Universität Regensburg
Mathematik

Causal fermion systems – an overview

Felix Finster

Preprint Nr. 09/2015

CAUSAL FERMION SYSTEMS – AN OVERVIEW

FELIX FINSTER

MAY 2015

Abstract. The theory of causal fermion systems is an approach to describe fundamental physics. We here introduce the mathematical framework and give an overview of the objectives and current results.

Contents

Causal fermion systems were introduced in [17] as a reformulation and generalization of the setting used in the fermionic projector approach [8]. The theory of causal fermion systems is an approach to describe fundamental physics. It gives quantum mechanics, general relativity and quantum field theory as limiting cases and is therefore a candidate for a unified physical theory. In this article, we introduce the mathematical framework and give an overview of the different limiting cases. The presentation

is self-contained and includes references to the corresponding research papers. The aim is not only to convey the underlying physical picture, but also to lay the mathematical foundations in a conceptually convincing way. This includes technical issues like specifying the topologies on the different spaces of functions and operators, giving a mathematical definition of an ultraviolet regularization, or specifying the maps which identify the objects of the causal fermion system with corresponding objects in Minkowski space. Also, we use a basis-independent notation whenever possible. The reader interested in a non-technical introduction is referred to [20].

1. The Abstract Framework

1.1. Basic Definitions.

For conceptual clarity, we begin with the general definitions.

Definition 1.1. *(causal fermion system)* Given a separable complex Hilbert space $\mathcal{H}$ with scalar product $\langle .|. \rangle_{\mathcal{H}}$ and a parameter $n \in \mathbb{N}$ (the *"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. On $\mathcal{F}$ we are given a positive measure ρ (defined on a σ-algebra of subsets of $\mathcal{F}$), the so-called *universal measure*. We refer to $(\mathcal{H}, \mathcal{F}, \rho)$ as a *causal fermion system*.

We remark that the separability of the Hilbert space (i.e. the assumption that $\mathcal{H}$ admits an at most countable Hilbert space basis) is not essential and could be left out. We included the separability assumption because it seems to cover all cases of physical interest and is useful if one wants to work with basis representations.

A causal fermion system describes a space-time together with all structures and objects therein (like the causal and metric structures, spinors and interacting quantum fields). In order to single out the physically admissible causal fermion systems, one must formulate physical equations. This is accomplished with the help of an action principle which we now introduce. For any $x, y \in \mathcal{F}$, the product xy is an operator of rank at most $2n$. We denote its non-trivial eigenvalues (counting algebraic multiplicities) by $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy} \in \mathbb{C}$. We introduce the *spectral weight* $|\,.\,|$ of an operator as the sum of the absolute values of its eigenvalues. In particular, the spectral weight of the operator products xy and $(xy)^2$ is defined 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 and the action by

$$\textit{Lagrangian:} \qquad \mathcal{L}(x,y) = \left|(xy)^2\right| - \frac{1}{2n} |xy|^2 \tag{1.1}$$

$$\textit{action:} \qquad \mathcal{S}(\rho) = \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d\rho(x)\, d\rho(y) . \tag{1.2}$$

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

$$\textit{volume constraint:} \qquad \rho(\mathcal{F}) = \text{const} \tag{1.3}$$

$$\textit{trace constraint:} \qquad \int_{\mathcal{F}} \mathrm{tr}(x)\, d\rho(x) = \text{const} \tag{1.4}$$

$$\textit{boundedness constraint:} \qquad \mathcal{T} := \iint_{\mathcal{F}\times\mathcal{F}} |xy|^2\, d\rho(x)\, d\rho(y) \leq C , \tag{1.5}$$

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

In order to make the causal action principle mathematically well-defined, one needs to specify the class of measures in which to vary ρ. To this end, on $\mathcal{F}$ we consider the topology induced by the operator norm

$$\|A\| := \sup \big\{ \|Au\|_{\mathcal{H}} \text{ with } \|u\|_{\mathcal{H}} = 1 \big\} . \tag{1.6}$$

In this topology, the Lagrangian as well as the integrands in (1.4) and (1.5) are continuous. The σ-algebra generated by the open sets of $\mathcal{F}$ consists of the so-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 for example [30, §52]). The right prescription is to vary ρ within the class of regular Borel measures of $\mathcal{F}$. In the so-called *finite-dimensional setting* when $\mathcal{H}$ is finite-dimensional and the total volume $\rho(\mathcal{F})$ is finite, the existence of minimizers is proven in [9, 10], and the properties of minimizing measures are analyzed in [26, 1].

The causal action principle also makes mathematical sense in the so-called *infinite-dimensional setting* when $\mathcal{H}$ is infinite-dimensional and the total volume $\rho(\mathcal{F})$ is infinite. In this case, the volume constraint (1.3) is implemented by demanding that all variations $(\rho(\tau))_{\tau \in (-\varepsilon, \varepsilon)}$ should for all $\tau, \tau' \in (-\varepsilon, \varepsilon)$ satisfy the conditions

$$\big|\rho(\tau) - \rho(\tau')\big|(\mathcal{F}) < \infty \qquad \text{and} \qquad \big(\rho(\tau) - \rho(\tau')\big)(\mathcal{F}) = 0 \tag{1.7}$$

(where $|.|$ denotes the total variation of a measure; see [30, §28]). The existence theory in the infinite-dimensional setting has not yet been developed. But it is known that the Euler-Lagrange equations corresponding to the causal action principle still have a mathematical meaning (as will be explained in §4.1 below). This makes it possible to analyze the causal action principle without restrictions on the dimension of $\mathcal{H}$ nor on the total volume. One way of getting along without an existence theory in the infinite-dimensional setting is to take the point of view that on a fundamental physical level, the Hilbert space $\mathcal{H}$ is finite-dimensional, whereas the infinite-dimensional setting merely is a mathematical idealization needed in order to describe systems involving an infinite number of quantum particles.

We finally explain the significance of the constraints. Generally speaking, the constraints (1.3)–(1.5) are needed to avoid trivial minimizers and in order for the variational principle to be well-posed. More specifically, if we dropped the constraint of fixed total volume (1.3), the measure $\rho = 0$ would be trivial minimizer. Without the boundedness constraint (1.5), the loss of compactness discussed in [10, Section 2.2] implies that no minimizers exist. If, on the other hand, we dropped the trace constraint (1.4), a trivial minimizer could be constructed as follows. We let x be the operator with the matrix representation

$$x = \text{diag}\big(\underbrace{1, \ldots, 1}_{n \text{ times}}, \underbrace{-1, \ldots, -1}_{n \text{ times}}, 0, 0, \ldots \big)$$

and choose ρ as a multiple of the Dirac measure supported at x. Then $\mathcal{T} > 0$ but $\mathcal{S} = 0$.

1.2. **Space-Time and Causal Structure.** A causal fermion system $(\mathcal{H}, \mathcal{F}, \rho)$ encodes a large amount of information. In order to recover this information, one can for example form products of linear operators in $\mathcal{F}$, compute the eigenvalues of such operator products and integrate expressions involving these eigenvalues with respect to the universal measure. However, it is not obvious what all this information means.

In order to clarify the situation, we now introduce additional mathematical objects. These objects are *inherent* in the sense that we only use information already encoded in the causal fermion system.

We first define *space-time*, denoted by M, as the support of the universal measure,

$$M := \operatorname{supp} \rho \subset \mathcal{F} .$$

On M we consider the topology induced by $\mathcal{F}$ (generated by the sup-norm (1.6) on $\mathrm{L}(\mathcal{H})$). Moreover, the universal measure $\rho|_M$ restricted to M can be regarded as a volume measure on space-time. This makes space-time into a *topological measure space*. Furthermore, one has the following notion of causality:

Definition 1.2. *(causal structure)* For any $x, y \in \mathcal{F}$, the product xy is an operator of rank at most $2n$. We denote its non-trivial eigenvalues (counting algebraic multiplicities) by $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy}$. The points x and y are called *spacelike* separated if all the λ_j^{xy} have the same absolute value. They are said to be *timelike* separated if the λ_j^{xy} are all real and do not all have the same absolute value. In all other cases (i.e. if the λ_j^{xy} are not all real and do not all have the same absolute value), the points x and y are said to be *lightlike* separated.

Restricting the causal structure of $\mathcal{F}$ to M, we get causal relations in space-time. To avoid confusion, we remark that in earlier papers (see [15], [17]) a slightly different definition of the causal structure was used. But the modified definition used here seems preferable.

The Lagrangian (1.1) is compatible with the above notion of causality in the following sense. Suppose that two points $x, y \in \mathcal{F}$ are spacelike separated. Then the eigenvalues λ_i^{xy} all have the same absolute value. Rewriting (1.1) as

$$\mathcal{L} = \sum_{i=1}^{2n} |\lambda_i^{xy}|^2 - \frac{1}{2n} \sum_{i,j=1}^{2n} |\lambda_i^{xy}|\, |\lambda_j^{xy}| = \frac{1}{4n} \sum_{i,j=1}^{2n} \Big(|\lambda_i^{xy}| - |\lambda_j^{xy}| \Big)^2 ,$$

one concludes that the Lagrangian vanishes. Thus pairs of points with spacelike separation do not enter the action. This can be seen in analogy to the usual notion of causality where points with spacelike separation cannot influence each other[1]. This analogy is the reason for the notion "causal" in "causal fermion system" and "causal action principle."

The above notion of causality is *symmetric* in x and y, as we now explain. Since the trace is invariant under cyclic permutations, we know that

$$\operatorname{tr}\big((xy)^p\big) = \operatorname{tr}\big(x\,(yx)^{p-1}\,y\big) = \operatorname{tr}\big((yx)^{p-1}\,yx\big) = \operatorname{tr}\big((yx)^p\big) \tag{1.8}$$

(where tr again denotes the trace of a linear operator on $\mathcal{H}$). Since all our operators have finite rank, there is a finite-dimensional subspace I of $\mathcal{H}$ such that xy maps I to itself and vanishes on the orthogonal complement of I. Then the non-trivial eigenvalues of the operator product xy are given as the zeros of the characteristic polynomial of the restriction $xy|_I : I \to I$. The coefficients of this characteristic polynomial (like the trace, the determinant, etc.) are symmetric polynomials in the eigenvalues and can therefore be expressed in terms of traces of powers of xy. As a consequence, the identity (1.8) implies that the operators xy and yx have the same characteristic polynomial and are thus isospectral. This shows that the causal notions are indeed symmetric in

[1]For clarity, we point out that our notion of causality does allow for nonlocal correlations and entanglement between regions with space-like separation. This will become clear in §1.4 and Section 5.3.

the sense that x and y are spacelike separated if and only if y and x are (and similarly for timelike and lightlike separation). One also sees that the Lagrangian $\mathcal{L}(x,y)$ is symmetric in its two arguments.

A causal fermion system also distinguishes a *direction of time*. To this end, we let π_x be the orthogonal projection in $\mathcal{H}$ on the subspace $x(\mathcal{H}) \subset \mathcal{H}$ and introduce the functional

$$\mathcal{C} \,:\, M \times M \to \mathbb{R}\,, \qquad \mathcal{C}(x,y) := i\,\mathrm{Tr}\big(y\,x\,\pi_y\,\pi_x - x\,y\,\pi_x\,\pi_y\big) \tag{1.9}$$

(this functional was first stated in [18, Section 7.5], motivated by constructions in [15, Section 3.5]). Obviously, this functional is anti-symmetric in its two arguments. This makes it possible to introduce the notions

$$\begin{cases} y \text{ lies in the } \textit{future} \text{ of } x & \text{if } \mathcal{C}(x,y) > 0 \\ y \text{ lies in the } \textit{past} \text{ of } x & \text{if } \mathcal{C}(x,y) < 0\,. \end{cases} \tag{1.10}$$

By distinguishing a direction of time, we get a structure similar to a causal set (see for example [3]). But in contrast to a causal set, our notion of "lies in the future of" is not necessarily transitive. This corresponds to our physical conception that the transitivity of the causal relations could be violated both on the cosmological scale (there might be closed timelike curves) and on the microscopic scale (there seems no compelling reason why the causal relations should be transitive down to the Planck scale). This is the reason why we consider other structures (namely the universal measure and the causal action principle) as being more fundamental. In our setting, causality merely is a derived structure encoded in the causal fermion system.

1.3. **The Kernel of the Fermionic Projector.** The causal action principle depends crucially on the eigenvalues of the operator product xy with $x, y \in \mathcal{F}$. For computing these eigenvalues, it is convenient not to consider this operator product on the (possibly infinite-dimensional) Hilbert space $\mathcal{H}$, but instead to restrict attention to a finite-dimensional subspace of $\mathcal{H}$, chosen such that the operator product vanishes on the orthogonal complement of this subspace. This construction leads us to the spin spaces and to the kernel of the fermionic projector, which we now introduce. For every $x \in \mathcal{F}$ we define the *spin space* S_x by $S_x = x(\mathcal{H})$; it is a subspace of $\mathcal{H}$ of dimension at most $2n$. For any $x, y \in M$ we define the *kernel of the fermionic operator* $P(x,y)$ by

$$P(x,y) = \pi_x\, y|_{S_y} \,:\, S_y \to S_x \tag{1.11}$$

(where π_x is again the orthogonal projection on the subspace $x(\mathcal{H}) \subset \mathcal{H}$). Taking the trace of (1.11) in the case $x = y$, one finds that $\mathrm{tr}(x) = \mathrm{Tr}_{S_x}(P_\tau(x,x))$, making it possible to express the integrand of the trace constraint (1.4) in terms of the kernel of the fermionic operator. In order to also express the eigenvalues of the operator xy, we define the *closed chain* A_{xy} as the product

$$A_{xy} = P(x,y)\,P(y,x) \,:\, S_x \to S_x\,. \tag{1.12}$$

Computing powers of the closed chain, one obtains

$$A_{xy} = (\pi_x y)(\pi_y x)|_{S_x} = \pi_x\, yx|_{S_x}\,, \qquad (A_{xy})^p = \pi_x\,(yx)^p|_{S_x}\,.$$

Taking the trace, one sees in particular that $\mathrm{Tr}_{S_x}(A_{xy}^p) = \mathrm{tr}\big((yx)^p\big)$. Repeating the arguments after (1.8), one concludes that the eigenvalues of the closed chain coincide with the non-trivial eigenvalues $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy}$ of the operator xy in Definition 1.2. Therefore, the kernel of the fermionic operator encodes the causal structure of M. The main advantage of working with the kernel of the fermionic operator is that the closed

chain (1.12) is a linear operator on a vector space of dimension at most $2n$, making it possible to compute the $\lambda_1^{xy}, \ldots, \lambda_{2n}^{xy}$ as the eigenvalues of a finite matrix.

Next, it is very convenient to arrange that the kernel of the fermionic operator is symmetric in the sense that

$$P(x,y)^* = P(y,x)\,. \tag{1.13}$$

To this end, one chooses on the spin space S_x the *spin scalar product* $\prec .|. \succ_x$ by

$$\prec u|v\succ_x = -\langle u|xu\rangle_{\mathcal{H}} \qquad (\text{for all } u,v \in S_x)\,. \tag{1.14}$$

Due to the factor x on the right, this definition really makes the kernel of the fermionic operator symmetric, as is verified by the computation

$$\begin{aligned} \prec u \,|\, P(x,y)\, v \succ_x &= -\langle u \,|\, x\, P(x,y)\, v\rangle_{\mathcal{H}} = -\langle u \,|\, xy\, v\rangle_{\mathcal{H}} \\ &= -\langle \pi_y\, x\, u \,|\, y\, v\rangle_{\mathcal{H}} = \prec P(y,x)\, u \,|\, v \succ_y \end{aligned}$$

(where $u \in S_x$ and $v \in S_y$). The spin space $(S_x, \prec .|. \succ_x)$ is an *indefinite* inner product of signature (p,q) with $p,q \leq n$. In this way, indefinite inner product spaces arise naturally when analyzing the mathematical structure of the causal action principle.

The kernel of the fermionic operator as defined by (1.11) is also referred to as the kernel of the *fermionic projector*, provided that suitable normalization conditions are satisfied. Different normalization conditions have been proposed and analyzed (see the discussion in [28, Section 2.2]). More recently, it was observed in [21] that one of these normalization conditions is automatically satisfied if the universal measure is a minimizer of the causal action principle (see §4.2 below). With this in mind, we no longer need to be so careful about the normalization. For notational simplicity, we always refer to $P(x,y)$ as the kernel of the fermionic projector.

1.4. **Wave Functions and Spinors.** For clarity, we sometimes denote the spin space S_x at a space-time point $x \in M$ by S_xM. A *wave function* ψ is defined as a function which to every $x \in M$ associates a vector of the corresponding spin space,

$$\psi \,:\, M \to \mathcal{H} \qquad \text{with} \qquad \psi(x) \in S_xM \quad \text{for all } x \in M\,. \tag{1.15}$$

We now want to define what we mean by *continuity* of a wave function. For the notion of continuity, we need to compare the wave function at different space-time points, being vectors $\psi(x) \in S_xM$ and $\psi(y) \in S_y\mathrm{M}$ in different spin spaces. Using that both spin spaces S_xM and S_yM are subspaces of the same Hilbert space $\mathcal{H}$, an obvious idea is to simply work with the Hilbert space norm $\|\psi(x) - \psi(y)\|_{\mathcal{H}}$. However, in view of the factor x in the spin scalar product (1.14), it is preferable to insert a corresponding power of the operator x. Namely, the natural norm on the spin space $(S_x, \prec .|. \succ_x)$ is given by

$$\big|\psi(x)\big|_x^2 := \big\langle \psi(x) \,\big|\, |x|\, \psi(x) \big\rangle_{\mathcal{H}} = \Big\| \sqrt{|x|}\, \psi(x) \Big\|_{\mathcal{H}}^2$$

(where $|x|$ is the absolute value of the symmetric operator x on $\mathcal{H}$, and $\sqrt{|x|}$ the square root thereof). This leads us to defining that the wave function ψ is *continuous* at x if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$\big\| \sqrt{|y|}\, \psi(y) - \sqrt{|x|}\, \psi(x) \big\|_{\mathcal{H}} < \varepsilon \qquad \text{for all } y \in M \text{ with } \|y-x\| \leq \delta\,.$$

Likewise, ψ is said to be continuous on M if it continuous at every $x \in M$. We denote the set of continuous wave functions by $C^0(M, SM)$. Clearly, the space of continuous wave functions is a complex vector space with pointwise operations, i.e. $(\alpha\psi + \beta\phi)(x) := \alpha\psi(x) + \beta\phi(x)$ with $\alpha, \beta \in \mathbb{C}$.

It is an important observation that every vector $u \in \mathcal{H}$ of the Hilbert space gives rise to a unique wave function. To obtain this wave function, denoted by ψ^u, we simply project the vector u to the corresponding spin spaces,

$$\psi^u \;:\; M \to \mathcal{H}\,, \qquad \psi^u(x) = \pi_x u \in S_x M\,. \tag{1.16}$$

We refer to ψ^u as the *physical wave function* of $u \in \mathcal{H}$. The estimate[2]

$$\begin{aligned} \Big\| \sqrt{|y|}\, \psi^u(y) - \sqrt{|x|}\, \psi^u(x) \Big\|_{\mathcal{H}} &= \Big\| \sqrt{|y|}\, u - \sqrt{|x|}\, u \Big\|_{\mathcal{H}} \\ &\leq \Big\| \sqrt{|y|} - \sqrt{|x|} \Big\|\, \|u\|_{\mathcal{H}} \overset{(\star)}{\leq} \|y-x\|^{\frac{1}{4}}\, \|y+x\|^{\frac{1}{4}}\, \|u\|_{\mathcal{H}} \end{aligned}$$

shows that ψ^u is indeed continuous. The physical picture is that the physical wave functions ψ^u are those wave functions which are realized in the physical system. Using a common physical notion, one could say that the vectors in $\mathcal{H}$ correspond to the "occupied states" of the system, and that an occupied state $u \in \mathcal{H}$ is represented in space-time by the corresponding physical wave function ψ^u. The shortcoming of this notion is that an "occupied state" is defined only for free quantum fields, whereas the physical wave functions are defined also in the interacting theory. For this reason, we prefer not use the notion of "occupied states."

For a convenient notation, we also introduce the *wave evaluation operator* Ψ as an operator which to every Hilbert space vector associates the corresponding physical wave function,

$$\Psi \;:\; \mathcal{H} \to C^0(M, SM)\,, \qquad u \mapsto \psi^u\,. \tag{1.18}$$

Evaluating at a fixed space-time point gives the mapping

$$\Psi(x) \;:\; \mathcal{H} \to S_x M\,, \qquad u \mapsto \psi^u(x)\,.$$

The kernel of the fermionic projector can be expressed in terms of the wave evaluation operator:

Lemma 1.3. *For any $x, y \in M$,*

$$x = -\Psi(x)^*\, \Psi(x) \tag{1.19}$$

$$P(x,y) = -\Psi(x)\, \Psi(y)^*\,. \tag{1.20}$$

[2]For completeness, we derive the inequality $(\star)$: Since the operator $\sqrt{|y|} - \sqrt{|x|}$ is symmetric and has finite rank, there is a normalized vector $u \in \mathcal{H}$ such that

$$\Big(\sqrt{|y|} - \sqrt{|x|}\Big) u = \pm \Big\| \sqrt{|y|} - \sqrt{|x|} \Big\|\, u\,. \tag{1.17}$$

Possibly by exchanging the roles of x and y we can arrange the plus sign. Then

$$\Big\| \sqrt{|y|} - \sqrt{|x|} \Big\| = \langle u \,|\, \Big(\sqrt{|y|} - \sqrt{|x|}\Big) u \rangle \leq \langle u \,|\, \Big(\sqrt{|y|} + \sqrt{|x|}\Big) u \rangle\,,$$

where in the last step we used that the operator $\sqrt{|x|}$ is positive. Multiplying by $\big\| \sqrt{|y|} - \sqrt{|x|} \big\|$ and using (1.17) with the plus sign, we obtain

$$\begin{aligned} \Big\| \sqrt{|y|} - \sqrt{|x|} \Big\|^2 &\leq \frac{1}{2} \Big(\langle u \,|\, \Big(\sqrt{|y|} + \sqrt{|x|}\Big) \Big(\sqrt{|y|} - \sqrt{|x|}\Big) u \rangle + \langle \Big(\sqrt{|y|} - \sqrt{|x|}\Big) u \,|\, \Big(\sqrt{|y|} + \sqrt{|x|}\Big) u \rangle \Big) \\ &= \frac{1}{2}\, \langle u \,|\, \Big\{ \Big(\sqrt{|y|} + \sqrt{|x|}\Big), \Big(\sqrt{|y|} - \sqrt{|x|}\Big) \Big\}\, u \rangle = \langle u \,|\, \big(|y| - |x|\big)\, u \rangle \leq \big\| |y| - |x| \big\|\,. \end{aligned}$$

We thus obtain the inequality $\big\| \sqrt{|y|} - \sqrt{|x|} \big\|^2 \leq \big\| |y| - |x| \big\|$. Applying this inequality with x replaced by x^2 and y replaced by y^2, it also follows that $\big\| |y| - |x| \big\|^2 \leq \big\| y^2 - x^2 \big\| \leq \big\| y - x \big\|\, \big\| y + x \big\|$. Combining these inequalities gives $(\star)$.

Proof. For any $v \in S_xM$ and $u \in \mathcal{H}$,

$$\prec v \,|\, \Psi(x)\, u \succ_x = \prec v \,|\, \pi_x\, u \succ_x \overset{(1.14)}{=} -\langle v \,|\, x\, u \rangle_{\mathcal{H}} = \langle (-x)\, v \,|\, u \rangle_{\mathcal{H}}$$

and thus

$$\Psi(x)^* = -x|_{S_xM} \;:\; S_xM \to \mathcal{H}\,.$$

Hence

$$\Psi(x)^*\, \Psi(x)\, u = \Psi(x)^*\, \psi^u_x = -x\, \psi^u_x \overset{(1.16)}{=} -x\, \pi_x u = -xu\,,$$

proving (1.19). Similarly, the relation (1.20) follows from the computation

$$\Psi(x)\, \Psi(y)^* = -\pi_x\, y|_{S_y} = -P(x,y)\,.$$

This completes the proof. □

The structure of the wave functions (1.15) taking values in the spin spaces is reminiscent of sections of a vector bundle. The only difference is that our setting is more general in that the base space M does not need to be a manifold, and the fibres S_xM do not need to depend smoothly on the base point x. However, comparing to the setting of spinors in Minkowski space or on a Lorentzian manifold, one important structure is missing: we have no Dirac matrices and no notion of Clifford multiplication. The following definition is a step towards introducing these additional structures.

Definition 1.4. *(Clifford subspace)* We denote the space of symmetric linear operators on $(S_x, \prec.|.\succ_x)$ by $\mathrm{Symm}(S_x) \subset \mathrm{L}(S_x)$. A subspace $K \subset \mathrm{Symm}(S_x)$ is called a *Clifford subspace* of signature (r,s) at the point x (with $r, s \in \mathbb{N}_0$) if the following conditions hold:

(i) For any $u, v \in K$, the anti-commutator $\{u,v\} \equiv uv + vu$ is a multiple of the identity on S_x.

(ii) The bilinear form $\langle ., . \rangle$ on K defined by

$$\frac{1}{2}\{u,v\} = \langle u, v \rangle \mathbb{1} \qquad \text{for all } u, v \in K \tag{1.21}$$

is non-degenerate and has signature (r,s).

In view of the anti-commutation relations (1.21), a Clifford subspace can be regarded as a generalization of the space spanned by the usual Dirac matrices. However, the above definition has two shortcomings: First, there are many different Clifford subspaces, so that there is no unique notion of Clifford multiplication. Second, we are missing the structure of tangent vectors as well as a mapping which would associate a tangent vector to an element of the Clifford subspace.

These shortcomings can be overcome by using either geometric or measure-theoretic methods. In the geometric approach, one gets along with the non-uniqueness of the Clifford subspaces by working with suitable equivalence classes. Using geometric information encoded in the causal fermion system, one can then construct mappings between the equivalence classes at different space-time points. This method will be outlined in §1.6. In the measure-theoretic approach, on the other hand, one uses the local form of the universal measure with the aim of constructing a unique Clifford subspace at every space-time point. This will be outlined in §1.7. Before entering these geometric and measure-theoretic constructions, we introduce additional structures on the space of wave functions.

1.5. **The Fermionic Projector on the Krein Space.** The space of wave functions can be endowed with an inner product and a topology. The inner product is defined by

$$<\psi|\phi> = \int_M \prec\psi(x)|\phi(x)\succ_x d\rho(x)\,. \tag{1.22}$$

In order to ensure that the last integral converges, we also introduce the scalar product $\langle\!\langle .|. \rangle\!\rangle$ by

$$\langle\!\langle \psi|\phi \rangle\!\rangle = \int_M \langle \psi(x)|\,|x|\,\phi(x)\rangle_{\mathcal{H}}\, d\rho(x) \tag{1.23}$$

(where $|x|$ is again the absolute value of the symmetric operator x on $\mathcal{H}$). The *one-particle space* $(\mathcal{K}, <.|.>)$ is defined as the space of wave functions for which the corresponding norm $|||\,.\,|||$ is finite, with the topology induced by this norm, and endowed with the inner product $<.|.>$. Such an indefinite inner product space with a topology induced by an additional scalar product is referred to as a *Krein space* (see for example [2, 32]).

When working with the one-particle Krein space, one must keep in mind that the physical wave function ψ^u of a vector $u \in \mathcal{H}$ does not need to be a vector in $\mathcal{K}$ because the corresponding integral in (1.22) may diverge. Similarly, the scalar product $\langle\!\langle \psi^u|\psi^u \rangle\!\rangle$ may be infinite. One could impose conditions on the causal fermion system which ensure that the integrals in (1.22) and (1.23) are finite for all physical wave functions. Then the mapping $u \mapsto \psi^u$ would give rise to an embedding $\mathcal{H} \hookrightarrow \mathcal{K}$ of the Hilbert space $\mathcal{H}$ into the one-particle Krein space. However, such conditions seem too restrictive and are not really needed. Therefore, here we shall not impose any conditions on the causal fermion systems but simply keep in mind that the physical wave functions are in general no Krein vectors.

Despite this shortcoming, the Krein space is useful because the kernel of the fermionic projector gives rise to an operator on $\mathcal{K}$. Namely, choosing a suitable dense domain of definition[3] $\mathcal{D}(P)$, we can regard $P(x,y)$ as the integral kernel of a corresponding operator P,

$$P\,:\,\mathcal{D}(P) \subset \mathcal{K} \to \mathcal{K}\,, \qquad (P\psi)(x) = \int_M P(x,y)\,\psi(y)\,d\rho(y)\,, \tag{1.24}$$

referred to as the *fermionic projector*. The fermionic projector has the following two useful properties:

- P is *symmetric* in the sense that $<P\psi|\phi> = <\psi|P\phi>$ for all $\psi, \phi \in \mathcal{D}(P)$:
  The symmetry of the kernel of the fermionic projector (1.13) implies that
  $$\prec P(x,y)\psi(y)\,|\,\psi(x)\succ_x = \prec\psi(y)\,|\,P(y,x)\psi(x)\succ_y .$$
  Integrating over x and y and applying (1.24) and (1.22) gives the result.
- $(-P)$ is *positive* in the sense that $<\psi|(-P)\psi> \geq 0$ for all $\psi \in \mathcal{D}(P)$:

[3]For example, one may choose $\mathcal{D}(P)$ as the set of all vectors $\psi \in \mathcal{K}$ satisfying the conditions

$$\phi := \int_M x\,\psi(x)\,d\rho(x) \;\in\; \mathcal{H} \qquad \text{and} \qquad |||\,\phi\,||| < \infty\,.$$

This follows immediately from the calculation

$$\begin{aligned} <\psi|(-P)\psi> &= -\iint_{M\times M} \prec\psi(x)\,|\,P(x,y)\,\psi(y)\succ_x d\rho(x)\,d\rho(y) \\ &= \iint_{M\times M} \langle\psi(x)\,|\,x\,\pi_x\,y\,\psi(y)\rangle_{\mathcal{H}}\,d\rho(x)\,d\rho(y) = \langle\phi|\phi\rangle_{\mathcal{H}} \geq 0\,, \end{aligned}$$

where we again used (1.22) and (1.11) and set

$$\phi = \int_M x\,\psi(x)\,d\rho(x)\,.$$

1.6. **Geometric Structures.** A causal fermion system also encodes geometric information on space-time. More specifically, in the paper [15] notions of connection and curvature are introduced and analyzed. We now outline a few constructions from this paper. Recall that the kernel of the fermionic projector (1.11) is a mapping from one spin space to another, thereby inducing relations between different space-time points. The idea is to use these relations for the construction of a spin connection $D_{x,y}$, being a unitary mapping between the corresponding spin spaces,

$$D_{x,y} \,:\, S_y \to S_x$$

(we consistently use the notation that the subscript ${}_{xy}$ denotes an object at the point x, whereas the additional comma ${}_{x,y}$ denotes an operator which maps an object at y to an object at x). The simplest method for constructing the spin connection would be to form a polar decomposition, $P(x,y) = A_{xy}^{-\frac{1}{2}}\,U$, and to introduce the spin connection as the unitary part, $D_{x,y} = U$. However, this method is too naive, because we want the spin connection to be compatible with a corresponding metric connection $\nabla_{x,y}$ which should map Clifford subspaces at x and y (see Definition 1.4 above) isometrically to each other. A complication is that, as discussed at the end of §1.4, the Clifford subspaces at x and y are not unique. The method to bypass these problems is to work with several Clifford subspaces and to use so-called splice maps, as we now briefly explain.

First, it is useful to restrict the freedom in choosing the Clifford subspaces with the following construction. Recall that for any $x \in M$, the operator $(-x)$ on $\mathcal{H}$ has at most n positive and at most n negative eigenvalues. We denote its positive and negative spectral subspaces by S_x^+ and S_x^-, respectively. In view of (1.14), these subspaces are also orthogonal with respect to the spin scalar product,

$$S_x = S_x^+ \oplus S_x^-\,.$$

We introduce the *Euclidean sign operator* s_x as a symmetric operator on S_x whose eigenspaces corresponding to the eigenvalues ± 1 are the spaces S_x^+ and S_x^-, respectively. Since $s_x^2 = \mathbb{1}$, the span of the Euclidean sign operator is a one-dimensional Clifford subspace of signature $(1,0)$. The idea is to extend s_x to obtain higher-dimensional Clifford subspaces. We thus define a *Clifford extension* as a Clifford subspace which contains s_x. By restricting attention to Clifford extensions, we have reduced the freedom in choosing Clifford subspaces. However, there is still not a unique Clifford extension, even for fixed dimension and signature. But one can define the *tangent space* T_x as an equivalence class of Clifford extensions; for details see [15, Section 3.1]. The bilinear form $\langle .,. \rangle$ in (1.21) induces a Lorentzian metric on the tangent space.

Next, for our constructions to work, we need to assume that the points x and y are both regular and are properly timelike separated, defined as follows:

Definition 1.5. A space-time point $x \in M$ is said to be *regular* if x has the maximal possible rank, i.e. $\dim x(\mathcal{H}) = 2n$. Otherwise, the space-time point is called *singular*.

In most situations of physical interest (like Dirac see configurations to be discussed in Sections 2 and 5 below), all space-time points are regular. Singular points, on the other hand, should be regarded as exceptional points or "singularities" of space-time.

Definition 1.6. The space-time points $x, y \in M$ are *properly timelike* separated if the closed chain A_{xy}, (1.12), has a strictly positive spectrum and if all eigenspaces are definite subspaces of $(S_x, \prec.|.\succ_x)$.

By a definite subspace of S_x we mean a subspace on which the inner product $\prec.|.\succ_x$ is either positive or negative definite.

The two following observations explain why the last definition makes sense:

- Properly timelike separation implies timelike separation (see Definition 1.2):
  Before entering the proof, we give a simple counter example which shows why the assumption of definite eigenspaces in Definition 1.6 is necessary for the implication to hold. Namely, if the point x is regular and A_{xy} is the identity, then the eigenvalues $\lambda_1, \ldots, \lambda_{2n}$ are all strictly positive, but they are all equal.

  If $I \subset S_x$ is a definite invariant subspace of A_{xy}, then the restriction $A_{xy}|_I$ is a symmetric operator on the Hilbert space $(I, \pm\prec.|.\succ_{I\times I})$, which is diagonalizable with real eigenvalues. Moreover, the orthogonal complement $I^\perp$ of $I \subset S_x$ is again invariant. If $I^\perp$ is non-trivial, the restriction $A_{xy}|_{I^\perp}$ has at least one eigenspace. Therefore, the assumption in Definition 1.6 that all eigenspaces are definite makes it possible to proceed inductively to conclude that the operator A_{xy} is diagonalizable and has real eigenvalues.

  If x and y are properly timelike separated, then its eigenvalues are by definition all real and positive. Thus it remains to show that they are not all the same. If conversely they were all the same, i.e. $\lambda_1 = \cdots = \lambda_{2n} = \lambda > 0$, then S_x would necessarily have the maximal dimension $2n$. Moreover, the fact that A_{xy} is diagonalizable implies that A_{xy} would be a multiple of the identity on S_x. Therefore, the spin space $(S_x, \prec.|.\succ)$ would have to be definite, in contradiction to the fact that it has signature (n, n).
- The notion is symmetric in x and y:
  Suppose that $A_{xy}u = \lambda u$ with $u \in S_x$ and $\lambda \in \mathbb{R} \setminus \{0\}$. Then the vector $w := P(y,x)\, u \in S_y$ is an eigenvector of A_{yx} again to the eigenvalue λ,

$$\begin{aligned} A_{yx}\, w &= P(y,x) P(x,y)\, P(y,x)\, u \\ &= P(y,x)\, A_{xy}\, u = \lambda\, P(y,x)\, u = \lambda w\,. \end{aligned}$$

  Moreover, the calculation

$$\begin{aligned} \lambda \prec u|u\succ &= \prec u|A_{xy}u\succ = \prec u \,|\, P(x,y)\, P(y,x)\, u\succ \\ &= \prec P(y,x)u \,|\, P(y,x)u\succ = \prec w|w\succ \end{aligned}$$

  shows that w is a definite vector if and only if u is. We conclude that A_{yx} has positive eigenvalues and definite eigenspaces if and only if A_{xy} has these properties.

So far, the construction of the spin connection has been worked out only in the case of spin dimension $n = 2$. Then for two regular and properly timelike separated points $x, y \in M$, the spin space S_x can be decomposed uniquely into an orthogonal direct sum $S_x = I^+ \oplus I^-$ of a two-dimensional positive definite subspace I^+ and a two-dimensional negative definite subspace I^- of A_{xy}. We define the *directional sign operator* v_{xy} of A_{xy} as the unique operator with eigenvalues $-1, 1, 0$ such that the eigenspaces corresponding to the eigenvalues ± 1 are the subspaces $I^\pm$.

Having the Euclidean sign operator s_x and the directional sign operator v_{xy} to our disposal, under generic assumptions one can distinguish two Clifford subspaces at the point x: a Clifford subspace K_{xy} containing v_{xy} and a Clifford extension $K_x^{(y)}$ (for details see [15, Lemma 3.12]). Similarly, at the point y we have a distinguished Clifford subspace K_{yx} (which contains v_{yx}) and a distinguished Clifford extension $K_y^{(x)}$. For the construction of the *spin connection* $D_{x,y} : S_y \to S_x$ one works with the Clifford subspaces K_{xy} and K_{yx} and demands that these are mapped to each other. More precisely, the spin connection is uniquely characterized by the following properties (see [15, Theorem 3.20]):

(i) $D_{x,y}$ is of the form

$$D_{x,y} = e^{i\varphi_{xy}\, v_{xy}}\, A_{xy}^{-\frac{1}{2}}\, P(x,y) \quad \text{with} \quad \varphi_{xy} \in (-\frac{3\pi}{4}, -\frac{\pi}{2}) \cup (\frac{\pi}{2}, \frac{3\pi}{4})\,.$$

(ii) The spin connection maps the Clifford subspaces K_{xy} and K_{yx} to each other, i.e.

$$D_{y,x}\, K_{xy}\, D_{x,y} = K_{yx}\,.$$

The spin connection has the properties

$$D_{y,x} = (D_{x,y})^{-1} = (D_{x,y})^* \qquad \text{and} \qquad A_{xy} = D_{x,y}\, A_{yx}\, D_{y,x}\,.$$

All the assumptions needed for the construction of the spin connection are combined in the notion that x and y must be *spin-connectable* (see [15, Definition 3.17]). We remark that in the limiting case of a Lorentzian manifold, the points x and y are spin-connectable if they are timelike separated and sufficiently close to each other (see [15, Section 5]).

By composing the spin connection along a discrete "path" of space-time points, one obtains a "parallel transport" of spinors. When doing so, it is important to keep track of the different Clifford subspaces and to carefully transform them to each other. In order to illustrate in an example how this works, suppose that we want to compose the spin connection $D_{y,z}$ with $D_{z,x}$. As mentioned above, the spin connection $D_{z,x}$ at the point z is constructed using the Clifford subspace K_{zx}. The spin connection $D_{y,z}$, however, takes at the same space-time point z the Clifford subspace K_{zy} as reference. This entails that before applying $D_{y,z}$ we must transform from the Clifford subspace K_{zx} to the Clifford subspace K_{zy}. This is accomplished by the *splice map* $U_z^{(y|x)}$, being a uniquely defined unitary transformation of S_x with the property that

$$K_{zy} = U_z^{(y|x)}\, K_{zx}\, \big(U_z^{(y|x)}\big)^*\,.$$

The splice map must be sandwiched between the spin connections in combinations like

$$D_{y,z}\, U_z^{(y|x)}\, D_{z,x}\,.$$

In order to construct a corresponding metric connection $\nabla_{x,y}$, one uses a similar procedure to related the Clifford subspaces to corresponding Clifford extensions. More

precisely, one first unitarily transform the Clifford extension $K_y^{(x)}$ to the Clifford subspace K_{yx}. Unitarily transforming with the spin connection D_{xy} gives the Clifford subspace K_{xy}. Finally, one unitarily transforms to the Clifford extension $K_x^{(y)}$. Since the Clifford extensions at the beginning and end are representatives of the corresponding tangent spaces, we thus obtain an isometry

$$\nabla_{x,y} \,:\, T_y \to T_x$$

between the tangent spaces (for details see [15, Section 3.4]).

In this setting, *curvature* is defined as usual as the holonomy of the connection. Thus the curvature of the spin connection is given by

$$\mathfrak{R}(x,y,z) = U_x^{(z|y)}\, D_{x,y}\, U_y^{(x|z)}\, D_{y,z}\, U_z^{(y|x)}\, D_{z,x} \,:\, S_x \to S_x\,,$$

and similarly for the metric connection. In [15, Sections 4 and 5] it is proven that the above notions in fact reduce to the spinorial Levi-Civita connection and the Riemannian curvature on a globally hyperbolic Lorentzian manifold if the causal fermion system is constructed by regularizing solutions of the Dirac equation (similar as will explained in the next section for the Minkowski vacuum) and the regularization is suitably removed. These results show that the notions of connection and curvature defined above indeed generalize the corresponding notions in Lorentzian spin geometry.

1.7. **Topological Structures.** From a mathematical perspective, causal fermion systems provide a framework for non-smooth geometries or generalized "quantum geometries." In this context, it is of interest how the topological notions on a differentiable manifold or a spin manifold generalize to causal fermion systems. Such topological questions are analyzed in [18], as we now briefly summarize.

By definition, space-time M is a topological space (see §1.2). Attaching to every space-time point $x \in M$ the corresponding spin space S_x gives the structure of a *sheaf*, making it possible to describe the topology by sheaf cohomology. If one assumes in addition that all space-time points are regular (see Definition 1.5), then all spin spaces are isomorphic, giving rise to a *topological vector bundle*.

In order to get the connection to spinor bundles, one needs the additional structure of Clifford multiplication. As explained in §1.4, the notion of a Clifford subspace (see Definition 1.4) makes it possible to define Clifford structures at every space-time point, but the definition is not unique and does not give the connection to tangent vectors of the base space. In §1.6 these shortcomings where bypassed by working with suitable equivalence classes of Clifford subspaces. From the topological point of view, the basic question is whether one can choose a representative of this equivalence class at each space-time point in such a way that the representative depends continuously on the base point. This leads to the notion of a *Clifford section* $\mathcal{C}\ell$, being a continuous mapping which to every space-time point $x \in M$ associates a corresponding Clifford subspace $\mathcal{C}\ell_x$ (for details see [18, Section 4.1]). Choosing a Clifford section leads to the structure of a so-called *topological spinor bundle*. An advantage of working with topological spinor bundles is that no notion of differentiability is required.

If M has a differentiable structure, one would like to associate a tangent vector $u \in T_xM$ to a corresponding element of the Clifford subspace $\mathcal{C}\ell_x$. This leads to the notion of a *spin structure* γ on a topological spinor bundle, being a continuous mapping which to every $x \in M$ associates a mapping $\gamma_x : T_xM \to \mathcal{C}\ell_x$. The topological obstructions for the existence of a spin structure on a topological spinor bundle generalize the spin condition on a spin manifold (for details see [18, Sections 4.2 and 4.5]).

A useful analytic tool for the construction of Clifford sections are so-called *tangent cone measures* (see [18, Section 5]). These measures make it possible to analyze the local structure of space-time in a neighborhood of a point $x \in M$ (again without any differentiability assumptions). The tangent cone measures can be used to distinguish a specific Clifford subspace $\mathcal{C}\ell_x$ and to relate $\mathcal{C}\ell_x$ to neighboring space-time points.

We close with two remarks. First, all the above constructions generalize to the *Riemannian setting* if the definition of causal fermion systems is extended to so-called *topological fermion systems* (see [18, Definition 2.1]). We thus obtain a mathematical framework to describe *spinors on singular spaces* (see [18, Sections 7 and 8] for many examples). Second, one can introduce nontrivial topological notions even for discrete space-times by constructing neighborhoods of M in $\mathcal{F}$ (using the metric structure of $\mathcal{F}$ induced by the norm on the Banach space $\mathrm{L}(\mathcal{H})$) and by studying the topology of these neighborhoods.

2. Correspondence to Minkowski Space

In order to put the abstract framework in a simple and concrete context, we now explain how to describe Dirac spinors in Minkowski space as a causal fermion system.

2.1. **Concepts Behind the Construction of Causal Fermion Systems.** We let $(\mathscr{M}, \langle .,. \rangle)$ be Minkowski space (with the signature convention $(+---)$) and $d\mu$ the standard volume measure (thus $d\mu = d^4x$ in a reference frame $x = (x^0, \ldots, x^3)$). We denote the spinor space at a point $x \in \mathscr{M}$ by $S_x\mathscr{M}$, so that a Dirac 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 as in physics textbooks we denote by $\overline{\psi}\phi$ (where $\overline{\psi} = \psi^\dagger \gamma^0$ is the usual adjoint spinor). Clearly, in Minkowski space one has a trivial parallel transport of spinors, making it possible to identify the spinor spaces at different space-time points. Thus the space-time index $S_x\mathscr{M}$ of the spinor space is added only for notational clarity.

On the solutions of the Dirac equation

$$(i\gamma^j \partial_j - m)\psi = 0 \tag{2.1}$$

we consider the usual Lorentz invariant scalar product

$$(\psi|\phi) := 2\pi \int_{\mathbb{R}^3} (\overline{\psi}\gamma^0 \phi)(t, \vec{x})\, d^3x \,, \tag{2.2}$$

making the solution space to a separable Hilbert space. We choose $\mathcal{H}$ as a closed subspace of this Hilbert space with the induced scalar product $\langle .|. \rangle_{\mathcal{H}} := (.|.)|_{\mathcal{H}\times\mathcal{H}}$. Clearly, $\mathcal{H}$ is again a separable Hilbert space. In order to describe the vacuum (i.e. the physical system where no particles and anti-particles are present), one chooses $\mathcal{H}$ as the subspace spanned by all the negative-energy solutions (the "Dirac sea vacuum"). To describe particles or anti-particles, one includes positive-energy solutions or leaves out negative-energy solutions, respectively. But any other closed subspace of the solution space may be chosen as well. We remark for clarity that in this section, we only consider the vacuum Dirac equation (2.1), so that the Dirac particles do not interact (interacting systems will be discussed in Section 5 below).

In order to get into the framework of causal fermion systems, to every space-time point $x \in \mathscr{M}$ we want to associate a linear operator $F(x) \in \mathcal{F}$. Once this has been

accomplished, the resulting mapping

$$F \,:\, \mathscr{M} \to \mathcal{F}\,. \tag{2.3}$$

can be used to introduce a measure ρ on $\mathcal{F}$. Namely, we say that a subset $\Omega \subset \mathcal{F}$ is measurable if and only if its pre-image $F^{-1}(\Omega)$ is a measurable subset of $\mathscr{M}$. Moreover, we define the measure of Ω as the space-time volume of the pre-image, $\rho(\Omega) := \mu(F^{-1}(\Omega))$. This construction is commonly used in mathematical analysis and is referred to as the *push-forward measure*, denoted by

$$\rho = F_* \mu\,.$$

Then $(\mathcal{H}, \mathcal{F}, \rho)$ will be a causal fermion system.

The basic idea for constructing $F(x)$ is to represent the inner product on the spinors in terms of the Hilbert space scalar product, i.e.

$$\langle \psi | F(x) \phi \rangle_{\mathcal{H}} = -(\overline{\psi}\phi)(x) \qquad \text{for all } \psi, \phi \in \mathcal{H}\,. \tag{2.4}$$

The operator $F(x)$ gives information on the densities and correlations of the Dirac wave functions at the space-time point x. It is referred to as the *local correlation operator* at x. Relating the maximal number of positive and negative eigenvalues of $F(x)$ to the signature of the inner product $(\overline{\psi}\phi)(x)$, one sees that $F(x)$ indeed has at most two positive and at most two negative eigenvalues. However, the equation (2.4) suffers from the shortcoming that the right side is in general ill-defined because solutions $\psi, \phi \in \mathcal{H}$ are in general not continuous and thus cannot be evaluated pointwise. This is the reason why we need to introduce an *ultraviolet regularization* (UV regularization). Before entering the analysis, we first outline our method and explain the physical picture in a few remarks. The mathematical construction will be given afterwards in §2.2.

In order to put our constructions in the general physical context, we first note that UV regularizations are frequently used in relativistic quantum field theory as a technical tool to remove divergences. A common view is that the appearance of such divergences indicates that the physical theory is incomplete and should be replaced for very small distances by another, more fundamental theory. The renormalization program is a method to get along with standard quantum field theory by finding a way of dealing with the divergences. The first step is the UV regularization, which is usually a set of prescriptions which make divergent integrals finite. The next step of the renormalization program is to show that the UV regularization can be taken out if other parameters of the theory (like masses and coupling constants) are suitably rescaled. Conceptually, in the renormalization program the UV regularization merely is a technical tool. All predictions of theory should be independent of how the regularization is carried out.

In the context of causal fermion systems, however, the physical picture behind the UV regularization is quite different. Namely, in our setting the *regularized* objects are to be considered as the fundamental physical objects. Therefore, the regularization has a physical significance. It should describe the microscopic structure of physical space-time.

Before explaining this physical picture in more detail, we need to introduce a microscopic length scale $\varepsilon > 0$ on which the UV regularization should come into play. Regularization lengths are often associated to the Planck length $\ell_P \approx 1.6 \cdot 10^{-35}$ m. The analysis of the gravitational field in [6] suggests that ε should be chosen even much smaller than the Planck length (see [6, Section 4.9 and §5.4.3]). Even without entering

a detailed discussion of the length scales, it is clear that ε will be by many orders of magnitude smaller than most other physical length scales of the system. Therefore, it is a sensible method to analyze the causal action principle in the asymptotics when ε is very small. In order to make such an asymptotics mathematically precise, we necessarily need to consider the *regularization length* ε as a *variable parameter* taking values in an interval $(0, \varepsilon_{\max})$. Only for such a variable parameter, one can analyze the asymptotics as $\varepsilon \searrow 0$.

For any $\varepsilon \in (0, \varepsilon_{\max})$, similar to (2.3) we shall construct a mapping $F^\varepsilon : \mathscr{M} \to \mathcal{F}$ by suitably inserting an UV regularization in (2.4). Then we construct the corresponding universal measure as the push-forward by F^ε, i.e.

$$\rho^\varepsilon := F^\varepsilon_* \mu \,. \tag{2.5}$$

This shall give rise to a causal fermion system $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$. We will also explain how to identify the objects in Minkowski space with corresponding objects of the causal fermion system:

Minkowski space	**causal fermion system**
space-time point $x \in \mathscr{M}$	space-time point $x \in M^\varepsilon := \operatorname{supp} \rho^\varepsilon$
topology of $\mathscr{M}$	topology of M^ε
spinor space $S_x\mathscr{M}$	spin space $S_x M^\varepsilon$
causal structure of Minkowski space	causal structure of Definition 1.2

With these identifications made, the structures of Minkowski space are no longer needed. They are encoded in the causal fermion system, and we may describe the physical space-time exclusively by the causal fermion system. We consider the objects with UV regularization as described by the causal fermion system as the fundamental physical objects.

In the following remarks we elaborate on the physical picture behind the UV regularization and explain why our setting is sufficiently general to describe the physical situation we have in mind.

Remark 2.1. (method of variable regularization) As just explained, the only reason for considering a family of causal fermion systems is to give the asymptotics $\varepsilon \searrow 0$ a precise mathematical meaning. But from the physical point of view, a specific regularization for a specific value of ε should be distinguished by the fact that the corresponding causal fermion system $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$ describes our physical space-time. We again point out that this concept is different from standard quantum field theory, where the regularization merely is a technical tool used in order to remove divergences. In our setting, the regularization has a physical significance. The *regularized* objects are to be considered as the *fundamental* physical objects, and the regularization is a method to describe the microscopic structure of physical space-time.

This concept immediately raises the question how the "physical regularization" should look like. Generally speaking, the regularized space-time should look like Minkowski space down to distances of the scale ε. For distances smaller than ε, the structure of space-time may be completely different. The simplest method of regularizing is to "smear out" or "mollify" all wave functions on the scale ε (this corresponds to Example 2.4 below). But it is also conceivable that space-time has a non-trivial

microstructure on the scale ε, which cannot be guessed or extrapolated from the structures of Minkowski space. Since experiments on the length scale ε seem out of reach, it is completely unknown what the microscopic structure of space-time is. Nevertheless, we can hope that we can get along without knowing this micro-structure, because the detailed form of this micro-structure might have no influence on the effective physical equations which are valid on the energy scales accessible to experiments. More precisely, the picture is that the general structure of the effective physical equations should be independent of the micro-structure of space-time. Values of mass ratios or coupling constants, however, may well depend on the micro-structure (a typical example is the gravitational constant, which is closely tied to the Planck length, which in turn is related to ε as explained in [6, Section 4.9]). In more general terms, the unknown micro-structure of space-time should enter the effective physical equations only by a finite (hopefully small) number of free parameters, which can then be taken as empirical free parameters of the effective macroscopic theory.

Clearly, the above picture must be questioned and supported by mathematical results. To this end, one needs to analyze in detail how the effective macroscopic theory depends on the regularization. For this reason, it is not sufficient to consider a specific family of regularizations. Instead, one must analyze a whole class of regularizations which is so large that it covers all relevant regularization effects. This strategy is referred to as the *method of variable regularization* (for a longer explanation see [8, §4.1]). It is the reason why in Definition 2.3 below we shall only state properties of the regularization, but we do not specify how precisely it should look like. ♢

Remark 2.2. (sequences of finite-dimensional regularizations) The critical reader may wonder why we consider a family of regularizations $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$ parametrized by a continuous parameter $(0, \varepsilon_{\max})$. Would it not be more suitable to consider instead a sequence of causal fermion systems $(\mathcal{H}_\ell, \mathcal{F}_\ell, \rho_\ell)$ which asymptotically as $\ell \to \infty$ describes Minkowski space? A related question is why we constructed the measure ρ as the push-forward of the Lebesgue measure (2.5). Would it not be better to work with more general measures such as to allow for the possibility of discrete micro-structures? The answer to these questions is that it is no loss of generality and a simply a matter of convenience to work with the family $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$ with $\varepsilon \in (0, \varepsilon_{\max})$, as we now explain.

We first point out that we do not demand our family $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$ to be in any sense "continuous" in the parameter ε. Therefore, one can also describe a sequence $(\mathcal{H}, \mathcal{F}, \rho_\ell)$ simply by choosing the family ρ^ε to be piecewise constant, for example

$$\rho^\varepsilon = \rho_\ell \qquad \text{if} \qquad \frac{1}{\ell} \leq \varepsilon < \frac{1}{\ell+1}\,.$$

Similarly, it is no loss of generality to take ρ as the push-forward measure of the Lebesgue measure because $F^\varepsilon(x)$ need not depend continuously on $x \in M$. For example, one can arrange a discrete space-time like a space-time lattice by choosing F^ε as a mapping which is piecewise constant on little cubes of Minkowski space. Clearly, this mapping is not continuous, but it is continuous almost everywhere. Moreover, its image is a discrete set, corresponding to a discrete micro-structure of space-time. For the method for representing a general measure ρ as the push-forward of for example the Lebesgue measure we refer the interested reader to the proof of [10, Lemma 1.4].

The remaining question is why we may keep the Hilbert space $\mathcal{H}$ fixed. In particular, we noted in §1.1 that the existence of minimizers of the causal action principle has been proven only if $\mathcal{H}$ is finite-dimensional. Therefore, should one not consider

a filtration $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \cdots \subset \mathcal{H}$ of $\mathcal{H}$ by finite-dimensional subspaces? Indeed, from the conceptual point of view, this would be the correct way to proceed. Nevertheless, the following consideration explains why we can just as well replace all the Hilbert spaces $\mathcal{H}_\ell$ by the larger space $\mathcal{H}$: For a given causal fermion system $(\mathcal{H}_\ell, \mathcal{F}_\ell, \rho_\ell)$ with $\mathcal{H}_\ell \subset \mathcal{H}$, by extending all operators by zero to the orthogonal complement of $\mathcal{H}_\ell$, one obtains the so-called *extended causal fermion system* $(\mathcal{H}, \mathcal{F}, \rho_\ell)$. The fact that the causal fermion system was extended can still be seen by forming the so-called *effective Hilbert space* as

$$\mathcal{H}^{\text{eff}} = \overline{\text{span}\{x(\mathcal{H}) \,|\, x \in \text{supp}\,\rho\}}\,.$$

Namely, for an extended causal fermion system, the effective Hilbert space still is a subset of the original Hilbert space, $\mathcal{H}^{\text{eff}} \subset \mathcal{H}_\ell$. Moreover, the support of the extended causal fermion system is still contained in $\mathcal{F}_\ell \subset \text{L}(\mathcal{H}_\ell)$. Therefore, we do not lose any information by extending a causal fermion system. Conversely, when analyzing a causal fermion system, it seems preferable to always make the Hilbert space as small as possible by taking $\mathcal{H}^{\text{eff}}$ as the underlying Hilbert space.

The delicate point about extending causal fermion systems is that the causal action principle does depend sensitively on the dimension of the underlying Hilbert space $\mathcal{H}$. More specifically, the infimum of the action is known to be strictly decreasing in the dimension of $\mathcal{H}$ (see the estimates in [9, Lemma 5.1], which apply similarly in the more general setting of [10]). Therefore, a minimizer ρ of the causal action principle will no longer be a minimizer if the causal fermion system is extended. However, the first order *Euler-Lagrange equations* (for details see §4.1 below) are still satisfied for the extended causal fermion system. Therefore, for convenience we fix the Hilbert space $\mathcal{H}$ and consider a family of causal fermion systems $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$ thereon. In order for the causal action principle to be well-defined and for ρ^ε to be a minimizer, one should replace $\mathcal{H}$ by the corresponding effective Hilbert space $\mathcal{H}^{\text{eff}}$, which may depend on ε and should be arranged to be finite-dimensional. For the analysis of the Euler-Lagrange equations, however, the restriction to $\mathcal{H}^{\text{eff}}$ is unnecessary, and it is preferable to work with the extended Hilbert space $\mathcal{H}$. ♢

We finally remark that the hurried reader who wants to skip the following constructions may read instead the introductory section [17, Section 1.1] where formal considerations without UV regularization are given. Moreover, a more explicit analysis of four-dimensional Minkowski space with a particularly convenient regularization is presented in [15, Section 4]. For a somewhat simpler analysis of two-dimensional Minkowski space we refer to [18, Section 7.2].

2.2. **Introducing an Ultraviolet Regularization.** We now enter the construction of the UV regularization. We denote the continuous Dirac wave functions (i.e. the continuous sections of the spinor bundle, not necessarily solutions of the Dirac equation) by $C^0(\mathscr{M}, S\mathscr{M})$. Similarly, the smooth wave functions with compact support in a subset $K \subset \mathscr{M}$ are denoted by $C_0^\infty(K, S\mathscr{M})$. For the C^k-norms we use the notation

$$|\eta|_{C^k(K)} = \sum_{|\alpha| \leq k} \sup_{x \in K} |\partial^\alpha \eta(x)| \qquad \text{for } \eta \in C_0^\infty(K, S\mathscr{M})\,,$$

where the α are multi-indices. Here $|.|$ is any pointwise norm on the spinor spaces (we again identify all spinor spaces with the trivial parallel transport). Since any two such norms can be estimated from above and below by a constant, the C^k-norms corresponding to different choices of the norms $|.|$ are also equivalent. For example,

one can choose $|\psi|^2 := \overline{\psi}\gamma^0\psi$ similar to the integrand in the scalar product (2.2). But clearly, other choices are possible just as well.

The UV regularization is performed most conveniently with so-called regularization operators, which we now define.

Definition 2.3. *Consider a family of linear operators $(\mathfrak{R}_\varepsilon)$ with $0 < \varepsilon < \varepsilon_{\max}$ which map $\mathcal{H}$ to the continuous wave functions,*

$$\mathfrak{R}_\varepsilon \,:\, \mathcal{H} \to C^0(\mathcal{M}, S\mathcal{M})\,.$$

The family is called a family of **regularization operators** *if the following conditions hold:*

(i) The image of every regularization operator is pointwise bounded, meaning that for every $\varepsilon \in (0, \varepsilon_{\max})$ and all $x \in \mathcal{M}$ there is a constant $c > 0$ such that for all $u \in \mathcal{H}$,

$$\big|\big(\mathfrak{R}_\varepsilon u\big)(x)\big| \leq c\, \|u\|_{\mathcal{H}}\;. \tag{2.6}$$

(ii) The image of every regularization operator is equicontinuous almost everywhere in the sense that for every $\varepsilon \in (0, \varepsilon_{\max})$, almost all $x \in \mathcal{M}$ and every $\delta > 0$, there is an open neighborhood $U \subset \mathcal{M}$ of x such that for all $u \in \mathcal{H}$ and all $y \in U$,

$$\big|\big(\mathfrak{R}_\varepsilon u\big)(x) - \big(\mathfrak{R}_\varepsilon u\big)(y)\big| \leq \delta\, \|u\|_{\mathcal{H}}\;. \tag{2.7}$$

(iii) In the limit $\varepsilon \searrow 0$, the family converges weakly to the identity, meaning that for every compact subset $K \subset \mathcal{M}$ and every $\delta > 0$ there is a constant $\varepsilon_0 > 0$, such that for all $\varepsilon \in (0, \varepsilon_0)$, $u \in \mathcal{H}$ and $\eta \in C_0^\infty(K, S\mathcal{M})$,

$$\Big| \int_{\mathcal{M}} \overline{\eta(x)}\big(\mathfrak{R}_\varepsilon(u) - u\big)(x)\, d^4x \Big| \leq \delta\, \|u\|_{\mathcal{H}}\, |\eta|_{C^1(K)}\,. \tag{2.8}$$

We point out that we do not demand that the regularized wave function $\mathfrak{R}_\varepsilon\psi$ is again a solution of the Dirac equation. This could be imposed (as is done in [25, Section 4]), but doing so seems too restrictive for the physical applications. We also note that "almost all" in (ii) refers to the standard volume measure $d\mu$ on $\mathcal{M}$.

For the mathematically interested reader we remark that the above properties (i) and (ii) are very similar to the assumptions in the Arzelà-Ascoli theorem (see for example [5, Section VII.5] or [35, Theorem 7.25]). In fact, if we replaced "almost all" in (ii) by "all", one could apply the Arzelà-Ascoli theorem and restate the properties (i) and (ii) equivalently by saying that taking the image $\mathfrak{R}_\varepsilon(B_1(0))$ of the unit ball in $\mathcal{H}$ and restricting the resulting family of functions to any compact set $K \subset \mathcal{M}$, one obtains a relatively compact subset of $C^0(K, S\mathcal{M})$. It is remarkable that the properties (i) and (ii) come up naturally as conditions for a sensible UV regularization, although we shall never use compactness arguments in our proofs. Weakening "all" by "almost all" in (ii) makes it possible to describe discrete space-times like space-time lattices, as was mentioned in Remark 2.2 above.

Simple examples of regularization operators are obtained by mollifying the wave functions on the scale ε:

Example 2.4. (regularization by mollification) Let $h \in C_0^\infty(\mathcal{M}, \mathbb{R})$ be a non-negative test function with

$$\int_{\mathcal{M}} h(x)\, d^4x = 1\,.$$

We define the operators $\mathfrak{R}_\varepsilon$ for $\varepsilon > 0$ as the convolution operators

$$(\mathfrak{R}_\varepsilon u)(x) := \frac{1}{\varepsilon^4} \int_{\mathscr{M}} h\Big(\frac{x-y}{\varepsilon}\Big)\, u(y)\, d^4y\,.$$

Let us prove that the family $(\mathfrak{R}_\varepsilon)_{0<\varepsilon<1}$ is a family of regularization operators. First,

$$\big|\big(\mathfrak{R}_\varepsilon u\big)(x)\big| \leq \frac{|h|_{C^0}}{\varepsilon^4} \int_K |u(y)|\, d^4y \leq \frac{|h|_{C^0}}{\varepsilon^4}\, \sqrt{\mu(K)}\, \Big(\int_K |u(y)|^2\, d^4y \Big)^{\frac{1}{2}}\,,$$

where in the last step we used the Schwarz inequality. We now rewrite the obtained space-time integral of $|u|^2$ with the help of Fubini's theorem as a bounded time integral and a spatial integral. In view of (2.2), the spatial integral can be estimated by the Hilbert space norm. We thus obtain

$$\int_K |u(y)|^2\, d^4y \leq C \int_K \big(\overline{u} \gamma^0 u\big)(y)\, d^4y \leq C \int_{t_0}^{t_1} \|u\|^2_{\mathscr{H}} = C\,(t_1 - t_0)\, \|u\|^2_{\mathscr{H}}\,, \tag{2.9}$$

where t_0 and t_1 are chosen such that K is contained in the time strip $t_0 < t < t_1$. We conclude that

$$\big|\big(\mathfrak{R}_\varepsilon u\big)\big| \leq \frac{|h|_{C^0}}{\varepsilon^4}\, \sqrt{\mu(K)\, C\,(t_1 - t_0)}\, \|u\|^2_{\mathscr{H}}\,,$$

proving (2.6).

In order to derive the inequality (2.7), we begin with the estimate

$$\big|\big(\mathfrak{R}_\varepsilon u\big)(x) - \big(\mathfrak{R}_\varepsilon u\big)(y)\big| \leq \frac{1}{\varepsilon^4} \sup_{z \in \mathscr{M}} \Big| h\Big(\frac{x-z}{\varepsilon}\Big) - h\Big(\frac{y-z}{\varepsilon}\Big)\Big| \int_K |u(y)|\, d^4y\,.$$

Again applying (2.9) and using that h is uniformly continuous, one obtains (2.7).

It remains to prove (2.8). We first write the integral on the left as

$$\int_{\mathscr{M}} \overline{\eta(x)} \big(\mathfrak{R}_\varepsilon(u) - u\big)(x)\, d^4x = \int_{\mathscr{M}} \overline{\big(\eta_\varepsilon(y) - \eta(y)\big)}\, u(y)\, d^4y\,, \tag{2.10}$$

where we set

$$\eta_\varepsilon(y) = \frac{1}{\varepsilon^4} \int_{\mathscr{M}} \eta(x)\, h\Big(\frac{x-y}{\varepsilon}\Big)\, d^4x\,.$$

Now we use the standard estimate for convolutions

$$\begin{aligned} |\eta_\varepsilon(y) - \eta(y)| &= \frac{1}{\varepsilon^4} \Big| \int_{\mathscr{M}} \big(\eta(x) - \eta(y)\big)\, h\Big(\frac{x-y}{\varepsilon}\Big)\, d^4x \Big| \\ &= \Big| \int_{\mathscr{M}} \Big(\eta(y + \varepsilon z) - \eta(y)\Big)\, h(z)\, d^4z \Big| \leq |\eta|_{C^1(K)} \int_{\mathscr{M}} |\varepsilon z|\ h(z)\, d^4z \end{aligned}$$

(where in the last step we used the mean value theorem). This gives rise to the estimate

$$|\eta_\varepsilon - \eta|_{C^0(K)} \leq c\, \varepsilon\, |\eta|_{C^1(K)}\,,$$

where c may depend on K and the choice of h, but is independent of η. This makes it possible to estimate (2.10) by

$$\Big| \int_{\mathscr{M}} \overline{\eta(x)} \big(\mathfrak{R}_\varepsilon(u) - u\big)(x)\, d^4x \Big| \leq \varepsilon\, |\eta|_{C^1(K)} \int_K |u(y)|_y\, d^4y\,.$$

Again applying (2.9), we conclude that

$$\Big| \int_{\mathscr{M}} \overline{\eta(x)} \big(\mathfrak{R}_\varepsilon(u) - u\big)(x)\, d^4x \Big| \leq \delta\, |\eta|_{C^1(K)}\, \sqrt{\mu(K)}\, \sqrt{C\,(t_1 - t_0)}\, \|u\|_{\mathscr{H}}\,,$$

proving (2.8). $\lozenge$

Given a family of regularization operators, we can construct causal fermion systems as follows. We fix $\varepsilon \in (0, \varepsilon_{\max})$. For any $x \in \mathscr{M}$, we consider the bilinear form

$$b_x \,:\, \mathcal{H} \times \mathcal{H} \to \mathbb{C}\,, \quad b_x(u,v) = -\overline{(\mathfrak{R}_\varepsilon\, u)(x)}(\mathfrak{R}_\varepsilon\, v)(x)\,. \tag{2.11}$$

This bilinear form is well-defined and bounded because $\mathfrak{R}_\varepsilon$ is defined pointwise and because evaluation at x gives a linear operator of finite rank. Thus for any $v \in \mathcal{H}$, the anti-linear form $b_x(.,v) : \mathcal{H} \to \mathbb{C}$ is continuous. By the Fréchet-Riesz theorem (see for example [33, Section 6.3]), there is a unique vector $w \in \mathcal{H}$ such that $b_x(u,v) = \langle u|w\rangle_{\mathcal{H}}$ for all $u \in \mathcal{H}$. The mapping $v \mapsto w$ is linear and bounded. We thus obtain a bounded linear operator $F^\varepsilon(x)$ on $\mathcal{H}$ such that

$$b_x(u,v) = \langle u \,|\, F^\varepsilon(x)\, v\rangle_{\mathcal{H}} \qquad \text{for all } u, v \in \mathcal{H}\,.$$

Taking into account that the inner product on the Dirac spinors at x has signature $(2,2)$, the local correlation operator $F^\varepsilon(x)$ is a symmetric operator on $\mathcal{H}$ of rank at most four, which has at most two positive and at most two negative eigenvalues. Finally, we introduce the *universal measure* $\rho^\varepsilon = F^\varepsilon_* \mu$ as the push-forward of the volume measure on $\mathscr{M}$ under the mapping F^ε. In this way, for every $\varepsilon \in (0, \varepsilon_0)$ we obtain a causal fermion system $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$ of spin dimension $n = 2$.

2.3. **Correspondence of Space-Time.** We now explain the connection between points of Minkowski space and points of space-time $M^\varepsilon := \operatorname{supp} \rho^\varepsilon$ of the corresponding causal fermion system $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$. We begin with a general characterization of M^ε.

Proposition 2.5. *For any $\varepsilon \in (0, \varepsilon_{\max})$, there is a subset $E \subset \mathscr{M}$ of μ-measure zero such that the mapping $F^\varepsilon|_{\mathscr{M}\setminus E} : \mathscr{M} \setminus E \to \mathcal{F}$ is continuous. Moreover, the support of the universal measure $M^\varepsilon := \operatorname{supp} \rho^\varepsilon$ is given by*

$$M^\varepsilon = \overline{F^\varepsilon(\mathscr{M} \setminus E)}^{\mathrm{L}(\mathcal{H})}\,. \tag{2.12}$$

Proof. To show continuity, we need to to estimate the sup-norm $\|F^\varepsilon(x) - F^\varepsilon(y)\|$. We first write the expectation value of the corresponding operator by

$$\begin{aligned}
\langle u \,|\, \big(F^\varepsilon(x) - F^\varepsilon(y)\big)\, v\rangle_{\mathcal{H}} &= -\overline{(\mathfrak{R}_\varepsilon\, u)(x)}(\mathfrak{R}_\varepsilon\, v)(x) + \overline{(\mathfrak{R}_\varepsilon\, u)(y)}(\mathfrak{R}_\varepsilon\, v)(y) \\
&= -\overline{(\mathfrak{R}_\varepsilon\, u)(x)}\big((\mathfrak{R}_\varepsilon\, v)(x) - (\mathfrak{R}_\varepsilon\, v)(y)\big) - \overline{\big((\mathfrak{R}_\varepsilon\, u)(x) - (\mathfrak{R}_\varepsilon\, u)(y)\big)}(\mathfrak{R}_\varepsilon\, v)(y)\,,
\end{aligned}$$

giving rise to the estimate

$$\begin{aligned}
&\big|\langle u \,|\, \big(F^\varepsilon(x) - F^\varepsilon(y)\big) v\rangle_{\mathcal{H}}\big| \\
&\quad \leq |(\mathfrak{R}_\varepsilon\, u)(x)|\, \big|(\mathfrak{R}_\varepsilon\, v)(x) - \mathfrak{R}_\varepsilon\, v)(y)\big| + \big|(\mathfrak{R}_\varepsilon\, u)(x) - (\mathfrak{R}_\varepsilon\, u)(y)\big|\, |(\mathfrak{R}_\varepsilon\, v)(y)|\,.
\end{aligned}$$

We now estimate the resulting spinor norms with the help of properties (i) and (ii) of Definition 2.3. First, we denote the exceptional set of μ-measure zero where (2.7) does not hold by $E \subset \mathscr{M}$. Combining (2.6) and (2.7), one immediately sees that every point $x \in \mathscr{M} \setminus E$ has a neighborhood U such that the boundedness property (2.6) holds uniformly on U (i.e. $|(\mathfrak{R}_\varepsilon u)(y)| \leq c\, \|u\|_{\mathcal{H}}$ for all $y \in U$). We thus obtain the estimate

$$\big|\langle u \,|\, \big(F^\varepsilon(x) - F^\varepsilon(y)\big) v\rangle_{\mathcal{H}}\big| \leq 2c\,\delta\, \|u\|_{\mathcal{H}}\, \|v\|_{\mathcal{H}}\,,$$

valid for all $y \in U$ and $u, v \in \mathcal{H}$. Hence the sup-norm is bounded by $\|F^\varepsilon(x) - F^\varepsilon(y)\| \leq 2c\delta$, showing that F^ε is continuous on $\mathscr{M} \setminus E$.

It remains to prove (2.12). Since $\mu(E) = 0$, the set E can be disregarded when forming the push-forward measure. Therefore, taking into account that the support of a measure is by definition a closed set, it suffices to show that for every $x \in \mathscr{M} \setminus E$, the

operator $p := F^\varepsilon(x)$ lies in the support of ρ^ε. Let $U \subset \mathcal{F}$ be an open neighborhood of p. Then the continuity of F^ε at x implies that the preimage $(F^\varepsilon)^{-1}(U)$ is an open subset of $\mathscr{M}$. Hence the Lebesgue measure of this subset is non-zero, $\mu((F^\varepsilon)^{-1}(U)) > 0$. By definition of the push-forward measure, it follows that $\rho^\varepsilon(U) > 0$. Hence every neighborhood of p has a non-zero measure, implying that $p \in \operatorname{supp} \rho^\varepsilon$. This concludes the proof. □

In order to have a convenient notation, in what follows we always identify a point in Minkowski space with the corresponding operator of the causal fermion system,

$$\text{identify} \quad x \in \mathscr{M} \qquad \text{with} \qquad F^\varepsilon(x) \in \mathcal{F}\,. \tag{2.13}$$

In general, this identification is not one-to-one, because the mapping F^ε need not be injective. In the latter case, there are two points $x, y \in \mathscr{M}$ such that the bilinear forms b_x and b_y coincide (see (2.11)). In other words, all correlations between regularized wave functions coincide at the points x and y. Using a more physical language, this means that the points x, y of Minkowski space are not distinguishable by any experiments performed on the fermionic wave functions. We take the point of view that in such situations, the points x and y should not be distinguished physically, and that it is reasonable and desirable that the two points are identified in the causal fermion system with the same space-time point $F^\varepsilon(x) = F^\varepsilon(y) \in M^\varepsilon := \operatorname{supp} \rho^\varepsilon$. In philosophical terms, our construction realizes the principle of the identity of indiscernibles.

We also remark that, due to the closure in (2.12), it may happen that the space-time M^ε contains a point z which does *not* lie in the image of F^ε, but is merely a limit point in $F^\varepsilon(\mathscr{M})$. In this case, the corresponding bilinear form $b(u, v) := \langle u | z v \rangle_{\mathcal{H}}$ can be approximated with an arbitrarily small error by bilinear forms b_x with $x \in \mathscr{M}$. Since experiments always involve small imprecisions, we take the point of view that it is again reasonable and desirable mathematically to include z into the space-time points.

Generally speaking, the just-discussed cases that F^ε is not injective or its image is not closed seem mostly of academic interest. In most applications, the mapping F^ε will be injective and closed. In all these situations, Proposition 2.5 will give us a one-to-one correspondence between points $x \in \mathscr{M}$ and points $F^\varepsilon(x) \in M^\varepsilon$.

We finally note that, working with the push-forward measure (2.5), the volume measure on space-time M^ε as defined by the universal measure $d\rho^\varepsilon$ always agrees under the identification (2.13) with the Lebesgue measure $d\mu$ on $\mathscr{M}$.

2.4. **Correspondence of Spinors and Wave Functions.** We proceed by explaining the connection between the spinor space $S_x\mathscr{M}$ at a point $x \in \mathscr{M}$ of Minkowski space and the corresponding spin space $S_x M \subset \mathcal{H}$ of the causal fermion system (where we use the identification (2.13)). This will also make it possible to get a connection between Dirac wave functions in Minkowski space and wave functions as defined in §1.4. In preparation, we derive useful explicit formulas for the local correlation operators. To this end, for any $x \in \mathscr{M}$ we define the *evaluation map* e_x^ε by

$$e_x^\varepsilon \,:\, \mathcal{H} \to S_x\mathscr{M}\,, \qquad e_x^\varepsilon\, \psi = (\mathfrak{R}_\varepsilon \psi)(x)\,. \tag{2.14}$$

Its adjoint is defined as usual, taking into account the corresponding inner products on the domain and the target space, i.e.

$$\langle (e_x^\varepsilon)^* \chi \,|\, \psi \rangle_{\mathcal{H}} = \overline{\chi}\, \big(e_x^\varepsilon\, \psi\big)\,.$$

We denote this adjoint by ι_x^ε,

$$\iota_x^\varepsilon := (e_x^\varepsilon)^* \;:\; S_x\mathscr{M} \to \mathcal{H}\,.$$

Multiplying e_x^ε by ι_x^ε gives us back the local correlation operator $F^\varepsilon(x)$. Namely,

$$\langle \psi \,|\, F^\varepsilon(x)\, \phi\rangle_{\mathcal{H}} = -\overline{(\mathfrak{R}_\varepsilon\, \psi)(x)}(\mathfrak{R}_\varepsilon\, \phi)(x) = -\overline{\big(e_x^\varepsilon \psi\big)}\big(e_x^\varepsilon \phi\big) = -\langle \psi \,|\, \iota_x^\varepsilon e_x^\varepsilon\, \phi\rangle_{\mathcal{H}}$$

and thus

$$F^\varepsilon(x) = -\iota_x^\varepsilon\, e_x^\varepsilon = -\iota_x^\varepsilon \left(\iota_x^\varepsilon\right)^* \;:\; \mathcal{H} \to \mathcal{H}\,. \tag{2.15}$$

The next proposition gives the desired connection between the spinor space $S_x\mathscr{M}$ and the corresponding spin space S_xM. We first state and prove the proposition and explain it afterwards.

Proposition 2.6. *The mapping*

$$e_x^\varepsilon|_{S_x} \;:\; S_xM \to S_x\mathscr{M} \quad \textit{is an isometric embedding}\,.$$

Moreover, under this embedding, the physical wave function of a vector u at x is mapped to the regularized Dirac wave function at x,

$$e_x^\varepsilon|_{S_x}\, \psi^u(x) = \big(\mathfrak{R}_\varepsilon u\big)(x)\,. \tag{2.16}$$

If the point x is regular (see Definition 1.5), the inverse is given by

$$\big(e_x^\varepsilon|_{S_x}\big)^{-1} = -\big(x|_{S_x}\big)^{-1} \iota_x^\varepsilon \;:\; S_x\mathscr{M} \to S_xM\,. \tag{2.17}$$

Proof. Let $\psi, \phi \in S_xM$. Then

$$\overline{\big(e_x^\varepsilon \psi\big)}\big(e_x^\varepsilon \phi\big) = \langle \psi \,|\, (e_x^\varepsilon)^*\, e_x^\varepsilon\, \phi\rangle_{\mathcal{H}} = \langle \psi \,|\, \iota_x^\varepsilon\, e_x^\varepsilon\, \phi\rangle_{\mathcal{H}} \overset{(2.15)}{=} -\langle \psi \,|\, x\, \phi\rangle_{\mathcal{H}} = \prec\!\psi|\phi\!\succ\,.$$

Moreover, since the image of ι_x^ε coincides with S_xM, we know that e_x^ε vanishes on the orthogonal complement $S_x^\perp \subset \mathcal{H}$. Therefore,

$$e_x^\varepsilon|_{S_x}\, \psi^u(x) = e_x^\varepsilon|_{S_x}\, \pi_x\, u = e_x^\varepsilon\, u = \big(\mathfrak{R}_\varepsilon u\big)(x)\,.$$

Finally, if x is regular,

$$-\big(x|_{S_x}\big)^{-1} \iota_x^\varepsilon\, e_x^\varepsilon|_{S_xM} \overset{(2.15)}{=} \big(x|_{S_x}\big)^{-1}\, x|_{S_x} = \mathbb{1}_{S_x}\,,$$

proving that the inverse of $e_x^\varepsilon|_{S_x}$ is indeed given by the expression in (2.17). □

This proposition makes it possible to identify the spin space $S_xM \subset \mathcal{H}$ endowed with the inner product $\prec\!.|.\!\succ_x$ with a subspace of the spinor space $S_x\mathscr{M}$ with the inner product $\overline{\psi}\phi$. If the point x is singular, this is all we can expect, because in this case the spaces S_xM and $S_x\mathscr{M}$ have different dimensions and are clearly not isomorphic. As already mentioned after Definition 1.5, in most situations of physical interest the point x will be regular. In this case, we even obtain an isomorphism of S_xM and $S_x\mathscr{M}$ which preserves the inner products on these spaces. The identity (2.16) shows that, under the above identifications, the physical wave function ψ^u (as defined by (1.16)) goes over to the regularized Dirac wave function $(\mathfrak{R}_\varepsilon u)(x)$. This shows again that the causal fermion system involves the *regularized* objects. Moreover, one sees that the abstract formalism introduced in Section 1 indeed gives agreement with the usual objects in Minkowski space. We remark that the above isomorphism of S_xM and $S_x\mathscr{M}$ also makes it possible to use unambiguously the same notation for the corresponding inner product. Indeed, it is convenient denote the inner product on the Dirac spinors at a time point $x \in \mathscr{M}$ by

$$\prec\!.|.\!\succ_x \;:\; S_x\mathscr{M} \times S_x\mathscr{M} \to \mathbb{C}\,, \qquad \prec\!\psi|\phi\!\succ_x = \overline{\psi}\phi\,.$$

In order to avoid confusion, we avoided this notation so far. But from now on we will sometimes use it.

In the next proposition we compute the kernel of the fermionic projector $P^\varepsilon(x,y)$ (as defined by (1.11), where the subscript ε clarifies the dependence on the UV regularization) in Minkowski space. Moreover, we prove that the limit $\varepsilon \searrow 0$ exists in the distributional sense.

Proposition 2.7. *Assume that the points x and y are regular. Then, under the above identification of S_xM with $S_x\mathscr{M}$, the kernel of the fermionic projector has the representation*

$$P^\varepsilon(x,y) = -e_x^\varepsilon\, \iota_y^\varepsilon \;:\; S_y\mathscr{M} \to S_x\mathscr{M}\,.$$

Moreover, choosing an orthonormal basis (u_ℓ) of $\mathcal{H}$, the kernel of the fermionic projector can be written as

$$P^\varepsilon(x,y) = -\sum_\ell \big(\mathfrak{R}_\varepsilon u_\ell\big)(x)\, \overline{\big(\mathfrak{R}_\varepsilon u_\ell\big)(y)}\,. \tag{2.18}$$

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

$$P(x,y) := -\sum_\ell u_\ell(x)\, \overline{u_\ell(y)}\,. \tag{2.19}$$

More precisely, for every compact subset $K \subset \mathscr{M}$ and every $\delta > 0$, there is a constant $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ and for all test wave functions $\eta, \tilde\eta \in C_0^\infty(K, S\mathscr{M})$,

$$\left| \iint_{\mathscr{M}\times\mathscr{M}} \overline{\eta(x)}\, \big(P^\varepsilon(x,y) - P(x,y)\big)\, \tilde\eta(y)\, d^4x\, d^4y \right| \leq \delta\, |\eta|_{C^1(K)}\, |\tilde\eta|_{C^1(K)}\,. \tag{2.20}$$

We remark that, since $\mathcal{H}$ is separable, we can always choose an at most countable orthonormal basis (u_ℓ) of $\mathcal{H}$.

Proof of Proposition 2.7. We first note that

$$P^\varepsilon(x,y) = e_x^\varepsilon\, \pi_x\, y\, \big(e_y^\varepsilon|_{S_y}\big)^{-1} = -e_x^\varepsilon\, \pi_x\, y\, \big(y|_{S_y}\big)^{-1}\, \iota_y^\varepsilon = -e_x^\varepsilon\, \pi_x\, \iota_y^\varepsilon = -e_x^\varepsilon\, \iota_y^\varepsilon\,.$$

In an orthonormal basis $(u)_\ell$, the completeness relation yields for any spinor $\chi \in S_y\mathscr{M}$

$$P^\varepsilon(x,y)\,\chi = -e_x^\varepsilon\, \iota_y^\varepsilon\, \chi = -\sum_\ell \big(e_x^\varepsilon\, u_\ell\big) \langle u_\ell \,|\, \iota_y^\varepsilon\, \chi\rangle_{\mathcal{H}} = -\sum_\ell \big(e_x^\varepsilon\, u_\ell\big)\, \big(\overline{e_x^\varepsilon\, u_\ell}\, \chi\big)\,,$$

and using (2.14) gives (2.18).

In order to prove (2.20), we introduce the functionals

$$\Phi_\eta^\varepsilon \;:\; \mathcal{H} \to \mathbb{C}\,, \qquad\qquad \Phi_\eta^\varepsilon u = \int_{\mathscr{M}} \overline{\eta(x)} \big(\mathfrak{R}_\varepsilon u\big)(x)\, d^4x$$

and similarly without UV regularization,

$$\Phi_\eta \;:\; \mathcal{H} \to \mathbb{C}\,, \qquad\qquad \Phi_\eta u = \int_{\mathscr{M}} \overline{\eta(x)}\, u(x)\, d^4x\,.$$

Then the left side of (2.20) can be written in the compact form

$$\big| \Phi_\eta^\varepsilon \big(\Phi_{\tilde\eta}^\varepsilon\big)^* - \Phi_\eta \big(\Phi_{\tilde\eta}\big)^* \big|\,,$$

which can be estimated with the triangle inequality by

$$\left|\Phi_\eta^\varepsilon \left(\Phi_{\tilde\eta}^\varepsilon\right)^* - \Phi_\eta \left(\Phi_{\tilde\eta}\right)^*\right| \leq \|\Phi_\eta^\varepsilon\| \left\|\Phi_{\tilde\eta}^\varepsilon - \Phi_{\tilde\eta}\right\| + \left\|\Phi_\eta^\varepsilon - \Phi_\eta\right\| \|\Phi_{\tilde\eta}\| \,. \tag{2.21}$$

It remains to estimate the operator norms in (2.21). To this end, we use property (iii) of Definition 2.3 in the following way: First, the norm of Φ_η can be estimated by

$$\left|\Phi_\eta u\right| = \int_{\mathscr{M}} \overline{\eta(x)}\, u(x)\, d^4x \leq |\eta|_{C^0(K)} \sqrt{\mu(K)} \left(\int_K |u(x)|\, d^4x \right)^{\frac{1}{2}} ,$$

and again by applying (2.9). This gives

$$\|\Phi_\eta\| \leq c\, |\eta|_{C^0(K)} \,.$$

Next, we use the triangle inequality together with (2.8) to obtain the inequality

$$\left\|\Phi_\eta^\varepsilon\right\| \leq \left\|\Phi_\eta^\varepsilon - \Phi_\eta\right\| \leq \delta\, |\eta|_{C^1(K)} + c\, |\eta|_{C^0(K)} \leq 2c\, |\eta|_{C^1(K)} \,,$$

valid uniformly for all $\varepsilon \in (0, \varepsilon_0)$ (note that property (i) cannot be used to obtain such a uniform estimate because we have no control on how the constant c in (2.6) depends on ε). Finally, again applying (2.8), we also know that

$$\left\|\Phi_\eta^\varepsilon - \Phi_\eta\right\| \leq \delta\, |\eta|_{C^1(K)} \,.$$

Using these inequalities in (2.21) gives the result. □

2.5. **Correspondence of the Causal Structure.** We now explain how the causal structure of Minkowski space is related to corresponding notions of a causal fermion system (see Definition 1.2 and the time direction (1.10)). To this end, we need to specify $\mathcal{H}$ as a closed subspace of the solution space of the vacuum Dirac equation (2.1). Clearly, this Dirac equation can be solved by the plane-wave ansatz

$$\psi(x) = e^{-ikx}\, \chi_k$$

with a constant spinor χ_k. Evaluating the resulting algebraic equation for χ shows that the momentum k must lie on the mass shell $k^2 = m^2$. The solutions on the upper and lower mass shell are the solutions of positive respectively negative energy. In order to avoid potential confusion with other notions of energy (like energy densities or energy expectation values), we here prefer the notion of solutions of positive and negative *frequency*. Taking Dirac's original concept literally, we here describe the vacuum in Minkowski space by the completely filled Dirac sea. Thus we choose $\mathcal{H}$ as the subspace of the solution space spanned by all plane-wave solutions of negative frequency. We refer to this choice as a *Dirac sea configuration*.

Lemma 2.8. *If $\mathcal{H}$ is the subspace of the solution space of the Dirac equation* (2.1) *spanned by all negative-frequency solutions, then the unregularized kernel of the fermionic projector as defined by* (2.19) *is the tempered bi-distribution*

$$P(x,y) = \int \frac{d^4k}{(2\pi)^4}\, (\not{k} + m)\, \delta(k^2 - m^2)\, \Theta(-k_0)\, e^{-ik(x-y)} \,, \tag{2.22}$$

where Θ is the Heaviside function, and $k(x-y)$ is a short notation for the Minkowski inner product $k_j\, (x-y)^j$.

Proof. The integrand in (2.22) clearly is a tempered distribution. Hence its Fourier transform $P(x,y)$ is also a tempered distribution (in the vector $y-x$ and also in both vectors x and y). In addition, one verifies by direct computation that $P(x,y)$ is a distributional solution of the Dirac equation,

$$\begin{aligned}(i\partial\!\!\!/_x - m)\, P(x,y) &= \int \frac{d^4k}{(2\pi)^4}\, (k\!\!\!/ - m)(k\!\!\!/ + m)\, \delta(k^2-m^2)\, \Theta(-k_0)\, e^{-ik(x-y)} \\ &= \int \frac{d^4k}{(2\pi)^4}\, \big(k^2-m^2\big)\, \delta(k^2-m^2)\, \Theta(-k_0)\, e^{-ik(x-y)} = 0\,.\end{aligned}$$

Due to the factor $\Theta(-k_0)$, the distribution $P(x,y)$ is composed of solutions of negative frequency. Moreover, since the matrix $(k\!\!\!/ + m)$ has rank two, one sees that $P(x,y)$ is indeed composed of *all* negative-frequency solutions. It remains to show that the normalization of $P(x,y)$ is compatible with (2.19), meaning that

$$-2\pi \int_{\mathbb{R}^3} P\big(x,(t,\vec{y})\big)\, \gamma^0\, P\big((t,\vec{y}),z\big)\, d^3y = P(x,z)\,.$$

This identity follows by a straightforward computation: First,

$$\begin{aligned}&\int_{\mathbb{R}^3} P\big(x,(t,\vec{y})\big)\, \gamma^0\, P\big((t,\vec{y}),z\big)\, d^3y \\ &\quad= \int_{\mathbb{R}^3} d^3y \int \frac{d^4k}{(2\pi)^4}\, e^{-ik(x-y)} \int \frac{d^4q}{(2\pi)^4}\, e^{-iq(y-z)}\, P_m(k)\, \gamma^0\, P_m(q) \qquad (2.23)\\ &\quad= \int \frac{d^4k}{(2\pi)^4} \int_{\mathbb{R}} \frac{d\lambda}{2\pi}\, e^{-ikx+iqz}\, P_m(k)\, \gamma^0\, P_m(q)\Big|_{q=(\lambda,\vec{k})}\,.\end{aligned}$$

Setting $k=(\omega,\vec{k})$, we evaluate the δ-distributions inside the factors P_m,

$$\begin{aligned}\delta(k^2-m^2)\, \delta(q^2-m^2)\big|_{q=(\lambda,\vec{k})} &= \delta\big(\omega^2-|\vec{k}|^2-m^2\big)\, \delta\big(\lambda^2-|\vec{k}|^2-m^2\big) \\ &= \delta(\lambda^2-\omega^2)\, \delta\big(\omega^2-|\vec{k}|^2-m^2\big)\,.\end{aligned}$$

This shows that we only get a contribution if $\lambda = \pm\omega$. Using this fact together with the mass shell property $\omega^2 - |\vec{k}|^2 = m^2$, we can simplify the Dirac matrices according to

$$\begin{aligned}(k\!\!\!/ + m)\, \gamma^0\, (q\!\!\!/ + m) &= (\omega\gamma^0 + \vec{k}\vec{\gamma} + m)\, \gamma^0\, (\pm\omega\gamma^0 + \vec{k}\vec{\gamma} + m) \\ &= (\omega\gamma^0 + \vec{k}\vec{\gamma} + m)\, (\pm\omega\gamma^0 - \vec{k}\vec{\gamma} + m)\, \gamma^0 \\ &= \Big((\pm\omega^2 + |\vec{k}|^2 + m^2)\, \gamma^0 + (1\pm 1)\, \omega\, (\vec{k}\vec{\gamma}) + (1\pm 1)\, m\omega \Big) \\ &= \begin{cases} 2\omega\, (k\!\!\!/ + m) & \text{in case } + \\ 0 & \text{in case } -\,. \end{cases}\end{aligned}$$

Hence we only get a contribution if $\lambda = \omega$, giving rise to the identity

$$\delta(\lambda^2-\omega^2) = \frac{1}{2|\omega|}\, \delta(\lambda-\omega)\,.$$

Putting these formulas together, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^3} P\big(x,(t,\vec{y})\big)\,\gamma^0\,P\big((t,\vec{y}),z\big)\,d^3y \\
&\quad= \int \frac{d^4k}{(2\pi)^4} \int_{\mathbb{R}} \frac{d\lambda}{2\pi}\, e^{-ik(x-z)}\,\delta(\lambda-\omega)\,\delta(k^2-m^2)\,\frac{2\omega}{2|\omega|}\,(\not{k}+m)\,\Theta(-k_0) \\
&\quad= -\frac{1}{2\pi}\int \frac{d^4k}{(2\pi)^4}\, e^{-ik(x-z)}\,\delta(k^2-m^2)\,(\not{k}+m)\,\Theta(-k_0)\,.
\end{aligned}$$

This gives the result. □

The Fourier integral (2.22) can be computed in closed form, giving an expression involving Bessel functions. In preparation, it is useful to pull the Dirac matrices out of the Fourier integral. To this end, one rewrites the factor $(\not{k}+m)$ in (2.22) in terms of a differential operator in position space,

$$P(x,y) = (i\not{\partial}_x + m)\,T_{m^2}(x,y)\,, \tag{2.24}$$

where T_{m^2} is the scalar bi-distribution

$$T_{m^2}(x,y) := \int \frac{d^4k}{(2\pi)^4}\,\delta(k^2-m^2)\,\Theta(-k_0)\,e^{-ik(x-y)}\,.$$

In the next lemma, we determine the singular structure of this distribution. The method is to subtract an explicit singular distribution and to show that the difference is a *regular distribution* (i.e. a locally integrable function, denoted by L^1_{loc}). The distribution PP/ξ^2, denoted by *principal value*, is defined by evaluating weakly with a test function $\eta \in C_0^\infty(\mathscr{M})$ and by removing the positive and negative parts of the pole in a symmetric way. There are different equivalent ways of writing the principal part, each of which could serve as a possible definition:

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

Lemma 2.9. *On the light cone, the bi-distribution T_{m^2} has the following singularity structure,*

$$T_{m^2}(x,y) + \frac{1}{8\pi^3}\left(\frac{\text{PP}}{\xi^2} + i\pi\,\delta(\xi^2)\,\epsilon(\xi^0)\right) \in L^1_{\text{loc}}(\mathscr{M}\times\mathscr{M})\,, \tag{2.25}$$

where we set $\xi := y-x$. Away from the light cone (i.e. for $\xi^2 \neq 0$), $T_{m^2}(x,y)$ is a smooth function given by

$$T_{m^2}(x,y) = \begin{cases} \dfrac{m}{16\pi^2}\,\dfrac{Y_1\big(m\sqrt{\xi^2}\big)}{\sqrt{\xi^2}} + \dfrac{im}{16\pi^2}\,\dfrac{J_1\big(m\sqrt{\xi^2}\big)}{\sqrt{\xi^2}}\,\epsilon(\xi^0) & \text{if } \xi \text{ is timelike} \\[2ex] \dfrac{m}{8\pi^3}\,\dfrac{K_1\big(m\sqrt{-\xi^2}\big)}{\sqrt{-\xi^2}} & \text{if } \xi \text{ is spacelike}\,, \end{cases} \tag{2.26}$$

where J_1, Y_1 and K_1 are Bessel functions.

Proof. The Fourier integral is computed most conveniently by inserting a convergence-generating factor. Thus for any $\varepsilon > 0$ we consider the Fourier integral

$$T^\varepsilon_{m^2}(x,y) := \int \frac{d^4k}{(2\pi)^4}\, \delta(k^2-m^2)\, \Theta(-k_0)\, e^{-ik(x-y)}\, e^{-\varepsilon\,|k_0|}\,.$$

This Fourier integral can be computed pointwise, showing that $T^\varepsilon(x,y)$ is a regular distribution. Taking the limit $\varepsilon \searrow 0$ in the distributional sense, we will then obtain $T_{m^2}(x,y)$.

Setting $\xi = y - x$ and $t = \xi^0$, we first carry out the integral over k_0 to obtain

$$\begin{aligned} T^\varepsilon_{m^2}(x,y) &= \int \frac{d^4k}{(2\pi)^4}\, \delta(k^2-m^2)\, \Theta(-k_0)\, e^{ik\xi}\, e^{-\varepsilon\,|k_0|} \\ &= \int_{\mathbb{R}^3} \frac{d^3k}{(2\pi)^4}\, \frac{1}{2\sqrt{\vec{k}^2+m^2}}\, e^{-i\sqrt{\vec{k}^2+m^2}\,t - i\vec{k}\vec{\xi}}\, e^{-\varepsilon\sqrt{\vec{k}^2+m^2}}\,. \end{aligned}$$

Next, for the spatial momentum $\vec{k}$ we introduce polar coordinates $(p = |\vec{k}|, \vartheta, \varphi)$, where ϑ is the angle between $\vec{k}$ and $\vec{\xi}$, and φ is the azimuthal angle. Also setting $r = |\vec{\xi}|$, we get

$$\begin{aligned} T^\varepsilon_{m^2}(x,y) &= \int_0^\infty \frac{dp}{2(2\pi)^3} \int_{-1}^1 d\cos\theta\, \frac{p^2}{\sqrt{p^2+m^2}}\, e^{-(\varepsilon+it)\sqrt{p^2+m^2}}\, e^{-ipr\cos\theta} \\ &= \frac{1}{r} \int_0^\infty \frac{dp}{(2\pi)^3}\, \frac{p}{\sqrt{p^2+m^2}}\, e^{-(\varepsilon+it)\sqrt{p^2+m^2}}\, \sin(pr) \\ &= \frac{m^2}{(2\pi)^3} \frac{K_1\big(m\sqrt{r^2+(\varepsilon+it)^2}\big)}{m\sqrt{r^2+(\varepsilon+it)^2}}, \end{aligned} \tag{2.27}$$

where the last integral was carried out using [29, formula (3.961.1)]. Here the square root and the Bessel function K_1 is defined as usual using a branch cut along the negative real axis.

When taking the limit $\varepsilon \searrow 0$, one must be careful for two reasons. First, a pole forms on the light cone $t = \pm r$. Second, the Bessel function K_1 involves logarithms, which must be evaluated in the complex plane using the branch cut along the negative real axis. For clarity, we treat these two issues after each other. The asymptotic expansion of the Bessel function (see [34, (10.31.1)])

$$K_1(z) = \frac{1}{z} + \mathcal{O}\big(z \log z\big)$$

yields that the pole on the light cone is of the form

$$T^\varepsilon_{m^2}(x,y) = \frac{1}{(2\pi)^3}\, \frac{1}{r^2+(\varepsilon+it)^2} + \mathcal{O}\big(\log|\xi^2|\big)\,,$$

uniformly in ε. Therefore, after subtracting the pole, we can take the limit $\varepsilon \searrow 0$ as a locally integrable function, i.e.

$$\lim_{\varepsilon\searrow 0} \left(T^\varepsilon_{m^2}(x,y) - \frac{1}{(2\pi)^3}\, \frac{1}{r^2+(\varepsilon+it)^2} \right) \in L^1_{\text{loc}}(\mathscr{M}\times\mathscr{M})\,.$$

For the subtracted pole, the limit $\varepsilon \searrow 0$ can be computed in the distributional sense by

$$\lim_{\varepsilon\searrow 0} \frac{1}{r^2+(\varepsilon+it)^2} = \lim_{\varepsilon\searrow 0} \frac{1}{r^2-t^2+i\varepsilon t} = -\frac{\mathrm{PP}}{\xi^2} - i\pi\,\delta(\xi^2)\,\epsilon(\xi^0)\,,$$

where we used the distributional equation

$$\lim_{\varepsilon \searrow 0} \left(\frac{1}{x - i\varepsilon} - \frac{1}{x + i\varepsilon} \right) = 2\pi i\, \delta(x) \,.$$

Here "PP" again denotes the principal value, and ϵ is the step function $\epsilon(x) = 1$ for $x \geq 0$ and $\epsilon(x) = -1$ otherwise. This gives (2.25).

In order to compute the regular part of the distribution T_{m^2}, we may disregard the singularity on the light cone and may consider the case that ξ is either spacelike or timelike. In the first case, the argument $m\sqrt{r^2 + (\varepsilon + it)^2}$ of the Bessel function converges to the positive real axis, where the Bessel function is analytic. This gives the lower equation in (2.26). In the remaining case that ξ is timelike, the argument $m\sqrt{r^2 + (\varepsilon + it)^2}$ converges to the imaginary axis (more precisely, to the upper imaginary axis if $t > 0$ and to the lower imaginary axis if $t < 0$). Using the relations [34, (10.27.9) and (10.27.10)]

$$i\pi J_1(z) = -iK_1(-iz) - iK_1(iz) \qquad \text{and} \qquad -\pi Y_1(z) = -iK_1(-iz) + iK_1(iz)$$

(valid if $|\arg z| < \frac{\pi}{2}$), one can express K_1 near the upper and lower imaginary axis by

$$K_1(\pm iz) = -\frac{\pi}{2}\big(J_1(z) \mp iY_1(z)\big) \,.$$

Using these identities in (2.27) and using that the Bessel functions J_1 and K_1 are analytic in a neighborhood of the positive real axis, one can take the limit $\varepsilon \searrow 0$ to obtain the upper equation in (2.26). □

We point out that the Bessel functions in (2.26) are all real-valued. In particular, one sees that $T(x,y)$ is real-valued if the vector ξ is spacelike.

Using the result of Lemma 2.9 in (2.24), one can derive corresponding formulas for $P(x,y)$. In particular, differentiating (2.25), one sees that $P(x,y)$ has an even stronger singularity on the light cone which involves terms of the form $1/\xi^4$ and $\delta'(\xi^2)$. Differentiating (2.26), carrying out the derivatives with the chain rule and using formulas for the derivatives of Bessel functions (see [34, (10.6.6) and (10.29.4)]), one can also express the fermionic projector $P(x,y)$ in terms of Bessel functions. We do not give the resulting formulas, because we do not need the detailed form later on. Instead, we here prefer to argue with general properties of the distribution $P(x,y)$. This makes it possible to infer qualitative properties of the eigenvalues of A_{xy}, even without referring to the detailed form of the formulas in Lemma 2.9. From Lorentz symmetry, we know that for all x and y with spacelike or timelike separation, $P(x,y)$ can be written as

$$P(x,y) = \alpha\, \xi_j \gamma^j + \beta\, \mathbb{1} \tag{2.28}$$

with two complex-valued functions α and β (where again $\xi = y - x$). Taking the conjugate with respect to the spin scalar product, we see that

$$P(y,x) = \overline{\alpha}\, \xi_j \gamma^j + \overline{\beta}\, \mathbb{1} \,. \tag{2.29}$$

As a consequence,

$$A_{xy} = P(x,y)\, P(y,x) = a\, \xi_j \gamma^j + b\, \mathbb{1} \tag{2.30}$$

with two real parameters a and b given by

$$a = \alpha\overline{\beta} + \beta\overline{\alpha} \,, \qquad b = |\alpha|^2\, \xi^2 + |\beta|^2 \,. \tag{2.31}$$

Applying the formula $(A_{xy} - b\mathbb{1})^2 = a^2\,\xi^2\,\mathbb{1}$, the roots of the characteristic polynomial of A_{xy} are computed by

$$b \pm \sqrt{a^2\,\xi^2}\,. \tag{2.32}$$

Therefore, the eigenvalues of the closed chain are either real, or else they form a complex conjugate pair. Which of the two cases appears is determined by the sign of the factor ξ^2. This gives the agreement of the different notions of causality in the following sense:

Proposition 2.10. *Assume that $P(x,y)$ is the unregularized kernel of the fermionic projector of the vacuum* (2.22)*, and that the eigenvalues $\lambda_1^{xy}, \ldots, \lambda_4^{xy}$ are computed as the eigenvalues of the closed chain* (1.12)*. Then the following statements hold: If the points $x, y \in \mathscr{M}$ have spacelike separation in Minkowski space, then they are also spacelike separated in the sense of Definition 1.2. If, on the other hand, the points $x, y \in \mathscr{M}$ have timelike separation in Minkowski space, then they are also timelike separated in the sense of Definition 1.2. Even more, they are properly timelike separated (see Definition 1.6) in the sense that the closed chain A_{xy} has strictly positive eigenvalues and definite eigenspaces. Finally, if the points $x, y \in \mathscr{M}$ have lightlike separation in Minkowski space, then the causal structure of Definition 1.2 is ill-defined.*

The fact that the causal structure is ill-defined for lightlike separation again explains why an UV regularization must be introduced.

Proof of Proposition 2.10. If the vector $\xi = y - x$ is spacelike, then the term ξ^2 is negative. Thus the eigenvalues in (2.32) form a complex conjugate pair, implying that they all have the same absolute value. Thus the points are spacelike separated in the sense of Definition 1.2.

If the vector ξ is timelike, the term ξ^2 in (2.32) is positive, so that the λ_j are all real. In order to show that they do not have the same absolute value, we need to verify that the parameters a and b are both non-zero. This makes it necessary to refer to the explicit formula involving Bessel functions (2.26): The Bessel functions Y_1 and J_1 do not have joint zeros on the positive real axis. As a consequence, the parameter β in (2.28) is non-zero. Likewise, the derivatives Y_1' and J_1' do not have joint zeros (as can again be verified from the fact that the Bessel functions form a fundamental system). This implies that the parameter α in (2.28) is non-zero. We conclude that the parameter b in (2.31) is non-zero. The combination of α and β in the formula for a in (2.31) can be rewritten in terms of a Wronskian of the Bessel function. This Wronskian can be computed explicitly using [34, (10.5.2)], implying that a is non-zero. We conclude that the points x and y are timelike separated in the sense of Definition 1.2.

In order to get the connection to proper timelike separation, recall that if ξ is a timelike vector of Minkowski space, then the closed chain has the form (2.31) with $a, b \neq 0$. A direct computation shows that this matrix is diagonalizable and that the eigenspaces are definite with respect to the spin scalar product. Moreover, applying the Schwarz inequality to the explicit formulas (2.31), one obtains

$$|a|\,\sqrt{\xi^2} = 2\,\mathrm{Re}\left(\alpha\,\sqrt{\xi^2}\,\overline{\beta}\right) \overset{(\star)}{\leq} |\alpha|^2 \xi^2 + |\beta|^2 = b\,, \tag{2.33}$$

proving that the eigenvalues in (2.32) are non-negative. It remains to show that none of these eigenvalues vanishes. To this end, it suffices to show that the inequality $(\star)$

in (2.33) is strict, which in turn is equivalent to proving that

$$\mathrm{Im}\left(\alpha\overline{\beta}\right) \neq 0\,.$$

This inequality follows by a detailed analysis of the Bessel functions (see [15, proof of Lemma 4.3]). We conclude that x and y are indeed properly timelike separated.

If the vector ξ is lightlike, then $P(x,y)$ is not defined pointwise. As a consequence, the closed chain is ill-defined. □

This proposition cannot be applied directly to causal fermion systems because, as explained in §2.1 and §2.2, constructing a causal fermion system makes it necessary to introduce an UV regularization. Nevertheless, the above proposition also gives correspondence of the different notions of causality for causal fermion systems describing the Minkowski vacuum, as we now explain. Thus let us consider the causal fermion system corresponding to the regularized fermionic projector of the vacuum $P^\varepsilon(x,y)$. In the limit $\varepsilon \searrow 0$, the kernel of the fermionic projector $P^\varepsilon(x,y)$ converges to the unregularized kernel $P(x,y)$ (see (2.20) in Proposition 2.6). If this convergence is pointwise, i.e. if for given space-time points $x, y \in \mathscr{M}$,

$$\lim_{\varepsilon \searrow 0} P^\varepsilon(x,y) = P(x,y)\,, \tag{2.34}$$

then the results of Proposition 2.10 also apply to the causal fermion system, up to error terms which tend to zero as $\varepsilon \searrow 0$. Thinking of ε as the Planck scale, this means physically that the notion of causality of Definition 1.2 agrees with the usual notion of causality in Minkowski space, up to corrections which are so small that they cannot be observed. The subtle point of this argument is that it requires pointwise convergence (2.34). Clearly, such a pointwise convergence cannot hold if x and y are lightlike separated, because the right side of (2.34) is ill-defined pointwise. Expressed for a causal fermion system for fixed ε on the Planck scale, this means that the notion of causality of Definition 1.2 does *not* agree with the usual notion of causality if the vector ξ is almost lightlike in the sense that $\big||\xi^0| - |\vec{\xi}|\big| \lesssim \varepsilon$. This is not surprising because we cannot expect that the notion of causality in Minkowski space holds with a higher resolution than the regularization scale ε. The remaining question is whether we have pointwise convergence (2.34) if the points x and y have timelike or spacelike separation. The answer is yes for a large class of regularizations (like for example the regularization by mollification in Example 2.4). However, the general notion of Definition 2.3 only gives weak convergence of the kernels (2.20). This shortcoming could be removed by adding a condition to Definition 2.3 which ensures pointwise convergence away from the light cone. On the other hand, such an additional condition seems unnecessary, and therefore it seems preferable not to impose it. Nevertheless, the physical picture is that the regularized kernel should converge pointwise, at least for generic points x and y which lie sufficiently far away from the light cone. With this in mind, Proposition 2.10 indeed shows that the notion of causality of Definition 1.2 corresponds to the usual notion of causality in Minkowski space, up to corrections which are so small that they are irrelevant in most situations of interest.

We conclude this section by explaining why the functional $\mathcal{C}$ introduced in (1.9) gives information on the time direction. Our first task is to rewrite this functional in terms of the regularized kernel of the fermionic projector $P^\varepsilon(x,y)$.

Lemma 2.11. *Assume that the operator $P^\varepsilon(x,x) : S_x\mathscr{M} \to S_x\mathscr{M}$ is invertible. Then, setting*

$$\nu(x) = P^\varepsilon(x,x)^{-1} \;:\; S_x\mathscr{M} \to S_x\mathscr{M}\;, \tag{2.35}$$

the functional $\mathfrak{C}$, (1.9), can be written as

$$\mathfrak{C}(x,y) = i\,\mathrm{Tr}_{S_x}\Big(P^\varepsilon(x,y)\,\nu(y)\,P^\varepsilon(y,x)\,\big[\nu(x), A_{xy}\big]\Big)\,. \tag{2.36}$$

Proof. Since $P(x,x) = \pi_x x|_{S_x} = x|_{S_x}$, we know that $\nu(x) = (x|_{S_x})^{-1}$. Thus

$$\begin{aligned}\pi_x\, y\, x\, \pi_y\, \pi_x|_{S_x} &= \pi_x y\, \pi_y x\, \pi_x y\, \nu(y)\, \pi_y x\, \nu(x)|_{S_x} \\ &= P^\varepsilon(x,y)\,P^\varepsilon(y,x)\,P^\varepsilon(x,y)\,\nu(y)\,P^\varepsilon(y,x)\,\nu(x)|_{S_x}\,.\end{aligned}$$

Using this formula in (1.9), we obtain

$$\begin{aligned}\mathfrak{C}(x,y) &= i\,\mathrm{Tr}_{S_x}\left(y\,x\,\pi_y\,\pi_x|_{S_x} - y\,\pi_x\,\pi_y\,x|_{S_x}\right) \\ &= i\,\mathrm{Tr}_{S_x}\Big(P^\varepsilon(x,y)\,P^\varepsilon(y,x)\,P^\varepsilon(x,y)\,\nu(y)\,P^\varepsilon(y,x)\,\nu(x) \\ &\qquad - P^\varepsilon(x,y)\,P^\varepsilon(y,x)\,\nu(x)\,P^\varepsilon(x,y)\,\nu(y)\,P^\varepsilon(y,x)\Big) \\ &= i\,\mathrm{Tr}_{S_x}\Big(P^\varepsilon(x,y)\,\nu(y)\,P^\varepsilon(y,x)\,\nu(x)\,P^\varepsilon(x,y)\,P^\varepsilon(y,x) \\ &\qquad - P^\varepsilon(x,y)\,\nu(y)\,P^\varepsilon(y,x)\,P^\varepsilon(x,y)\,P^\varepsilon(y,x)\,\nu(x)\Big)\,.\end{aligned}$$

This gives the result. □

We point out that the operator $\nu(x)$ in (2.35) is ill-defined without UV regularization because evaluating the distribution $P(x,y)$ on the diagonal $x=y$ has no mathematical meaning. As a consequence, the functional $\mathfrak{C}$ is ill-defined without UV regularization, even if x and y have timelike separation. This makes the following computation somewhat delicate. In order to keep the analysis reasonably simple, we assume that the regularized kernel of the fermionic projector has *vector-scalar structure*, meaning that it is of the general form

$$P^\varepsilon(x,y) = v_j^\varepsilon(x,y)\,\gamma^j \;+\; \beta^\varepsilon(x,y)\,\mathbb{1} \tag{2.37}$$

with a vectorial and a scalar component. Here $v^\varepsilon(x,y)$ is a complex vector field (i.e. it can be written as $v^\varepsilon = u^\varepsilon + iw^\varepsilon$ with Minkowski vectors u^ε and w^ε which need not be collinear). Then, evaluating (2.37) for $x=y$, one sees that $P^\varepsilon(x,x)$ can be written as

$$P^\varepsilon(x,x) = v_j^\varepsilon(x)\,\gamma^j \;+\; \beta^\varepsilon(x)\,\mathbb{1}$$

(where we set $v^\varepsilon(x) = v^\varepsilon(x,x)$ and $\beta^\varepsilon(x) = \beta^\varepsilon(x,x)$). Since $P^\varepsilon(x,x)$ is a symmetric operator on $S_x\mathscr{M}$, it follows that v^ε is a real vector field, and β a real-valued function. For a large class of regularizations, the matrix $P^\varepsilon(x,x)$ is invertible because the vectorial component dominates the scalar component. With this in mind, we here assume that $\nu(x)$ exists. Then it is given by

$$\nu(x) = \frac{1}{\rho(x)}\left(v_j^\varepsilon(x)\,\gamma^j - \beta^\varepsilon(x)\,\mathbb{1}\right), \tag{2.38}$$

where $\rho := v_j^\varepsilon (v^\varepsilon)^j - (\beta^\varepsilon)^2$. Now we can compute the composite expression in (2.36), working for all other terms with the unregularized formulas (which is again justified if we have pointwise convergence (2.34)). This gives the following result.

Proposition 2.12. *Using* (2.38) *and replacing* $P^\varepsilon(x,y)$, $P^\varepsilon(y,x)$ *and* A_{xy} *by the unregularized expressions* (2.28), (2.29) *and* (2.30), *the functional* $\mathfrak{C}$ *is given by*

$$\mathfrak{C}(x,y) = \frac{16a}{\rho(x)\,\rho(y)}\; \mathrm{Im}\left(\alpha\overline{\beta}\right) \left(v^\varepsilon(x)^j\, \xi_j\, v^\varepsilon(y)^k\, \xi_k - \xi^2\, v^\varepsilon(x)^j\, v^\varepsilon(y)_j\right). \tag{2.39}$$

Proof. Using (2.38) and (2.30) in (2.36) gives

$$\begin{aligned}\mathfrak{C}(x,y) &= i\,\mathrm{Tr}_{S_x}\left(P(x,y)\,\nu(y)\,P(y,x)\,\left[\nu(x), A_{xy}\right]\right)\\ &= \frac{ia}{\rho(x)}\,\mathrm{Tr}_{S_x}\left(P(x,y)\,\nu(y)\,P(y,x)\,\left[\not{v}^\varepsilon(x), \not{\xi}\right]\right),\end{aligned}$$

where in the last step we used that the scalar components of A_{xy} and $\nu(x)$ drop out of the commutator. Taking the scalar component of $\nu(y)$, the two factors $P(x,y)$ and $P(y,x)$ combine to the closed chain, which according to (2.30) has no bilinear component, so that the trace vanishes. Therefore, we only need to take into account the vectorial component of $\nu(y)$. Using (2.28) and (2.29), we obtain

$$\begin{aligned}\mathfrak{C}(x,y) &= \frac{ia}{\rho(x)\,\rho(y)}\,\mathrm{Tr}_{S_x}\left(\left(\alpha\not{\xi} + \beta\,\mathbb{1}\right)\not{v}^\varepsilon(y)\left(\overline{\alpha}\not{\xi} + \overline{\beta}\,\mathbb{1}\right)\left[\not{v}^\varepsilon(x), \not{\xi}\right]\right)\\ &= -\frac{a}{\rho(x)\,\rho(y)}\;\mathrm{Im}\left(\alpha\overline{\beta}\right)\;\mathrm{Tr}_{S_x}\left(\left[\not{\xi}, \not{v}^\varepsilon(y)\right]\left[\not{v}^\varepsilon(x), \not{\xi}\right]\right).\end{aligned}$$

Computing the trace of the product of Dirac matrices gives the result. □

For the interpretation of the formula (2.39), we first consider the case that y and x have space-like separation. In this case, it turns out that the prefactor $\mathrm{Im}(\alpha\overline{\beta})$ vanishes, so that (2.39) gives no information on a time direction. This is consistent with the fact that for points in Minkowski space with space-like separation, the notions of future- and past-directed depend on the observer and cannot be defined in a covariant manner. However, if y and x have timelike separation, then the factors a and $\mathrm{Im}(\alpha\overline{\beta})$ are indeed both non-zero (see the proof of Proposition 2.10). Therefore, the functional $\mathfrak{C}$ is non-zero, provided that the vector ξ is *non-degenerate* in the sense that it is linearly independent of both $v^\varepsilon(x)$ and $v^\varepsilon(y)$. Since the set of directions ξ for which these vectors are linearly dependent has measure zero, we may always restrict attention to non-degenerate directions. Moreover, the formula (2.39) shows that the functional $\mathfrak{C}$ does not change sign for ξ inside the upper or lower light cone. On the other hand, $\mathfrak{C}$ is antisymmetric under sign flips of ξ because interchanging x and y in (1.9) obviously gives a minus sign.

We conclude that for the regularized Dirac sea vacuum, the sign of the functional $\mathfrak{C}$ distinguishes a time direction. Asymptotically as $\varepsilon \searrow 0$, this time direction agrees with the distinction of the causal past and causal future in Minkowski space.

To summarize, in this section we saw how the intrinsic structures of a causal fermion system correspond to the usual structures in Minkowski space. To this end, we constructed causal fermion systems from a regularized Dirac sea configuration and analyzed the asymptotics as the UV regularization is removed. For brevity, we only considered the topological and causal structure of space-time as well as spinors and wave functions. The reader interested in geometric structures like connection and curvature is referred to the detailed exposition in [15]. Moreover, in Section 5 below we shall explain how the methods and results introduced in this section can be generalized to interacting systems.

3. Underlying Physical Principles

In order to clarify the physical concepts, we now briefly discuss the underlying physical principles. Causal fermion systems evolved from an attempt to combine several physical principles in a coherent mathematical framework. As a result, these principles appear in the framework in a specific way:

- The **principle of causality** is built into a causal fermion system in a specific way, as was explained in §1.2 above.
- The **Pauli exclusion principle** is incorporated in a causal fermion system, as can be seen in various ways. One formulation of the Pauli exclusion principle states that every fermionic one-particle state can be occupied by at most one particle. In this formulation, the Pauli exclusion principle is respected because every wave function can either be represented in the form ψ^u (the state is occupied) with $u \in \mathcal{H}$ or it cannot be represented as a physical wave function (the state is not occupied). Via these two conditions, the fermionic projector encodes for every state the occupation numbers 1 and 0, respectively, but it is impossible to describe higher occupation numbers. More technically, one may obtain the connection to the fermionic Fock space formalism by choosing an orthonormal basis $u_1, \ldots, u_f$ of $\mathcal{H}$ and forming the f-particle Hartree-Fock state
$$\Psi := \psi^{u_1} \wedge \cdots \wedge \psi^{u_f} .$$
Clearly, the choice of the orthonormal basis is unique only up to the unitary transformations
$$u_i \to \tilde{u}_i = \sum_{j=1}^{f} U_{ij}\, u_j \quad \text{with} \quad U \in \mathrm{U}(f) .$$
Due to the anti-symmetrization, this transformation changes the corresponding Hartree-Fock state only by an irrelevant phase factor,
$$\psi^{\tilde{u}_1} \wedge \cdots \wedge \psi^{\tilde{u}_f} = \det U\; \psi^{u_1} \wedge \cdots \wedge \psi^{u_f} .$$
Thus the configuration of the physical wave functions can be described by a fermionic multi-particle wave function. The Pauli exclusion principle becomes apparent in the total anti-symmetrization of this wave function.
- A **local gauge principle** becomes apparent once we choose basis representations of the spin spaces and write the wave functions in components. Denoting the signature of $(S_x, \prec.|.\succ_x)$ by $(p(x), q(x))$, we choose a pseudo-orthonormal basis $(\mathfrak{e}_\alpha(x))_{\alpha=1,\ldots,p+q}$ of S_x. Then a wave function ψ can be represented as
$$\psi(x) = \sum_{\alpha=1}^{p+q} \psi^\alpha(x)\, \mathfrak{e}_\alpha(x)$$
with component functions $\psi^1, \ldots, \psi^{p+q}$. The freedom in choosing the basis $(\mathfrak{e}_\alpha)$ is described by the group $\mathrm{U}(p,q)$ of unitary transformations with respect to an inner product of signature (p,q). This gives rise to the transformations
$$\mathfrak{e}_\alpha(x) \to \sum_{\beta=1}^{p+q} U^{-1}(x)^\beta_\alpha\, \mathfrak{e}_\beta(x) \qquad \text{and} \qquad \psi^\alpha(x) \to \sum_{\beta=1}^{p+q} U(x)^\alpha_\beta\, \psi^\beta(x)$$
with $U \in \mathrm{U}(p,q)$. As the basis $(\mathfrak{e}_\alpha)$ can be chosen independently at each spacetime point, one obtains *local gauge transformations* of the wave functions, where

the gauge group is determined to be the isometry group of the spin scalar product. The causal action is *gauge invariant* in the sense that it does not depend on the choice of spinor bases.

- The **equivalence principle** is incorporated in the following general way. Space-time $M := \operatorname{supp} \rho$ together with the universal measure ρ form a topological measure space, being a more general structure than a Lorentzian manifold. Therefore, when describing M by local coordinates, the freedom in choosing such coordinates generalizes the freedom in choosing general reference frames in a space-time manifold. Therefore, the equivalence principle of general relativity is respected. The causal action is *generally covariant* in the sense that it does not depend on the choice of coordinates.

4. The Dynamics of Causal Fermion Systems

Similar to the Einstein-Hilbert action in general relativity, in the causal action principle one varies space-time as well as all structures therein globally. This global viewpoint implies that it is not obvious what the causal action principle tells us about the dynamics of the system. The first step for clarifying the situation is to derive the Euler-Lagrange (EL) equations corresponding to the causal action principle (§4.1). Similar to the Einstein or Maxwell equations, these EL equations should describe the dynamics. Additional insight is gained by studying Noether-like theorems which specify the quantities which are conserved in the dynamics (§4.2). Finally, we review results on the initial value problem (§4.3). We remark that more explicit information on the dynamics is obtained by considering limiting cases in which the EL equations corresponding to the causal action reduce to equations of a structure familiar from classical field theory and quantum field theory (see Section 5).

4.1. **The Euler-Lagrange Equations.** We now return to the abstract setting of Section 1. Our goal is to derive the EL equations corresponding to the causal action principle in the form most useful for our purposes. The method is to consider so-called *variations of the physical wave functions* which we now introduce (for more general variations see Remark 4.3 below). Let $(\mathcal{H}, \mathcal{F}, \rho)$ be a causal fermion system. We assume that ρ is a minimizer of the causal action principle. However, we do not want to assume that the total volume $\rho(\mathcal{F})$ be finite. Instead, we merely assume that ρ is *locally finite* in the sense that $\rho(K) < \infty$ for every compact subset $K \subset \mathcal{F}$. Our starting point is the wave evaluation operator Ψ introduced in (1.18),

$$\Psi \,:\, \mathcal{H} \to C^0(M, SM)\,, \qquad u \mapsto \psi^u\,.$$

We now vary the wave evaluation operator. Thus for any $\tau \in (-\delta, \delta)$ we consider a mapping $\Psi_\tau : \mathcal{H} \to C^0(M)$. For $\tau = 0$, this mapping should coincide with the wave evaluation operator Ψ. The family $(\Psi_\tau)_{\tau \in (-\delta,\delta)}$ can be regarded as a simultaneous variation of all physical wave functions of the system. In fact, for any $u \in \mathcal{H}$, the variation of the corresponding physical wave function is given by

$$\psi^u_\tau := \Psi_\tau(u) \in C^0(M, SM)\,.$$

Next, we introduce the corresponding local correlation operators F_τ by

$$F_\tau(x) := -\Psi_\tau(x)^* \Psi_\tau(x) \qquad \text{so that} \qquad F_\tau \,:\, M \to \mathcal{F}\,.$$

In view of (1.19), we know that $F_0(x) = x$. Therefore, the family $(F_\tau)_{\tau \in (-\delta,\delta)}$ is a variation of the local correlation operators. Taking the push-forward measure gives

rise to a family of universal measures,

$$\rho_\tau := (F_\tau)_* \rho \,. \tag{4.1}$$

Since F_0 is the identity, we know that $\rho_0 = \rho$. Therefore, the family $(\rho_\tau)_{\tau \in (-\delta, \delta)}$ is indeed a variation of the universal measure.

We now work out the EL equations for the resulting class of variations of the universal measure. In order for the constructions to be mathematically well-defined, we need a few technical assumptions which are summarized in the following definition.

Definition 4.1. *The* **variation of the physical wave functions** *is* **smooth** *and* **compact** *if the family of operators $(\Psi_\tau)_{\tau \in (-\delta, \delta)}$ has the following properties:*

(a) The variation is trivial on the orthogonal complement of a finite-dimensional subspace $I \subset \mathcal{H}$, i.e.

$$\Psi_\tau|_{I^\perp} = \Psi \qquad \text{for all } \tau \in (-\delta, \delta) \,.$$

(b) There is a compact subset $K \subset M$ outside which the variation is trivial, i.e.

$$\big(\Psi_\tau(u)\big)\big|_{M \setminus K} = \big(\Psi(u)\big)\big|_{M \setminus K} \qquad \text{for all } \tau \in (-\delta, \delta) \text{ and } u \in \mathcal{H} \,.$$

(c) The Lagrangian is continuously differentiable in the sense that the derivative

$$\frac{d}{d\tau} \mathcal{L}\big(x, F_\tau(y)\big)\big|_{\tau=0} \tag{4.2}$$

exists and is continuous on $M \times M$.

With the conditions (a) and (b) we restrict attention to variations which are sufficiently well-behaved (similar as in the classical calculus of variations, where one restricts attention to smooth and compactly supported variations). It is a delicate point to satisfy the condition (c), because (due to the absolute values of the eigenvalues in (1.1)) the Lagrangian is only Lipschitz continuous on $\mathcal{F} \times \mathcal{F}$. Therefore, the derivative in (4.2) does not need to exist, even if $F_\tau(y)$ is smooth. This means that in the applications, one must verify that the condition (c) holds (for details see the computations in [6]). Here we simply assume that the variation of the wave functions is smooth and compact.

By definition of the push-forward measure (4.1), we know that for any integrable function f on $\mathcal{F}$,

$$\int_{\mathcal{F}} f(x)\, d\rho_\tau = \int_{\mathcal{F}} f(F_\tau(x))\, d\rho \,. \tag{4.3}$$

In this way, the variation of the measure can be rewritten as a variation of the arguments of the integrand. In particular, the variation of the action can be written as

$$\iint_{M \times M} \mathcal{L}\big(F_\tau(x), F_\tau(y)\big)\, d\rho(x)\, d\rho(y)$$

(and similarly for the other integrals). Another benefit of working with the push-forward measure (4.1) is that the total volume is preserved. Namely, combining the identity (4.3) with the assumption in Definition 4.1 (b), one readily verifies that the volume constraint (1.3) is satisfied in the sense that ρ_τ satisfies the conditions (1.7).

We consider first variations, treating the constraints with Lagrange multipliers (this procedure is justified in [1]). Since the volume constraint is already respected, it

remains to consider the trace constraint (1.4) and the boundedness constraint (1.5). We conclude that first variations of the functional

$$\mathcal{S}_{\kappa,\lambda} := \mathcal{S} + \kappa \left(\mathcal{T} - C\right) - \lambda \left(\int_{\mathcal{F}} \mathrm{tr}(x)\, d\rho - c \right) \tag{4.4}$$

vanish for suitable values of the Lagrange parameters $\kappa, \lambda \in \mathbb{R}$, where the constants C and c are the prescribed values of the constraints. For clarity, we point out that the boundedness constraint merely is an inequality. The method for handling this inequality constraint is to choose $\kappa = 0$ if $\mathcal{T}(\rho) < C$, whereas in the case $\mathcal{T}(\rho) = C$ the Lagrange multiplier κ is in general non-zero (for details see again [1]). Introducing the short notation

$$\mathcal{L}_\kappa(x,y) := \mathcal{L}(x,y) + \kappa\, |xy|^2 \,,$$

we can write the effective action as

$$\mathcal{S}_{\kappa,\lambda}(\rho_\tau) = \iint_{M \times M} \mathcal{L}_\kappa(x,y)\, d\rho(x)\, d\rho(y) - \lambda \int_M \mathrm{tr}\left(F_\tau(x)\right) d\rho(x) - \kappa C + \lambda c \,.$$

Now we can compute the first variation by differentiating with respect to τ. It is most convenient to express the causal action and the constraints in terms of the kernel of the fermionic projector (just as explained at the beginning of §1.3). Moreover, it is preferable to consider the Lagrangian $\mathcal{L}_\kappa(x,y)$ as a function only of $P_\tau(x,y)$ by writing the closed chain as

$$A_{xy} = P_\tau(x,y)\, P_\tau(x,y)^* \tag{4.5}$$

(where $P_\tau(x,y)^*$ denotes similar to (1.13) the adjoint with respect to the spin scalar product). We use the notation

$$\delta P(x,y) = \frac{d}{d\tau}\, P_\tau(x,y)\Big|_{\tau=0} \,,$$

and similarly for other functions. When computing the variation of the Lagrangian, one must keep in mind that $\mathcal{L}_\kappa(x,y)$ depends both on $P_\tau(x,y)$ and on its adjoint $P_\tau(x,y)^*$ (cf. (4.5)). Therefore, when applying the chain rule, we obtain contributions which are complex linear and complex anti-linear in $\delta P_\tau(x,y)$. We write the first variation with traces as

$$\delta \mathcal{L}_\kappa(x,y) = \mathrm{Tr}_{S_y}\left(B\, \delta P(x,y)\right) + \mathrm{Tr}_{S_x}\left(C\, \delta P(x,y)^*\right)$$

with linear operators $B : S_x \to S_y$ and $C : S_y \to S_x$. Since $\delta P(x,y)$ can be chosen arbitrarily, this equation uniquely defines both B and C. Since the variation of the Lagrangian is always real-valued, it follows that $C = B^*$. Using furthermore the symmetry of the Lagrangian in the arguments x and y, we conclude that the first variation of the Lagrangian can be written as (see also [8, Section 5.2])

$$\delta \mathcal{L}_\kappa(x,y) = \mathrm{Tr}_{S_y}\left(Q(y,x)\, \delta P(x,y)\right) + \mathrm{Tr}_{S_x}\left(Q(x,y)\, \delta P(x,y)^*\right) \tag{4.6}$$

with a kernel $Q(x,y) : S_y \to S_x$ which is symmetric in the sense that

$$Q(x,y)^* = Q(y,x) \,. \tag{4.7}$$

The EL equations are expressed in terms of the kernel $Q(x,y)$ as follows.

Proposition 4.2. (Euler-Lagrange equations) *Let ρ be a minimizer of the causal action principle. Then for a suitable choice of the Lagrange parameters λ and κ, the integral operator Q with kernel defined by* (4.6) *satisfies the equations*

$$\int_M Q(x,y)\, \psi^u(y)\, d\rho(y) = \frac{\lambda}{2}\, \psi^u(x) \qquad \text{for all } u \in \mathcal{H} \text{ and } x \in M \,. \tag{4.8}$$

We note for clarity that by writing the equation (4.8) we imply that the integral must exist and be finite.

Proof of Proposition 4.2. Using (4.6), the first variation of $\mathcal{S}_{\kappa,\lambda}$ is computed by

$$\delta\mathcal{S}_{\kappa,\lambda} = \iint_{M\times M} \Big(\mathrm{Tr}_{S_y}\big(Q(y,x)\,\delta P(x,y)\big) + \mathrm{Tr}_{S_x}\big(Q(x,y)\,\delta P(x,y)^*\big)\Big)\, d\rho(x)\, d\rho(y) \\ - \lambda \int_M \mathrm{Tr}\big(\delta P(x,x)\big)\, d\rho(x)\,.$$

Noting that $\delta P(x,y) = \delta P(y,x)$, after renaming the integration variables in the first summand of the double integral, we obtain

$$\delta\mathcal{S}_{\kappa,\lambda} = 2 \iint_{M\times M} \mathrm{Tr}_{S_x}\big(Q(x,y)\,\delta P(y,x)\big) - \lambda \int_M \mathrm{Tr}_{S_x}\big(\delta P(x,x)\big)\, d\rho(x)\,. \tag{4.9}$$

Next, we express δP in terms of the variation of the physical wave functions. By Lemma 1.3, we know that

$$P_\tau(x,y) = -\Psi_\tau(x)\Psi_\tau(y)^*\,.$$

Differentiating this relation gives

$$\delta P(x,y) = -(\delta\Psi)(x)\;\Psi(y)^* - \Psi(x)\;(\delta\Psi)(y)^*\,.$$

We now specialize to the case that the variation is trivial on the orthogonal complement of a one-dimensional subspace $I = \mathrm{span}(u) \subset \mathcal{H}$. Then for any $\phi \in S_y$,

$$\delta P(x,y)\,\phi = -\delta\psi^u(x) \prec \psi^u(y)\,|\,\phi\succ_y - \psi^u(x) \prec\delta\psi^u(y)\,|\,\phi\succ_y\,.$$

By inserting a phase factor according to

$$\delta\psi^u \to e^{i\varphi}\,\delta\psi^u\,,$$

one sees that $\delta\psi^u$ can be varied independently inside and outside the spin scalar product. Therefore, it suffices to consider variations inside the spin scalar product. Thus the vanishing of the first variation (4.9) yields the condition

$$0 = 2 \iint_{M\times M} \prec\delta\psi^u(x)\,|\,Q(x,y)\,\psi^u(y)\succ_x - \lambda \int_M \prec\delta\psi^u(x)\,|\,\psi^u(x)\succ_x\,.$$

Since the variation $\delta\psi^u$ is arbitrary (within the class of smooth and compactly supported variations), the result follows. □

We remark that the kernel $Q(x,y)$ also gives rise to an operator on the one-particle Krein space $(\mathcal{K}, <.|.>)$ as introduced in §1.5. Thus, in analogy to (1.24), one sets

$$Q\;:\;\mathcal{D}(Q) \subset \mathcal{K} \to \mathcal{K}\,, \qquad (Q\psi)(x) = \int_M Q(x,y)\,\psi(y)\,d\rho(y)\,,$$

where the domain $\mathcal{D}(Q)$ can be chosen for example as the continuous wave functions with compact support. The symmetry property of the kernel (4.7) implies that the operator Q is symmetric on the Krein space $(\mathcal{K}, <.|.>)$. The equation (4.8) can be written in a compact form as the operator equation

$$\big(2Q - \lambda\mathbb{1}\big)\,\Psi = 0 \tag{4.10}$$

(where Ψ is again the wave evaluation operator (1.18)). In words, this equation means that the operator $(2Q - \lambda\mathbb{1})$ vanishes on the physical wave functions. However, the operator equation (4.10) is not satisfying mathematically because the physical wave

functions in the image of Ψ are in general not vectors of the Krein space $(\mathcal{K}, <.|.>)$ (see §1.5). Nevertheless, (4.10) is useful as a short notation for the EL equations (4.8).

Remark 4.3. (more general variations) Clearly, only a special class of variations of the universal measure can be described by variations of the physical wave functions. As a consequence, the resulting EL equations (4.8) are only *necessary* conditions for ρ to be a critical point of the action (4.4). We now explain how these necessary conditions are related to the stronger EL equations as derived in [1].

As an example of variations which are not covered by the ansatz (4.1), one can multiply the universal measure by weight functions

$$d\rho_\tau = f_\tau \, d\rho \,, \tag{4.11}$$

where $(f_\tau)_{\tau \in (-\delta, \delta)}$ is a family of non-negative functions which are integrable and have mean zero, i.e.

$$f_\tau \geq 0 \qquad \text{and} \qquad \int_M f_\tau \, d\rho = 0 \,.$$

Computing first variations of the action (4.4) gives rise to the equation

$$2 \int_M \mathcal{L}_\kappa(x,y) \, d\rho(y) + \lambda \, \mathrm{tr}(x) = \text{const} \qquad \text{on } M \,. \tag{4.12}$$

This is an additional EL equation which minimizers of the causal action principle must satisfy. It turns out that in the limiting case of an interacting system in Minkowski space (to be discussed in §5.2 and §5.3 below), this equation can be satisfied simply by a rescaling of the local correlation operators.

Variations of the physical wave functions as well as variations of the form (4.11) have the property that the support of the universal measure changes continuously (in the sense that for every compact set $K \subset \mathcal{F}$ and every open neighborhood U of $K \cap \text{supp}\, \rho$ there is $\varepsilon > 0$ such that $\text{supp}\, \rho_\tau \cap K \subset U$ for all τ with $|\tau| < \varepsilon$). Such variations can be regarded as the analogs of variations of the potentials, the metric or the wave functions in classical field theory or quantum mechanics. However, in the setting of causal fermion systems there are also more general smooth variations for which the support of the measure ρ_τ changes discontinuously. A typical example is to let ρ be a bounded measure and to set

$$\rho_\tau = (1 - \tau^2)\, \rho + \tau^2 \, \rho(\mathcal{F}) \, \delta_x \,, \tag{4.13}$$

where δ_x is the Dirac measure supported at $x \notin \text{supp}\, \rho$. The EL equations corresponding to such variations have a different mathematical structure, which we cannot explain in detail here. Generally speaking, for interacting systems in Minkowski space, the EL equations of Proposition 4.2 give rise to an effective interaction via *classical* gauge fields (this so-called *continuum limit* will be discussed in §5.2). The EL equations corresponding to more general variations like (4.13), however, give rise to an effective interaction via bosonic *quantum* fields. We will come back to this point in §5.3. ♢

4.2. **Symmetries and Conserved Surface Layer Integrals.** In [21] it is shown that symmetries of the Lagrangian give rise to conservation laws. These results can be understood as adaptations of Noether's theorem to the causal action principle. Since the mathematical structure of the causal action principle is quite different from that of the Lagrangian formulation of classical field theory, these adaptations are not straightforward. We now explain a few concepts and results from [21] which are important for understanding the general physical picture.

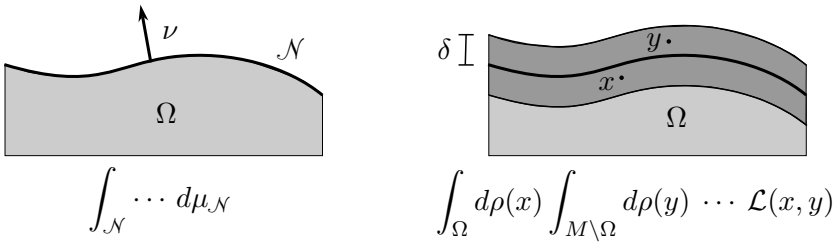

FIGURE 1. A surface integral and a corresponding surface layer integral.

We first recall that the conservation laws obtained from the classical Noether theorem state that the integral of a certain density over a Cauchy surface $\mathcal{N}$ does not depend on the choice of $\mathcal{N}$. For example, charge conservation states that the spatial integral of the charge density gives a constant. As another example, energy conservation states that in a static space-time background, the integral of the energy density is a constant. In general terms, the conserved quantities are spatial integrals over a Cauchy surface $\mathcal{N}$ (see the left of Figure 1). In the setting of causal fermion systems, it is unclear how such surface integrals should be defined, in particular because we do not have a measure on hypersurfaces and because it is not clear what the normal ν on the hypersurface should be. This is the reason why in the Noether-like theorems in [21] one works instead of surface integrals with so-called *surface layer integrals* where one integrates over a boundary layer of a set $\Omega \subset M$ (see the right of Figure 1). The width δ of this layer is the length scale on which $\mathcal{L}(x,y)$ decays. For a system composed of Dirac particles (similar as explained in Section 2 for the Minkowski vacuum and in §5.2 for interacting systems), this length scale can be identified with the *Compton scale* $\sim m^{-1}$ of the Dirac particles. Thus the width of the surface layer is a non-zero macroscopic length scale. In particular, the surface layer integrals cannot be identified with or considered as a generalization of the surface integrals of the classical Noether theorem. However, in most situations of interest, when the surface N is almost flat on the Compton scale (like for a spatial hyperplane in Minkowski space), the surface layer integral can be well-approximated by a corresponding surface integral. It turns out that in this limiting case, the conservation laws obtained from the Noether-like theorems in [21] go over to corresponding classical conservation laws.

From the conceptual point of view, the most interesting conservation law is *charge conservation*. In order to construct the underlying symmetry, we let $\mathcal{A}$ be a bounded symmetric operator on $\mathcal{H}$ and let

$$\mathcal{U}_\tau := \exp(i\tau \mathcal{A})$$

be the corresponding one-parameter family of unitary transformations. We introduce the family of transformations

$$\Phi_\tau : \mathcal{F} \to \mathcal{F}, \qquad \Phi_\tau(x) = \mathcal{U}_\tau\, x\, \mathcal{U}_\tau^{-1}.$$

Since the Lagrangian is defined via the spectrum of operators on $\mathcal{H}$, it clearly remains unchanged if all operators are unitarily transformed, i.e.

$$\mathcal{L}\big(\Phi_\tau(x), \Phi_\tau(y)\big) = \mathcal{L}(x,y). \tag{4.14}$$

In other words, the transformations Φ_τ describe a *symmetry of the Lagrangian*. Next, one constructs a corresponding one-family of universal measures by taking the push-forward,

$$\rho_\tau := (\Phi_\tau)_* \rho \,.$$

As a consequence of the symmetry (4.14), this variation of the universal measure leaves the action invariant. Under suitable differentiability assumptions, this symmetry gives rise to the identity

$$\frac{d}{d\tau} \int_\Omega d\rho(x) \int_{M \setminus \Omega} d\rho(y) \Big(\mathcal{L}\big(\Phi_\tau(x), y\big) - \mathcal{L}\big(\Phi_{-\tau}(x), y\big) \Big) \Big|_{\tau=0} = 0 \,, \tag{4.15}$$

valid for any compact subset $\Omega \subset M$.

We now explain how the identity (4.15) is related to a conservation law. To this end, for simplicity we consider a system in Minkowski space (similar as explained for the vacuum in Section 2) and choose a sequence of compact sets Ω_n which exhaust the region between two Cauchy surfaces at times $t = t_0$ and $t = t_1$. Then the surface layer integral (4.15) reduces to the difference of integrals over surface layers at times $t \approx t_0$ and $t \approx t_1$. Next, we choose $\mathcal{A} = \pi_{\langle u \rangle}$ as the projection operator on the one-dimensional subspace generated by a vector $u \in \mathcal{H}$. Then in the limit $\varepsilon \searrow 0$ in which the UV regularization is removed, the resulting surface layer integral at time $t \approx t_0$ reduces to the integral

$$\int_{\mathbb{R}^3} \prec u(t_0, \vec{x}) \,|\, \gamma^0 u(t_0, \vec{x}) \succ_{(t_0, \vec{x})} d^3x \,,$$

thereby reproducing the probability integral in Dirac theory. As a consequence, the representation of the scalar product $\langle . | . \rangle_{\mathcal{H}}$ as an integral over a Cauchy surface (2.2) has a natural generalization to the setting of causal fermion systems, if the surface integral is replaced by a corresponding surface layer integral. This result also shows that the spatial normalization of the fermionic projector (where one works with spatial integrals of the form (2.23); for details see [28]) really is the correct normalization method which reflects the intrinsic conservation laws of the causal fermion system.

The conservation laws in [21] also give rise to the *conservation of energy and momentum*, as we now outline. In the classical Noether theorem, these conservation laws are a consequence of space-time symmetries as described most conveniently using the notion of Killing fields. Therefore, one must extend this notion to the setting of causal fermion systems. Before explaining how this can be accomplished, we recall the procedure in the classical Noether theorem: In the notion of a Killing field, one distinguishes the background geometry from the additional particles and fields. The background geometry must have a symmetry as described by the Killing equation. The additional particles and fields, however, do not need to have any symmetries. Nevertheless, one can construct a symmetry of the whole system by actively transporting the particles and fields along the flow lines of the Killing field. The conservation law corresponding to this symmetry transformation gives rise to the conservation of energy and momentum.

In a causal fermion system, there is no clear-cut distinction between the background geometry and the particles and fields of the system, because all of these structures are encoded in the underlying causal fermion system and mutually depend on each other. Therefore, instead of working with a symmetry of the background geometry, we work with the notion of an approximate symmetry. By actively transforming those physical wave functions which do not respect the symmetry, such an approximate

symmetry again gives rise to an exact symmetry transformation, to which the Noether-like theorems in [21] can be applied. More precisely, one begins with a C^1-family of transformations $(f_\tau)_{\tau \in (-\delta, \delta)}$ of space-time,

$$f_\tau \,:\, M \to M \qquad \text{with} \qquad f_0 = \mathbb{1}\,, \tag{4.16}$$

which preserve the universal measure in the sense that $(f_\tau)_* \rho = \rho$. The family (f_τ) can be regarded as the analog of a flow in space-time along a classical Killing field. Moreover, one considers a family of unitary transformations $(\mathcal{U}_\tau)_{\tau \in (-\delta, \delta)}$ on $\mathcal{H}$ with the property that

$$\mathcal{U}_{-\tau}\, \mathcal{U}_\tau = \mathbb{1} \qquad \text{for all } \tau \in (-\delta, \delta)\,.$$

Combining these transformations should give rise to an *approximate symmetry* of the wave evaluation operator (1.18) in the sense that if we compare the transformation of the space-time point with the unitary transformation by setting

$$E_\tau(u, x) := (\Psi u)\big(f_\tau(x)\big) - (\Psi \mathcal{U}_\tau^{-1} u)(x) \qquad (x \in M, u \in \mathcal{H})\,, \tag{4.17}$$

then the operator $E_\tau : \mathcal{H} \to C^0(M, SM)$ should be sufficiently small. Here "small" means for example that E vanishes on the orthogonal complement of a finite-dimensional subspace of $\mathcal{H}$; for details see [21, Section 6]. Introducing the variation Φ_τ by

$$\Phi_\tau \,:\, M \to \mathcal{F}\,, \qquad \Phi_\tau(x) = \mathcal{U}_\tau\, x\, \mathcal{U}_\tau^{-1}\,,$$

we again obtain a symmetry of the Lagrangian (4.14). This gives rise to conserved surface layer integrals of the form (4.15). In order to bring these surface layer integrals into a computable form, one decomposes the first variation of Φ_τ as

$$\delta\Phi(x) := \partial_\tau \Phi_\tau(x)\big|_{\tau=0} = \delta f(x) + v(x)\,, \tag{4.18}$$

where δf is the first variation of f_τ, (4.16), and $v(x)$ is a vector field on $\mathcal{F}$ along M which is transversal to $M \subset \mathcal{F}$. Expressing v in terms of the operator E in (4.17) shows that v is again small, making it possible to compute the corresponding variation of the Lagrangian in (4.15). We remark that in the decomposition (4.18), the vector field δf describes a transformation of the space-time points. The vector field v, however, can be understood as an active transformation of all the objects in space-time which do *not* have the space-time symmetry (similar as described above for the parallel transport of the particles and fields along the flow lines of the Killing field in the classical Noether theorem).

In order to get the connection to classical conservation laws, one again studies a system in Minkowski space and considers the limiting case where a sequence Ω_n exhausts the region between two Cauchy surfaces at times $t = t_0$ and $t = t_1$. In this limiting case, the conserved surface layer integral reduces to the surface integral

$$\int_{\mathbb{R}^3} T_{i0}\, K^i\, d^3x\,,$$

where T_{ij} is the energy-momentum tensor of the Dirac particles and $K = \delta f$ is a Killing field. This shows that the conservation of energy and momentum is a special case of more general conservation laws which are intrinsic to causal fermion systems.

4.3. **The Initial Value Problem and Time Evolution.** In order to get a better understanding of the dynamics described by the causal action principle, it is an important task to analyze the initial value problem. The obvious questions are: What is the initial data? Is it clear that a solution exists? Is the solution unique? How do solutions look like? Giving general answers to these questions is a difficult mathematical problem. In order to evaluate the difficulties, one should recall that ρ describes space-time as well as all structures therein. Therefore, similar as in the Cauchy problem for the Einstein equations, solving the initial value problem involves finding the geometry of space-time together with the dynamics of all particles and fields. In view of the complexity of this problem, the only results known at present are contained in the paper [16], where an initial value problem is formulated and some existence and uniqueness theorems are proven. We now review a few methods and results of this paper. Moreover, at the end of this section we mention an approach proposed in [19] for obtaining more explicit information on the dynamics by analyzing perturbations of a given minimizing measure.

Since the analysis of the causal action principle is technically demanding, in [16] one considers instead so-called *causal variational principles in the compact setting*. In order to get into this simplified setting, one replaces $\mathcal{F}$ by a compact metric space (or a smooth manifold). The Lagrangian is replaced by a non-negative continuous function $\mathcal{L} \in C^{0,1}(\mathcal{F} \times \mathcal{F}, \mathbb{R}^+_0)$ which is symmetric in its two arguments. Similar to (1.2) one minimizes the action

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

in the class of all normalized regular Borel measures on $\mathcal{F}$, but now leaving out the constraints (1.4) and (1.5). Space-time is again defined by $M := \operatorname{supp} \rho$. The resulting causal structure is defined by saying that two space-time points $x, y \in M$ are called *timelike* separated if $\mathcal{L}(x,y) > 0$, and *spacelike* separated if $\mathcal{L}(x,y) = 0$. Clearly, in this setting there are no wave functions. Nevertheless, causal variational principles in the compact setting incorporate basic features of the causal action principle and are therefore a good starting point for the analysis (for a more detailed introduction and structural results on the minimizing measures see [26]).

When solving the classical Cauchy problem, instead of searching for a global solution, it is often easier to look for a local solution around a given initial value surface. This concept of a local solution also reflects the common physical situation where the physical system under consideration is only a small subsystem of the whole universe. With this in mind, we would like to "localize" the variational principle to a subset $\mathfrak{I} \subset \mathcal{F}$, referred to as the *inner region*. There is the complication that the Lagrangian $\mathcal{L}(x,y)$ is nonlocal in the sense that it may be non-zero for points $x \in \mathfrak{I}$ and $y \in \mathcal{F} \setminus \mathfrak{I}$. In order to take this effect into account, one describes the influence of the "outer region" $\mathcal{F} \setminus \mathfrak{I}$ by a so-called *external potential* $\phi : \mathcal{F} \to \mathbb{R}^+_0$. In the limiting case when the outer region becomes large, this gives rise to the so-called *inner variational principle*, where the action defined by

$$\mathcal{S}_{\mathfrak{I}}[\rho, \phi] = \iint_{\mathfrak{I} \times \mathfrak{I}} \mathcal{L}(x,y)\, d\rho(x)\, d\rho(y) + 2 \int_{\mathfrak{I}} \big(\phi(x) - \mathfrak{s}\big)\, d\rho(x) \tag{4.19}$$

is minimized under variations of ρ in the class of regular Borel measures on $\mathfrak{I}$ (not necessarily normalized because the volume constraint is now taken care of by the corresponding Lagrange parameter $\mathfrak{s} > 0$).

The *initial values* are described by a regular Borel measure ρ_0 (which is to be thought of as the universal measure restricted to a time slice around the initial value surface in space-time). The *initial conditions* are implemented by demanding that

$$\rho \geq \rho_0 \,. \tag{4.20}$$

The naive method of minimizing (4.19) under the constraint (4.20) is not a sensible concept because the constraint (4.20) would give rise to undesirable Lagrange multiplier terms in the EL equations. Instead, one minimizes (4.19) without constraints, but chooses the external potential ϕ in such a way that the minimizing measure satisfies the initial values (4.20). It turns out that this procedure does not determine the external potential uniquely. Therefore, the method proposed in [16] is to *optimize the external potential* by making it in a suitable sense "small." As is made precise in [16] in various situations, the resulting interplay between minimizing the action and optimizing the external potential gives rise to unique solutions of the initial-value problem with an optimal external potential.

We point out that, due to the mathematical simplifications made, the results in [16] do not apply to physically interesting situations like the initial value problem for interacting Dirac sea configurations. Moreover, the methods in [16] do not seem to give explicit information on the dynamics of causal fermion systems. Therefore, it is a promising complementary approach to consider perturbations of a given minimizing measure (which should describe the "vacuum configuration") and to analyze the dynamics of the perturbations by studying the resulting EL equations. This approach is pursued in [19] in the following way. In order to describe the perturbations of the minimizing measure ρ, one considers smooth variations for which the support of ρ changes continuously. Combining (4.1) and (4.11), these variations can be written as

$$\tilde{\rho}_\tau = (F_\tau)_* \big(f_\tau\, \rho\big)$$

with a family of mappings $F_\tau : M \to \mathcal{F}$ and a family of non-negative functions f_τ. Expanding in powers of τ, these variations can be described conveniently in terms of sections of *jet bundles* over M. The EL equations yield conditions on the jets, which can be rewritten as dynamical equations in space-time.

5. Limiting Cases

We now discuss different limiting cases of causal fermion systems.

5.1. The Quasi-Free Dirac Field and Hadamard States.

We now turn attention to interacting systems. The simplest interaction is obtained by inserting an *external potential* into the Dirac equation (2.1),

$$\big(i\gamma^j \partial_j + \mathcal{B} - m\big)\, \psi(x) = 0 \,. \tag{5.1}$$

Another situation of physical interest is to consider the Dirac equation in an external classical gravitational field as described mathematically by a globally hyperbolic Lorentzian manifold $(\mathscr{M}, g)$. In this section, we explain how the methods and results of Section 2 generalize to the situation when an external field is present. This will also give a connection to quasi-free Dirac fields and Hadamard states. In order to keep the explanations as simple as possible, we here restrict attention to an external potential $\mathcal{B}$ in Minkowski space, but remark that many methods and results could or have been worked out also in the presence of a gravitational field.

The obvious conceptual difficulty when extending the constructions of Section 2 is that one no longer has the notion of "negative-frequency solutions" which were essential for introducing Dirac sea configurations (see Lemma 2.8). In order to overcome this difficulty, one needs to decompose the solution space of the Dirac equation (5.1) into two subspace, in such a way that without external potential the two subspaces reduce to the subspaces of positive and negative frequency. This *external field problem* was solved perturbatively in [7, 14] and non-perturbatively in [25, 24, 23] (for a more detailed exposition see [8, §2.1]).

We now briefly outline the non-perturbative treatment, which relies on the construction on the so-called *fermionic signature operator*. Choosing again the scalar product (2.2), the solution space of the Dirac equation (5.1) forms a Hilbert space denoted by $(\mathcal{H}_m, (.|.)_m)$. Moreover, on the Dirac wave functions (not necessarily solutions of the Dirac equations) one may introduce a dual pairing by integrating the spin scalar product over all of space-time,

$$<.|.> \,:\, C^\infty(\mathscr{M}, S\mathscr{M}) \times C_0^\infty(\mathscr{M}, S\mathscr{M}) \to \mathbb{C}\,, \quad <\psi|\phi> = \int_{\mathscr{M}} \prec\psi|\phi\succ_x d^4x\,. \tag{5.2}$$

The basic idea is to extend this dual pairing to a bilinear form on the Hilbert space $\mathcal{H}_m$ and to represent this bilinear form in terms of the Hilbert space scalar product

$$<\phi_m|\psi_m> = (\phi_m \,|\, \mathcal{S}\, \psi_m)_m\,.$$

If $\mathscr{M}$ is a space-time of *finite lifetime*, this construction can indeed be carried out and defines the *fermionic signature operator* $\mathcal{S}$ being a bounded symmetric operator on $\mathcal{H}_m$ (see [25]). The positive and negative spectral subspaces of $\mathcal{S}$ give the desired decomposition of $\mathcal{H}_m$ into two subspaces. We remark that the fermionic signature operator makes it possible to study *spectral geometry* for Lorentzian signature (see [22] and [12] for the connection to index theory).

In space-times of infinite lifetime like Minkowski space, the above method does not work because (5.2) does not extend to a continuous bilinear form on $\mathcal{H}_m \times \mathcal{H}_m$. The underlying problem is that the time integral in (5.2) in general diverges for solutions of the Dirac equation. In order to circumvent this problem, one considers families of Dirac solutions $(\psi_m)_{m\in I}$ (for an open interval $I = (m_a, m_b) \subset (0,\infty)$) and makes use of the fact that integrating over the mass parameter generates decay of the wave functions for large times (for details see [24]). As a result, one can make sense of the equation

$$<\int_I \psi_m\, dm \,|\, \int_I \psi_{m'}\, dm'> = \int_I (\psi_m \,|\, \mathcal{S}_m\, \phi_m)_m\, dm\,,$$

which uniquely defines a family of bounded symmetric operators $(\mathcal{S}_m)_{m\in I}$. Now the positive and negative spectral subspaces of the operator $\mathcal{S}_m$ again give the desired decomposition of $\mathcal{H}_m$ into two subspaces.

Having decomposed the solution space, one may choose the Hilbert space $\mathcal{H}$ of the causal fermion system as one of the two subspaces of the solution space. Choosing an orthonormal basis (u_ℓ) of $\mathcal{H}$ and introducing the unregularized kernel of the fermionic projector again by (2.19), one obtains a two-point distribution $P(x,y)$. Using that this two-point distribution comes from a projection operator in the Hilbert space $\mathcal{H}_m$, there is a canonical construction which gives a quasi-free Dirac field together with a Fock representation such that the two-point distribution coincides with $P(x,y)$. In the language of algebraic quantum field theory, this result is stated as follows (see [23, Theorem 1.4]):

Theorem 5.1. *There is an algebra of smeared fields generated by $\Psi(g)$, $\Psi^*(f)$ together with a quasi-free state ω with the following properties:*

(a) The canonical anti-commutation relations hold:

$$\{\Psi(g), \Psi^*(f)\} = <g^* \,|\, \tilde{k}_m\, f> , \qquad \{\Psi(g), \Psi(g')\} = 0 = \{\Psi^*(f), \Psi^*(f')\} .$$

(b) The two-point function of the state is given by

$$\omega\big(\Psi(g)\, \Psi^*(f)\big) = - \iint_{\mathscr{M} \times \mathscr{M}} g(x) P(x,y) f(y)\, d^4x\, d^4y .$$

This theorem means that before introducing an UV regularization, the description of the Dirac system using the fermionic projector is equivalent to the usual description of a quasi-free Dirac field in quantum field theory.

Moreover, it is shown in [23] that the two-point distribution $P(x,y)$ is of *Hadamard form*, provided that $\mathcal{B}$ is smooth, not too large and decays faster than quadratically for large times (for details see [23, Theorem 1.3] and the references in this paper). This result implies that the representation of the quasi-free Dirac field as obtained from the fermionic projector is a suitable starting point for a perturbative treatment of the resulting interacting theory (see for example [4]).

In our context, the fact that $P(x,y)$ is of Hadamard form implies that that the results in §1.2 also apply in the presence of an external potential, as we now explain. The Hadamard property means in words that the bi-distribution $P(x,y)$ in the presence of the external potential has the same singularity structure as in the Minkowski vacuum. As a consequence, the arguments in §1.2 remain true if the points x and y are sufficiently close to each other. More precisely, the relevant length scale is given by the inverse of the amplitude $|\mathcal{B}(x)|^{-1}$ of the external potential. On the other hand, the separation of the points x and y must be larger than the scale ε on which regularization effects come into play. Therefore, the causal structure of a causal fermion system agrees with that of Minkowski space on the scale $\varepsilon \ll \big|x^0 - y^0\big| + \big|\vec{x} - \vec{y}\big| \ll |\mathcal{B}|^{-1}$ (where $|\mathcal{B}|$ is any matrix norm). Thinking of ε as being at least as small as the Planck length, in most situations of interest the lower bound is no restriction. The upper bound is also unproblematic because the causal structure on the macroscopic scale can still be recovered by considering paths in space-time and subdividing the path on a scale $\delta \ll |\mathcal{B}|^{-1}$ (similar as explained in [15, Section 4.4] for the spin connection). With this in mind, we conclude that the causal structure of a causal fermion system indeed agrees with that of Minkowski space, even in the presence of an external potential.

5.2. **Effective Interaction via Classical Gauge Fields.** We now outline how to describe interacting systems in Minkowski space by analyzing the EL equations corresponding to the causal action principle as worked out in Proposition 4.2. In this so-called *continuum limit* the interaction is described by classical gauge fields. For brevity, we can only explain a few basic concepts and refer the interested reader to the detailed computations in the book [6].

Let us begin with the Minkowski vacuum. As shown in §2.2, regularizing a vacuum Dirac sea configuration gives rise to a causal fermion system $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$. Moreover, we saw in the following sections §2.3–§2.4 that the inherent structures of the causal fermion system can be identified with those of Minkowski space (in particular, see (2.13) as well

as Propositions 2.6 and 2.7). This makes it possible to write the EL equations (4.8) as

$$\int_{\mathscr{M}} Q^\varepsilon(x,y)\,\big(\mathfrak{R}_\varepsilon u_\ell\big)(y)\,d^4y = \frac{\lambda}{2}\,\big(\mathfrak{R}_\varepsilon u_\ell\big)(x) \qquad \text{for all } u \in \mathcal{H}\,, \tag{5.3}$$

where the regularized kernel $Q^\varepsilon(x,y)$ is again defined via (4.6) as the derivative of the Lagrangian. Next, one chooses the Hilbert space $\mathcal{H}$ as in §2.5 as the Dirac sea configuration formed of all negative-energy solutions of the Dirac equation. Then $P^\varepsilon(x,y)$ can be computed explicitly by regularizing the distribution $P(x,y)$ as given in momentum space by (2.22) and in position space by (2.24) and Lemma 2.9. Computing $Q^\varepsilon(x,y)$, it turns out that the EL equations are mathematically well-defined if the convolution integral in (5.3) is rewritten with the help of Plancherel's theorem as a multiplication in momentum space. The analysis of the *continuum limit* gives a procedure for studying these equations in the asymptotics $\varepsilon \searrow 0$ when the regularization is removed. The effective equations obtained in this asymptotic limit are evaluated most conveniently in a formalism in which the unknown microscopic structure of space-time (as described by the regularization) enters only in terms of a finite (typically small) number of so-called *regularization parameters*. According to the method of variable regularization (see Remark 2.1), one needs to analyze the dependence of the regularization parameters in detail. It turns out that the causal fermion systems obtained from the vacuum Dirac sea configuration satisfy the EL equations in the continuum limit, for any choice of the regularization parameters.

The first step towards interacting systems is to consider systems involving *particles* and/or *anti-particles*. To this end, one simply modifies the constructions in §2.5 by choosing the Hilbert space $\mathcal{H}$ differently. Namely, instead of choosing all negative-energy solutions, one chooses $\mathcal{H}$ as a subspace of the solution space which differs from the space of all negative-energy solutions by a finite-dimensional subspace. In other words, $\mathcal{H}$ is obtained from the space of all negative-energy solutions by taking out a finite number n_a of states and by adding a finite number of states n_p of positive energy. Thus, denoting the regularized kernel of the fermionic projector of the Minkowski vacuum for clarity by $P^\varepsilon_\text{sea}(x,y)$, the kernel of the fermionic projector (2.18) can be written as

$$P^\varepsilon(x,y) = P^\varepsilon_\text{sea}(x,y) - \sum_{k=1}^{n_\text{p}} \big(\mathfrak{R}_\varepsilon \psi_k\big)(x)\overline{\big(\mathfrak{R}_\varepsilon \psi_k\big)(y)} + \sum_{l=1}^{n_\text{a}} \big(\mathfrak{R}_\varepsilon \phi_l\big)(x)\overline{\big(\mathfrak{R}_\varepsilon \phi_l\big)(y)}\,, \tag{5.4}$$

where ψ_k and ϕ_l are suitably normalized bases of the particle and anti-particle states, respectively. In this procedure, we again take Dirac's concept of a "sea" of particles literally and describe particles and anti-particles by occupying positive-energy states and creating "holes" in the Dirac sea, respectively. We also remark that the construction (5.4) modifies the kernel of the fermionic projector only by smooth contributions and thus preserves the singularity structure of $P^\varepsilon(x,y)$ as $\varepsilon \searrow 0$. As a consequence, the correspondence of the inherent structures of the causal fermion systems to the structures in Minkowski space remains unchanged (just as explained at the end of §5.1 for an external potential).

According to (5.4), the particle and anti-particle states modify the kernel of the fermionic projector. It turns out that this has the effect that the EL equations in the continuum limit no longer hold. In order to again satisfy these equations, we need to introduce an interaction. In mathematical terms, this means that the universal measure ρ must be modified. The basic question is how to modify the universal

measure in such a way that the EL equations in the continuum limit again hold. It turns out that it is a useful first step to insert an external potential $\mathcal{B}$ into the Dirac equation (2.1) by going over to the Dirac equation (5.1). Choosing $\mathcal{H}$ as a subspace of the solution space of this Dirac equation, the constructions of Section 2 again apply and give rise to causal fermion systems $(\mathcal{H}, \mathcal{F}, \rho^\varepsilon)$. The potential $\mathcal{B}$ modifies the dynamics of all physical wave functions in a collective way. Now one can ask the question whether the resulting causal fermion systems satisfy the EL equations in the continuum limit. It turns out that this is the case if and only if the potential $\mathcal{B}$ satisfies certain equations, which can be identified with classical field equations for the potential $\mathcal{B}$. In this way, the causal action principle gives rise to classical field equations. In order to make our concepts clear, we point out that the potential $\mathcal{B}$ merely is a convenient device in order to describe the collective behavior of all physical wave functions. It should not be considered as a fundamental object of the theory. We also note that, in order to describe variations of the physical wave functions, the potential in (5.1) can be chosen arbitrarily (in particular, the potential does not need to satisfy any field equations). Each choice of $\mathcal{B}$ describes a different variation of the physical wave functions. It is the EL equations in the continuum limit which single out the physically admissible potentials as being those which satisfy the field equations.

Before going on, we briefly explain how the subspace $\mathcal{H}$ is chosen. Clearly, the Dirac equation (5.1) cannot in general be solved in closed form. Therefore, for an explicit analysis one must use perturbative methods. When performing the perturbation expansion, one must be careful about the proper normalization of the fermionic states (in the sense that spatial integrals of the form (2.23) should be preserved). Moreover, one must make sure that the singular structure of $P(x, y)$ in position space is compatible with the causal action principle (meaning that the light-cone expansion of $P(x, y)$ only involves bounded integrals of $\mathcal{B}$ and its derivatives). Satisfying these two requirements leads to the *causal perturbation expansion* (see [28] and the references therein). We also mention that regularizing the perturbation expansion is a delicate issue. This can already be understood for the simple regularization by mollification in Example 2.4, in which case it is not clear whether one should first mollify and then introduce the interaction or vice versa. The correct method for regularizing the perturbation expansion is obtained by demanding that the behavior under gauge transformations should be preserved by the regularization. This leads to the *regularized causal perturbation expansion* as developed in [8, Appendix D] and [6, Appendix F].

We proceed with a brief overview of the results of the analysis of the continuum limit. In [6] the continuum limit is worked out in several steps beginning from simple systems and ending with a system realizing the fermion configuration of the standard model. For each of these systems, the continuum limit gives rise to effective equations for second-quantized fermion fields coupled to classical bosonic gauge fields (for the connection to second-quantized bosonic fields see §5.3 below). To explain the structure of the obtained results, it is preferable to first describe the system modelling the leptons as analyzed in [6, Chapter 4]. The input to this model is the configuration of the leptons in the standard model without interaction. Thus the fermionic projector of the vacuum is assumed to be composed of three generations of Dirac particles of masses $m_1, m_2, m_3 > 0$ (describing e, μ, τ) as well as three generations of Dirac particles of masses $\tilde{m}_1, \tilde{m}_2, \tilde{m}_3 \geq 0$ (describing the corresponding neutrinos). Furthermore,

we assume that the regularization of the neutrinos breaks the chiral symmetry (implying that we only see their left-handed components). We point out that the definition of the model does not involve any assumptions on the interaction.

The detailed analysis in [6, Chapter 4] reveals that the effective interaction in the continuum limit has the following structure. The fermions satisfy the Dirac equation coupled to a left-handed SU(2)-gauge potential $A_L = \big(A_L^{ij}\big)_{i,j=1,2}$,

$$\left[i\partial\!\!\!/ + \begin{pmatrix} A\!\!\!/_L^{11} & A\!\!\!/_L^{12}\, U_{\text{MNS}}^* \\ A\!\!\!/_L^{21}\, U_{\text{MNS}} & -A\!\!\!/_L^{11} \end{pmatrix} \chi_L - mY \right] \psi = 0\,,$$

where we used a block matrix notation (in which the matrix entries are 3×3-matrices). Here mY is a diagonal matrix composed of the fermion masses,

$$mY = \text{diag}(\tilde{m}_1, \tilde{m}_2, \tilde{m}_3,\ m_1, m_2, m_3)\,, \tag{5.5}$$

and U_{MNS} is a unitary 3×3-matrix (taking the role of the Maki-Nakagawa-Sakata matrix in the standard model). The gauge potentials A_L satisfy a classical Yang-Mills-type equation, coupled to the fermions. More precisely, writing the isospin dependence of the gauge potentials according to $A_L = \sum_{\alpha=1}^3 A_L^\alpha \sigma^\alpha$ in terms of Pauli matrices, we obtain the field equations

$$\partial^k \partial_l (A_L^\alpha)^l - \Box (A_L^\alpha)^k - M_\alpha^2\, (A_L^\alpha)^k = c_\alpha\, \overline{\psi}\big(\chi_L \gamma^k\, \sigma^\alpha\big)\psi\,, \tag{5.6}$$

valid for $\alpha = 1,2,3$ (for notational simplicity, we wrote the Dirac current for one Dirac particle; for a second-quantized Dirac field, this current is to be replaced by the expectation value of the corresponding fermionic field operators). Here M_α are the bosonic masses and c_α the corresponding coupling constants. The masses and coupling constants of the two off-diagonal components are equal, i.e. $M_1 = M_2$ and $c_1 = c_2$, but they may be different from the mass and coupling constant of the diagonal component $\alpha = 3$. Generally speaking, the mass ratios M_1/m_1, M_3/m_1 as well as the coupling constants c_1, c_3 depend on the regularization. For a given regularization, they are computable.

Finally, our model involves a gravitational field described by the Einstein equations

$$R_{jk} - \frac{1}{2}\, R\, g_{jk} + \Lambda\, g_{jk} = \kappa\, T_{jk}\,, \tag{5.7}$$

where R_{jk} denotes the Ricci tensor, R is scalar curvature, and T_{jk} is the energy-momentum tensor of the Dirac field. Moreover, κ and Λ denote the gravitational and the cosmological constants, respectively. We find that the gravitational constant scales like $\kappa \sim \delta^{-2}$, where $\delta \geq \varepsilon$ is the length scale on which the chiral symmetry is broken.

In [6, Chapter 5] a system is analyzed which realizes the configuration of the leptons and quarks in the standard model. The result is that the field equation (5.6) is replaced by field equations for the electroweak and strong interactions after spontaneous symmetry breaking (the dynamics of the corresponding Higgs field has not yet been analyzed). Furthermore, the system again involves gravity (5.7).

A few clarifying remarks are in order. First, the above field equations come with corrections which for brevity we cannot discuss here (see [6, Sections 3.8, 4.4 and 4.6]). Next, it is worth noting that, although the states of the Dirac sea are explicitly taken into account in our analysis, they do not enter the field equations. More specifically, in a perturbative treatment, the divergences of the Feynman diagram describing the vacuum polarization drop out of the EL equations of the causal action. Similarly, the naive "infinite negative energy density" of the sea drops out of the Einstein equations,

making it unnecessary to subtract any counter terms. We finally remark that the only free parameters of the theory are the masses in (5.5) as well as the parameter δ which determines the gravitational constant. The coupling constants, the bosonic masses and the mixing matrices are functions of the regularization parameters which are unknown due to our present lack of knowledge on the microscopic structure of space-time. The regularization parameters cannot be chosen arbitrarily because they must satisfy certain relations. But except for these constraints, the regularization parameters are currently treated as free empirical parameters.

To summarize, the dynamics in the continuum limit is described by Dirac spinors coupled to classical gauge fields and gravity. The effective continuum theory is manifestly covariant under general coordinate transformations. The only limitation of the continuum limit is that the bosonic fields are merely classical. We shall come back to second-quantized bosonic fields in §5.3 below.

5.3. **Effective Interaction via Bosonic Quantum Fields.** In §5.2 it was outlined that and in which sense the regularized Dirac sea vacuum satisfies the EL equations (4.8). In simple terms, these results mean that the regularized Dirac sea vacuum is a critical point of the causal action under variations of the physical wave functions (see Definition 4.1). We now explain why the regularized Dirac sea vacuum is *not a minimizer* of the causal action principle. This argument will lead us to a method for further decreasing the causal action. It also gives some insight on the structure of the minimizing measure. In particular, we shall see that the effective interaction in the resulting space-time is to be described effectively by bosonic *quantum* fields.

Suppose that $(\mathcal{H}, \mathcal{F}, \rho)$ is a causal fermion system describing a regularized Dirac sea configuration (see §2.5). In order to explain the basic idea, it suffices to consider the case that ρ has finite total volume (which can be arranged for example by considering the system in a four-dimensional box). For a unitary transformation $V \in \mathrm{U}(\mathcal{H})$, we define the measure $V(\rho)$ by

$$(V\rho)(\Omega) = \rho(V \Omega V^{-1}) \,. \tag{5.8}$$

We choose a finite number of unitary transformations $V_1, \ldots, V_L$ and introduce a new measure $\tilde{\rho}$ as the convex combination of the unitarily transformed measures,

$$\tilde{\rho} = \frac{1}{L} \sum_{\mathfrak{a}=1}^{L} V_{\mathfrak{a}} \rho \,.$$

Obviously, all linear constraints like the volume constraint (1.3) and the trace constraint (1.4) are preserved by this transformation. The action becomes

$$\begin{aligned} \mathcal{S}(\tilde{\rho}) &= \frac{1}{L^2} \sum_{\mathfrak{a},\mathfrak{b}=1}^{L} \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d(V_{\mathfrak{a}}\rho)(x)\, d(V_{\mathfrak{b}}\rho)(y) \\ &= \frac{\mathcal{S}(\rho)}{L} + \frac{1}{L^2} \sum_{\mathfrak{a}\neq\mathfrak{b}} \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d(V_{\mathfrak{a}}\rho)(x)\, d(V_{\mathfrak{b}}\rho)(y) \,. \end{aligned} \tag{5.9}$$

Due to the factor $1/L$, the first summand becomes small as L increases. The second summand involves all the contributions for $\mathfrak{a} \neq \mathfrak{b}$. If we can arrange that these contributions become small, then the action of the new measure $\tilde{\rho}$ will indeed be smaller than the action of ρ.

Let us consider the contributions for $\mathfrak{a} \neq \mathfrak{b}$ in more detail. In order to simplify the explanations, it is convenient to assume that the measures $V_\mathfrak{a}\rho$ have mutually disjoint supports (this can typically be arranged by a suitable choice of the unitary transformations $V_\mathfrak{a}$). Then the space-time $\tilde{M} := \operatorname{supp} \tilde{\rho}$ can be decomposed into L "sub-space-times" $M_\mathfrak{a} := \operatorname{supp} \rho_\mathfrak{a}$,

$$\tilde{M} = M_1 \cup \cdots \cup M_L \qquad \text{and} \qquad M_\mathfrak{a} \cap M_\mathfrak{b} = \varnothing \quad \text{if } \mathfrak{a} \neq \mathfrak{b}\,.$$

Likewise, a physical wave function ψ^u can be decomposed into the contributions in the individual sub-space-times,

$$\psi^u = \sum_{\mathfrak{a}=1}^{L} \psi^u_\mathfrak{a} \qquad \text{with} \qquad \psi^u_\mathfrak{a} := \chi_{M_\mathfrak{a}}\, \psi^u$$

(and $\chi_{M_\mathfrak{a}}$ is the characteristic function). This also gives rise to a corresponding decomposition of the fermionic projector:

Lemma 5.2. *Every sub-space-time $M_\mathfrak{a}$ of $\tilde{M}$ is homeomorphic to M, with a homeomorphism given by*

$$\phi_\mathfrak{a} \,:\, M \to M_\mathfrak{a}\,, \qquad \phi_\mathfrak{a}(x) := V_\mathfrak{a}^*\, x\, V_\mathfrak{a}\,.$$

Moreover, the mapping

$$V_\mathfrak{a}^*\big|_{S_x} \,:\, S_x \to S_{\phi_\mathfrak{a}(x)} \tag{5.10}$$

is an isomorphism of the corresponding spinor spaces. Identifying the spinor spaces in different sub-space-times via this isomorphism, the fermionic projector can be written as

$$P(x,y) = -\sum_{\mathfrak{a},\mathfrak{b}=1}^{L} \chi_{M_\mathfrak{a}}(x)\, P_{\mathfrak{a},\mathfrak{b}}(x,y)\, \chi_{M_\mathfrak{b}}(y) \qquad \textit{with} \tag{5.11}$$

$$P_{\mathfrak{a},\mathfrak{b}}(x,y) := \Psi(x)\, V_\mathfrak{a}\, V_\mathfrak{b}^*\, \Psi(y)^*\,. \tag{5.12}$$

Proof. The definition of $V\rho$, (5.8), immediately implies that the transformation (1.16) maps M to $M_\mathfrak{a}$ and is a homeomorphism. By definition of the physical wave function (1.16),

$$\psi^u(\phi_\mathfrak{a}(x)) = \pi_{\phi_\mathfrak{a}(x)} = \pi_{V_\mathfrak{a}^* x V_\mathfrak{a}} u = V_\mathfrak{a}^*\, \pi_x\, V_\mathfrak{a} u\,.$$

The identification (5.10) makes it possible to leave out the factor $V_\mathfrak{a}^*$. Then we can write the wave evaluation operator (1.18) as

$$\tilde{\Psi}(x) = \sum_{\mathfrak{a}=1}^{L} \chi_{M_\mathfrak{a}}(x)\, \Psi(x)\, V_\mathfrak{a}\,.$$

Applying (1.20) gives the result. □

This lemma makes it possible to rewrite the action (5.9) as

$$\mathcal{S}(\tilde{\rho}) = \frac{\mathcal{S}(\rho)}{L} + \frac{1}{L^2} \sum_{\mathfrak{a} \neq \mathfrak{b}} \iint_{M \times M} \mathcal{L}\big[P_{\mathfrak{a},\mathfrak{b}}(x,y)\big]\, d\rho(x)\, d\rho(y)\,, \tag{5.13}$$

where the square bracket means that the Lagrangian is computed as a function of the kernel of the fermionic projector $P_{\mathfrak{a},\mathfrak{b}}(x,y)$ (just as explained after (1.12) for the kernel $P(x,y)$). The identities (5.12) and (5.13) give a good intuitive understanding of how the action depends on the unitary operators $V_\mathfrak{a}$. We first note that in the case $\mathfrak{a} = \mathfrak{b}$, the unitary operators in (5.12) drop out, so that $P_{\mathfrak{a},\mathfrak{a}}(x,y) = P(x,y)$. This

also explains why the first summand in (5.13) involves the original action $\mathcal{S}(\rho)$. In the $\mathfrak{a} \neq \mathfrak{b}$, however, the unitary operators in (5.12) do not drop out. In particular, this makes it possible to introduce phase factors into the fermionic projector. For example, one may change the phase of each physical wave function $\psi^u_\mathfrak{a}$ arbitrarily while keeping the physical wave functions $\psi^u_\mathfrak{b}$ for $\mathfrak{b} \neq \mathfrak{a}$ unchanged. Choosing the resulting phases randomly, one gets destructive interference, implying that the kernel $P_{\mathfrak{a},\mathfrak{b}}(x,y)$ becomes small. Making use of this *dephasing effect*, one can make the summands in (5.13) for $\mathfrak{a} \neq \mathfrak{b}$ small. A detailed analysis of the involved scalings reveals that this indeed makes it possible to decrease the causal action (see [13]).

In words, this result means that minimizing the causal action triggers a mechanism which tends to decompose space-time M into many small sub-space-times $M_1, \ldots, M_L$. The physical wave functions in the different sub-space-times involve relative phases, with the effect that the correlations between the sub-space-times (as described by the kernels $P_{\mathfrak{a},\mathfrak{b}}(x,y)$) become small. Since the dephasing takes place on a microscopic length scale, this effect is referred to as *microscopic mixing*.

Let us discuss what microscopic mixing implies for the effective macroscopic interaction. One must distinguish two situations. One limiting case is complete dephasing, in which case $P_{\mathfrak{a},\mathfrak{b}}$ is approximately zero. As a result, there are no relations or structures between the two sub-space-times (note that for example the causal structure is encoded in the kernel of the fermionic projector; see §1.3). This entails that the two sub-space-times do not interact with each other. The resulting picture is that space-time looks effectively like a "superposition" of the different sub-space-times. This scenario is referred to as the *microscopic mixing of space-time regions*. The dephasing can be understood similar to decoherence effects in standard quantum field theory (see for example [31]).

If each of the microscopically mixed sub-space-times involves a different classical bosonic field, one obtains effectively a superposition of classical field configurations. This makes it possible to describe second-quantized bosonic fields (see [11]). However, as the different sub-space-times do not interact with each other, each sub-space-time has it own independent dynamics. This dynamics is described by the classical bosonic field in the corresponding sub-space-time.

In order to obtain an interaction via second-quantized bosonic fields, one needs to consider another limiting case in which the dephasing involves only some of the physical wave functions. In this case, the fermionic projector $P_{\mathfrak{a},\mathfrak{b}}$ is not necessarily small. This also implies that relations arising as a consequence of the collective behavior of all physical wave functions (like the causal relations or classical bosonic fields) still exist between the sub-space-times $M_\mathfrak{a}$ and $M_\mathfrak{b}$. In more physical terms, the sub-space-times still interact with each other. This scenario is studied in [13] and is referred to as the *microscopic mixing of wave functions*. In order to describe the effective interaction, one describes the unitary operators $V_\mathfrak{a}$ by random matrices. Taking averages over the random matrices, one finds that the effective interaction can be described perturbatively in terms of Feynman diagrams which involve both fermionic and bosonic loops. The appearance of bosonic loops can be understood by working with second-quantized bosonic fields. Working out the detailed combinatorics and the implications of the

resulting quantum field theory is work in progress (for the first step in this program see [27]).

Acknowledgments: I would like to thank the referee for helpful comments on the manuscript.

References

[1] Y. Bernard and F. Finster, *On the structure of minimizers of causal variational principles in the non-compact and equivariant settings*, arXiv:1205.0403 [math-ph], Adv. Calc. Var. **7** (2014), no. 1, 27–57.

[2] J. Bognár, *Indefinite Inner Product Spaces*, Springer-Verlag, New York, 1974, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78.

[3] L. Bombelli, J. Lee, D. Meyer, and R.D. Sorkin, *Space-time as a causal set*, Phys. Rev. Lett. **59** (1987), no. 5, 521–524.

[4] R. Brunetti, M. Dütsch, and K. Fredenhagen, *Perturbative algebraic quantum field theory and the renormalization groups*, arXiv:0901.2038 [math-ph], Adv. Theor. Math. Phys. **13** (2009), no. 5, 1541–1599.

[5] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York-London, 1969, Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.

[6] F. Finster, *The Continuum Limit of Causal Fermion Systems*, book based on the preprints arXiv:0908.1542 [math-ph], arXiv:1211.3351 [math-ph] and arXiv:1409.2568 [math-ph], in preparation.

[7] ———, *Definition of the Dirac sea in the presence of external fields*, arXiv:hep-th/9705006, Adv. Theor. Math. Phys. **2** (1998), no. 5, 963–985.

[8] ———, *The Principle of the Fermionic Projector*, hep-th/0001048, hep-th/0202059, hep-th/0210121, AMS/IP Studies in Advanced Mathematics, vol. 35, American Mathematical Society, Providence, RI, 2006.

[9] ———, *A variational principle in discrete space-time: Existence of minimizers*, arXiv:math-ph/0503069, Calc. Var. Partial Differential Equations **29** (2007), no. 4, 431–453.

[10] ———, *Causal variational principles on measure spaces*, arXiv:0811.2666 [math-ph], J. Reine Angew. Math. **646** (2010), 141–194.

[11] ———, *Entanglement and second quantization in the framework of the fermionic projector*, arXiv:0911.0076 [math-ph], J. Phys. A: Math. Theor. **43** (2010), 395302.

[12] ———, *The chiral index of the fermionic signature operator*, arXiv:1404.6625 [math-ph] (2014).

[13] ———, *Perturbative quantum field theory in the framework of the fermionic projector*, arXiv:1310.4121 [math-ph], J. Math. Phys. **55** (2014), no. 4, 042301.

[14] F. Finster and A. Grotz, *The causal perturbation expansion revisited: Rescaling the interacting Dirac sea*, arXiv:0901.0334 [math-ph], J. Math. Phys. **51** (2010), 072301.

[15] ———, *A Lorentzian quantum geometry*, arXiv:1107.2026 [math-ph], Adv. Theor. Math. Phys. **16** (2012), no. 4, 1197–1290.

[16] ———, *On the initial value problem for causal variational principles*, arXiv:1303.2964 [math-ph], to appear in J. Reine Angew. Math. (2016).

[17] F. Finster, A. Grotz, and D. Schiefeneder, *Causal fermion systems: A quantum space-time emerging from an action principle*, arXiv:1102.2585 [math-ph], Quantum Field Theory and Gravity (F. Finster, O. Müller, M. Nardmann, J. Tolksdorf, and E. Zeidler, eds.), Birkhäuser Verlag, Basel, 2012, pp. 157–182.

[18] F. Finster and N. Kamran, *Spinors on singular spaces and the topology of causal fermion systems*, arXiv:1403.7885 [math-ph] (2014).

[19] F. Finster and J. Kleiner, *The jet bundle dynamics of causal fermion systems*, in preparation.

[20] ———, *Causal fermion systems as a candidate for a unified physical theory*, arXiv:1502.03587 [math-ph] (2015).

[21] ———, *Noether-like theorems for causal variational principles*, arXiv:1506.09076 [math-ph] (2015).

[22] F. Finster and O. Müller, *Lorentzian spectral geometry for globally hyperbolic surfaces*, arXiv:1411.3578 [math-ph] (2014).

[23] F. Finster, S. Murro, and C. Röken, *The fermionic projector in a time-dependent external potential: Mass oscillation property and Hadamard states*, arXiv:1501.05522 [math-ph] (2015).
[24] F. Finster and M. Reintjes, *A non-perturbative construction of the fermionic projector on globally hyperbolic manifolds II – Space-times of infinite lifetime*, arXiv:1312.7209 [math-ph] (2013).
[25] ———, *A non-perturbative construction of the fermionic projector on globally hyperbolic manifolds I – Space-times of finite lifetime*, arXiv:1301.5420 [math-ph], to appear in Adv. Theor. Math. Phys. (2015).
[26] F. Finster and D. Schiefeneder, *On the support of minimizers of causal variational principles*, arXiv:1012.1589 [math-ph], Arch. Ration. Mech. Anal. **210** (2013), no. 2, 321–364.
[27] F. Finster and J. Tolksdorf, *A microscopic derivation of quantum electrodynamics*, in preparation.
[28] ———, *Perturbative description of the fermionic projector: Normalization, causality and Furry's theorem*, arXiv:1401.4353 [math-ph], J. Math. Phys. **55** (2014), no. 5, 052301.
[29] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Fourth edition prepared by Ju. V. Geronimus and M. Ju. Ceutlin., Academic Press, New York, 1965.
[30] P.R. Halmos, *Measure Theory*, Springer, New York, 1974.
[31] E. Joos, H.D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory*, second ed., Springer-Verlag, Berlin, 2003.
[32] H. Langer, *Spectral functions of definitizable operators in Krein spaces*, Functional Analysis (Dubrovnik, 1981), Lecture Notes in Math., vol. 948, Springer, Berlin, 1982, pp. 1–46.
[33] P.D. Lax, *Functional Analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2002.
[34] F.W.J. Olver, D.W. Lozier, R.F. Boisvert, and C.W. Clark (eds.), *Digital Library of Mathematical Functions*, National Institute of Standards and Technology from http://dlmf.nist.gov/ (release date 2011-07-01), Washington, DC, 2010.
[35] W. Rudin, *Principles of Mathematical Analysis*, third ed., McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976, International Series in Pure and Applied Mathematics.

Fakultät für Mathematik, Universität Regensburg, D-93040 Regensburg, Germany
E-mail address: finster@ur.de