromagnetism and matter*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1964.
47. ———, *The Feynman Lectures on Physics. Vol. 3: Quantum mechanics*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1965.
48. F. Finster, *Ableitung von Feldgleichungen aus dem Prinzip des fermionischen Projektors*, arXiv preprint gr-qc/9606040 (1996).
49. ———, *Derivation of local gauge freedom from a measurement principle*, arXiv preprint funct-an/9701002 (1997).
50. ———, *Definition of the Dirac sea in the presence of external fields*, Adv. Theor. Math. Phys. **2** (1998), no. 5, 963–985.
51. ———, *Local U(2,2) symmetry in relativistic quantum mechanics*, Journal of Mathematical Physics **39** (1998), no. 12, 6276–6290.
52. ———, *Light-cone expansion of the Dirac sea in the presence of chiral and scalar potentials*, Journal of Mathematical Physics **41** (2000), no. 10, 6689–6746.
53. ———, *The Principle of the Fermionic Projector*, AMS/IP Studies in Advanced Mathematics, vol. 35, American Mathematical Society, Providence, RI; International Press, Somerville, MA, 2006.
54. ———, *A variational principle in discrete space-time: Existence of minimizers*, Calc. Var. Partial Differential Equations **29** (2007), no. 4, 431–453.
55. ———, *From discrete space-time to Minkowski space: Basic mechanisms, methods and perspectives*, Quantum Field Theory, Springer, 2009, pp. 235–259.

56. ———, *Causal variational principles on measure spaces*, J. Reine Angew. Math. **646** (2010), 141–194.
57. ———, *A formulation of quantum field theory realizing a sea of interacting Dirac particles*, Letters in Mathematical Physics **97** (2011), no. 2, 165–183.
58. ———, *Perturbative quantum field theory in the framework of the fermionic projector*, J. Math. Phys. **55** (2014), no. 4, 042301, 53.
59. ———, *The Continuum Limit of Causal Fermion Systems*, Fundamental Theories of Physics, vol. 186, Springer, 2016, From Planck scale structures to macroscopic physics.
60. F. Finster and A. Grotz, *A Lorentzian quantum geometry*, Advances in Theoretical and Mathematical Physics **16** (2012), no. 4, 1197–1290.
61. F. Finster, A. Grotz and D. Schiefeneder, *Causal fermion systems: a quantum space-time emerging from an action principle*, (2012), 157–182.
62. F. Finster and M. Jokel, *Causal fermion systems: An elementary introduction to physical ideas and mathematical concepts*, Progress and Visions in Quantum Theory in View of Gravity (2020), 63–92.
63. F. Finster and N. Kamran, *Complex structures on jet spaces and bosonic Fock space dynamics for causal variational principles*, arXiv preprint arXiv:1808.03177 (2018).
64. F. Finster and J. Kleiner, *Causal fermion systems as a candidate for a unified physical theory*, Journal of Physics: Conference Series **626** (2015), 012020.
65. ———, *A Hamiltonian formulation of causal variational principles*, Calc. Var. Partial Differential Equations **56** (2017), no. 3, Paper No. 73, 33.
66. F. Finster and C. Langer, *Causal variational principles in the sigma-locally compact setting: Existence of minimizers*, Advances in Calculus of Variations **1** (2020) no. ahead-of-print.
67. F. Finster and M. Lottner, *Banach manifold structure and infinite-dimensional analysis for causal fermion systems*, arXiv preprint arXiv:2101.11908 (2021).
68. F. Finster and D. Schiefeneder, *On the support of minimizers of causal variational principles*, Arch. Ration. Mech. Anal. **210** (2013), no. 2, 321–364.
69. F. Finster and J. Tolksdorf, *Perturbative description of the fermionic projector: Normalization, causality, and Furry's theorem*, Journal of Mathematical Physics **55** (2014), no. 5, 052301.
70. T. Fließbach, *Mechanik*, Bibliographisches Institut, Mannheim, 1992.
71. G.B. Folland, *A Course in Abstract Harmonic Analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995.
72. ———, *Real Analysis: Modern Techniques and Their Applications*, Pure and Applied Mathematics (New York), Second, John Wiley & Sons, Inc., New York, 1999.
73. ———, *Quantum Field Theory: A tourist guide for mathematicians*, Mathematical Surveys and Monographs, vol. 149, American Mathematical Society, Providence, RI, 2008.
74. T. Friedrich, *Dirac Operators in Riemannian Geometry*, Graduate Studies in Mathematics, vol. 25, American Mathematical Society, Providence, RI, 2000, Translated from the 1997 German original by Andreas Nestke.
75. S.A. Gaal, *Point Set Topology*, Pure and Applied Mathematics, Vol. XVI, Academic Press, New York-London, 1964.
76. R.J. Gardner and W.F. Pfeffer, *Borel measures*, Handbook of set-theoretic topology (1984), 961–1043.
77. C. Gattringer and C.B. Lang, *Quantum Chromodynamics on the Lattice*, Lecture Notes in Physics, vol. 788, An introductory presentation, Springer-Verlag, Berlin, 2010, An introductory presentation.
78. M. Giaquinta and S. Hildebrandt, *Calculus of Variations I: The Lagrangian formalism*, Grundlehren der Mathematischen Wissenschaften, vol. 310, Springer-Verlag, Berlin, 1996.
79. I. Gohberg and P. Lancaster and L. Rodman, *Indefinite Linear Algebra and Applications*, Birkhäuser Verlag, Basel, 2005.
80. J. Gribbin, *Q is for Quantum: An Encyclopedia of Particle Physics*, The Free Press, New York, 1998.
81. D. Griffiths, *Introduction to Elementary Particles*, John Wiley & Sons, 2008.
82. D.J. Griffiths, *Introduction to Quantum Mechanics*, Pearson, 2005.
83. O. Grøn and S. Hervik, *Einstein's General Theory of Relativity*, Springer, New York, 2007, With modern applications in cosmology.
84. R.C. Gunning and H. Rossi, *Analytic Functions of Several Complex Variables*, AMS Chelsea Publishing, Providence, RI, 2009, Reprint of the 1965 original.
85. B.C. Hall, *Lie Groups, Lie Algebras, and Representations*, Graduate Texts in Mathematics, vol. 222, Springer-Verlag, New York, 2003, An elementary introduction.

86. ———, *Quantum Theory for Mathematicians*, Graduate Texts in Mathematics, vol. 267, Springer, New York, 2013.
87. P.R. Halmos, *Measure Theory*, D. Van Nostrand Company, Inc., New York, N. Y., 1950.
88. R.S. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. (N.S.) **7**, (1982), no. 1, 65–222.
89. S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, London-New York, 1973, Cambridge Monographs on Mathematical Physics, No. 1.
90. D. Hilbert, *Die Grundlagen der Physik. (Erste Mitteilung.)*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1915** (1915), 395–408.
91. W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton Mathematical Series, v. 4, Princeton University Press, Princeton, N. J., 1941.
92. J.D. Jackson, *Classical Electrodynamics*, Second, John Wiley & Sons, Inc., New York-London-Sydney, 1975.
93. J. Jost and X. Li-Jost, *Calculus of Variations*, Cambridge Studies in Advanced Mathematics, vol. 64, Cambridge University Press, Cambridge, 1998.
94. M. Kaku, *Quantum Field Theory*, The Clarendon Press, Oxford University Press, New York, 1993, A modern introduction.
95. T. Kamihigashi, *Interchanging a limit and an integral: necessary and sufficient conditions*, Journal of Inequalities and Applications **2020** (2020), no. 1, 1–9.
96. G. Kane, *Modern Elementary Particle Physics: Explaining and Extending the Standard Model*, Cambridge University Press, 2017.
97. T. Kato, *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition.
98. L. Kaup and B. Kaup, *Holomorphic Functions of Several Variables*, De Gruyter Studies in Mathematics, vol. 3, Walter de Gruyter & Co., Berlin, 1983, An introduction to the fundamental theory, With the assistance of Gottfried Barthel, Translated from the German by Michael Bridgland.
99. A.S. Kechris, *Classical Descriptive Set Theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.
100. H. König, *Measure and Integration: An Advanced Course in Basic Procedures and Applications*, Springer Science & Business Media, 2009.
101. G. Köthe, *Topological Vector Spaces. I*, Translated from the German by D. J. H. Garling. Die Grundlehren der mathematischen Wissenschaften, Band 159, Springer-Verlag New York Inc., New York, 1969.
102. S. Kobayashi and K. Nomizu,*Foundations of Differential Geometry Vol. I*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1996, Reprint of the 1963 original, A Wiley-Interscience Publication.
103. H. Kragh, *Dirac: A Scientific Biography*, Cambridge University Press, 1990.
104. I. Kupka, *A generalised uniform convergence and Dini's theorem*, New Zealand J. Math. **27** (1998).
105. L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields. Course of Theoretical Physics, Vol. 2*, fourth revised English ed., Pergamon Press, Oxford-London-New York-Paris; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1975, Translated from the Russian by M. Hamermesh.
106. ———, *Mechanics. Course of Theoretical Physics, Vol. 1*, third ed, Pergamon Press, Oxford-London-New York-Paris; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1976, Translated from the Russian by J.B. Sykes and J.S. Bell.
107. ———, *Quantum mechanics: non-relativistic theory. Course of Theoretical Physics, Vol. 3*, third edition, revised and enlarged ed., Addison-Wesley Series in Advanced Physics, Pergamon Press Ltd., Oxford-New York-Beijing; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1977, Translated from the Russian by J.B. Sykes and J.S. Bell.
108. S. Lang, *Linear Algebra*, Undergraduate Texts in Mathematics, third ed., Springer-Verlag, New York, 1987.
109. C. Langer, *Homogeneous Causal Fermion Systems*, Master thesis, Universität Regensburg (2017).
110. ———, *Causal Variational Principles in the Infinite-Dimensional Setting: Existence of Minimizers*, arXiv preprint arXiv:2101.06140 (2021).
111. H. Langer, *Zur Spektraltheorie J-selbstadjungierter Operatoren*, Math. Ann. **146** (1962), 60–85.
112. H.B. Lawson, Jr. and M.-L. Michelsohn, *Spin Geometry*, Princeton Mathematical Series, vol. 38, Princeton University Press, Princeton, NJ, 1989.

113. P.D. Lax, *Functional Analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2002.
114. ———, *Linear Algebra and its Applications*, Pure and Applied Mathematics, second ed., Wiley-Interscience, Hoboken, NJ, 2007.
115. A. Lichnerowicz, *Elements of Tensor Calculus*, Methuen & Co. Ltd., London; John Wiley & Sons Inc., New York, 1962, Translated by J.W. Leech and D.J. Newman.
116. H.A. Lorentz, *Die relative Bewegung der Erde und des Äthers*, Abhandlungen über Theoretische Physik, BG Teubner, Leipzig **1907** (1892), 443–447.
117. T.-W. Ma, *Banach-Hilbert Spaces, Vector Measures and Group Representations*, World Scientific Publishing Co., Inc., River Edge, NJ, 2002.
118. K.B. Marathe and G. Martucci, *The Mathematical Foundations of Gauge Theories*, Studies in Mathematical Physics, vol. 5, North-Holland Publishing Co., Amsterdam, 1992.
119. R. Meise and D. Vogt, *Introduction to Functional Analysis*, Oxford Graduate Texts in Mathematics, vol. 2, The Clarendon Press, Oxford University Press, New York, 1997.
120. C.W. Misner and K.S. Thorne and J.A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Francisco, Calif., 1973.
121. J.W. Morgan, *An introduction to gauge theory*, Gauge theory and the topology of four-manifolds (Park City, UT, 1994), IAS/Park City Math. Ser., vol. 4, Amer. Math. Soc., Providence, RI, 1998, pp. 51–143.
122. J.R. Munkres, *Topology*, second ed., Prentice Hall, Inc., Upper Saddle River, NJ, 2000.
123. G.L. Naber, *The Geometry of Minkowski Spacetime*, Applied Mathematical Sciences, vol. 92, Springer, New York, 2012, An introduction to the mathematics of the special theory of relativity, Second edition.
124. I. Newton, *The* Principia*: mathematical principles of natural philosophy*, University of California Press, Berkeley, CA, 1999, A new translation by I. Bernard Cohen and Anne Whitman, assisted by Julia Budenz, Preceded by "A guide to Newton's *Principia*" by I.B. Cohen.
125. B. O'Neill, *Semi-Riemannian Geometry*, Pure and Applied Mathematics, vol. 103, Academic Press, Inc., New York, 1983, With applications to relativity.
126. M. Oppio, *On the mathematical foundations of causal fermion systems in Minkowski spacetime*, arXiv preprint arXiv:1909.09229 (2019).
127. P.B. Pal, *An Introductory Course of Particle Physics*, Taylor & Francis, 2014.
128. L.E. Parker and D.J. Toms, *Quantum Field Theory in Curved Spacetime*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 2009, Quantized fields and gravity.
129. M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1995, Edited and with a foreword by David Pines.
130. M. Planck, *Über irreversible Strahlungsvorgänge*, Annalen der Physik **306** (1900), no. 1, 69–122.
131. C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, second ed., Princeton University Press, Princeton and Oxford, 2013.
132. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I. Functional analysis*, Academic Press, New York-London, 1972.
133. ———, *Methods of Modern Mathematical Physics II. Fourier analysis, self-adjointness*, Academic Press, New York-London, 1975.
134. M. Renardy and R.C. Rogers, *An Introduction to Partial Differential Equations*, Texts in Applied Mathematics, vol. 13, second ed., Springer-Verlag, New York, 2004.
135. J. Roe, *Index Theory, Coarse Geometry, and Topology of Manifolds*, CBMS Regional Conference Series in Mathematics, vol. 90, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1996.
136. ———, *Lectures on Coarse Geometry*, University Lecture Series, vol. 31, American Mathematical Society, Providence, RI, 2003.
137. C.A. Rogers, *Hausdorff Measures*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1998, Reprint of the 1970 original, With a foreword by K. J. Falconer.
138. C. Rovelli, *Quantum Gravity*, Cambridge University Press, 2004.
139. W. Rudin, *Real and Complex Analysis*, third ed., McGraw-Hill Book Co., New York, 1987.
140. ———, *Functional Analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, 1991.

141. B. Russell, *History of Western Philosophy: Collectors Edition*, Routledge, 2013.
142. J.J. Sakurai, *Modern Quantum Mechanics*, revised ed., Addison-Wesley, 1994.
143. L. Schwartz, *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Published for the Tata Institute of Fundamental Research, Bombay by Oxford University Press, London, 1973, Tata Institute of Fundamental Research Studies in Mathematics, No. 6.
144. K. Schwarzschild, *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin) (1916), 189–196.
145. S.S. Schweber, *QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga*, Princeton Series in Physics, Princeton University Press, Princeton, NJ, 1994.
146. ———, *An Introduction to Relativistic Quantum Field Theory*, Courier Corporation, 2011.
147. R.U. Sexl and H.K. Urbantke, *Relativität, Gruppen, Teilchen*, second ed., Springer-Verlag, Vienna, 1982, Spezielle Relativitätstheorie als Grundlage der Feld- und Teilchenphysik.
148. R. Shankar, *Principles of Quantum Mechanics*, second ed., Plenum Press, New York, 1994.
149. D. Simon, *Albert Einstein: Akademie-Vorträge: Sitzungsberichte der Preußischen Akademie der Wissenschaften 1914-1932*, John Wiley & Sons, 2012.
150. L.A. Steen and J.A. Seebach, Jr., *Counterexamples in Topology*, Dover Publications, Inc., Mineola, NY, 1995, Reprint of the second (1978) edition.
151. S. Sternberg, *Group Theory and Physics*, Cambridge University Press, Cambridge, 1994.
152. B. Thaller, *The Dirac Equation*, Texts and Monographs in Physics, Springer-Verlag, Berlin, 1992.
153. F. Vester, *Denken, Lernen, Vergessen: Was geht in unserem Kopf vor, wie lernt das Gehirn, und wann lässt es uns im Stich?*, Deutsche Verlagsanstalt, 1975.
154. R.M. Wald, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, University of Chicago press, 1994.
155. F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York-Berlin, 1983, Corrected reprint of the 1971 edition.
156. S. Weinberg, *The Quantum Theory of Fields. Vol. I*, Cambridge University Press, Cambridge, 1996, Foundations, Corrected reprint of the 1995 original.
157. D. Werner, *Funktionalanalysis*, extended ed., Springer-Verlag, Berlin, 2000.
158. H. Weyl, *Quantenmechanik und Gruppentheorie*, Zeitschrift für Physik **46** (1927), no. 1-2, 1–46.
159. ———, *Gruppentheorie und Quantenmechanik*, second, Wissenschaftliche Buchgesellschaft, Darmstadt, 1977.
160. J.A. Wheeler and K. Ford, *Geons, Black Holes and Quantum Foam: A Life in Physics*, American Association of Physics Teachers, 2000.
161. S. Willard, *General Topology*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970.
162. R. Williamson and L. Janos, *Constructing metrics with the Heine-Borel property*, Proc. Amer. Math. Soc. **100** (1987), no. 3, 567–573.
163. C.-S. Wu and E. Ambler and R.W. Hayward and D.D. Hoppes and R.P. Hudson, *Experimental test of parity conservation in beta decay*, Physical review **105** (1957), no. 4, 1413.
164. C.-N. Yang and R.L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Physical review **96** (1954), no. 1, 191.
165. S.-T. Yau and S. Nadis, *The Shape of Inner Space: String Theory and the Geometry of the Universe's Hidden Dimensions*, Basic Books, 2010.
166. K. Yosida, *Functional Analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the sixth (1980) edition.
167. A. Zee, *Quantum Field Theory in a Nutshell*, second ed., Princeton University Press, Princeton, NJ, 2010.
168. E. Zeidler, *Nonlinear Functional Analysis and its Applications III*, Springer-Verlag, New York, 1985, Variational methods and optimization, Translated from the German by Leo F. Boron.
169. ———, *Nonlinear Functional Analysis and its Applications. IV: Applications to mathematical physics*, Springer-Verlag, New York, 1988.
170. ———, *Quantum Field Theory III. Gauge theory*, Heidelberg, 2011, A bridge between mathematicians and physicists, Springer.