Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

A Variational Principle in Discrete Space-Time - Existence of Minimizers

by

Felix Finster

Preprint no.: 36 2005

A Variational Principle in Discrete Space-Time – Existence of Minimizers

Felix Finster

March 2005

Abstract

We formulate a variational principle for a collection of projectors in an indefinite inner product space. The existence of minimizers is proved under general assumptions.

In a recent book it was proposed to formulate physics with a new variational principle in space-time [2]. In the present paper we construct minimizers of this variational principle. In order to make the presentation self-contained and easily accessible, we introduce the mathematical framework from the basics (see Sections 1 and 2). Thus this paper can be used as an introduction to the mathematical setting of the principle of the fermionic projector. However, the reader who wants to get a physical understanding is referred to [2].

Our variational principle is set up in finite dimension, and thus the continuity of the action is not an issue. The difficulties are the lack of compactness and the fact that there is no notion of convexity. Therefore, we need to derive suitable estimates (Sections 4 and 6) before we can use the direct method of the calculus of variations (Sections 7 and 8). Our main results can be understood already after reading Sections 1 and 2 and are stated in Section 5 (see Theorem 5.2, Theorem 5.3 and Corollary 5.5).

1 Discrete Space-Time and the Fermionic Projector

Let H be a finite-dimensional complex vector space, endowed with a sesquilinear form $<.|.> \,:\, H \times H \to \mathbb{C}$, i.e. for all $u, v, w \in H$ and $\alpha, \beta \in \mathbb{C}$,

$$\begin{aligned} <u \,|\, \alpha v + \beta w> &= \alpha <u \,|\, v> + \beta <u \,|\, w> \\ <\alpha u + \beta v \,|\, w> &= \overline{\alpha} <u \,|\, w> + \overline{\beta} <v \,|\, w> . \end{aligned}$$

We assume that $<.|.>$ is symmetric,

$$\overline{<u \,|\, v>} \;=\; <v \,|\, u> ,$$

and non-degenerate,

$$<u \,|\, v> \;=\; 0 \quad \forall\, v \in H \quad \Longrightarrow \quad u \;=\; 0 .$$

Note that $<.|.>$ is in general not positive, and it is therefore not a scalar product. We also refer to $(H, <.|.>)$ as an *indefinite inner product space*. To a non-degenerate subspace of H we can associate its *signature* (p, q), where p and q are the maximal dimensions of positive and negative definite subspaces, respectively (for more details see [1, 3] and the examples in Section 3).

Many constructions familiar from scalar product spaces can be carried over to indefinite inner product spaces. In particular, we define the *adjoint* of a linear operator $A : H \to H$ by the relation

$$<u \,|\, Av> \;=\; <A^* u \,|\, v> \qquad \forall\, u, v \in H\,.$$

A linear operator A is said to be *unitary* if $A^* = A^{-1}$ and *symmetric* if $A^* = A$. It is called a *projector* if it is symmetric and idempotent,

$$A^* \;=\; A \;=\; A^2\,.$$

Let M be a finite set. To every point $x \in M$ we associate a projector E_x. We assume that these projectors are orthogonal and complete in the sense that

$$E_x\, E_y \;=\; \delta_{xy}\, E_x \qquad \text{and} \qquad \sum_{x \in M} E_x \;=\; \mathbb{1}\,. \tag{1}$$

Equivalently, we can say that the images of the projectors E_x give a decomposition of H into orthogonal subspaces,

$$H \;=\; \bigoplus_{x \in M} E_x(H)\,. \tag{2}$$

Furthermore, we assume that the images $E_x(H) \subset H$ of these projectors are non-degenerate and all have the same signature (n, n). We refer to (n, n) as the *spin dimension.* Relation (2) shows that the dimension of H must be equal to $m \cdot 2n$, where $m = \#M$ denotes the number of points of M. The points $x \in M$ are called *discrete space-time points*, and the corresponding projectors E_x are the *space-time projectors.* The structure $(H, <.|.>, (E_x)_{x \in M})$ is called *discrete space-time.*

We now introduce one more projector P on H, the so-called *fermionic projector*, which has the additional property that its image $P(H)$ is *negative* definite. In other words, $P(H)$ has signature $(0, f)$ with $f \in \mathbb{N}$. The vectors in the image of P have the interpretation as the quantum mechanical states of the particles of our system, and we call $f = \dim P(H)$ the *number of particles.* We remark that in physical applications [2] these particles are Dirac particles, which are fermions, giving rise to the name "fermionic projector".

A space-time projector E_x can be used to restrict an operator to the subspace $E_x(H) \subset H$. Using a more graphic notion, we also refer to this restriction as the *localization* at the space-time point x. For example, using the completeness of the space-time projectors (1), we readily see that

$$f \;=\; \mathrm{Tr}\, P \;=\; \sum_{x \in M} \mathrm{Tr}(E_x P)\,. \tag{3}$$

The expression $\mathrm{Tr}(E_x P)$ can be understood as the localization of the trace at the space-time point x, and summing over all space-time points gives the total trace. We call $\mathrm{Tr}(E_x P)$ the *local trace* of P. When forming more complicated composite expressions in the projectors P and $(E_x)_{x \in M}$, it is convenient to use the short notations

$$P(x, y) \;=\; E_x\, P\, E_y \qquad \text{and} \qquad u(x) \;=\; E_x\, u\,.$$

Referring to the orthogonal decomposition (2), $P(x, y)$ maps $E_y(H)$ to $E_x(H)$ and vanishes otherwise. It is often useful to regard $P(x, y)$ as a mapping only between these subspaces,

$$P(x, y) \;:\; E_y(H) \;\mapsto\; E_x(H)\,.$$

Using (1), we can write the product Pu as follows,

$$(Pu)(x) \;=\; E_x\,Pu \;=\; \sum_{y\in M} E_x\,P\,E_y\,u \;=\; \sum_{y\in M} (E_x\,P\,E_y)\,(E_y\,u)\,,$$

and thus

$$(Pu)(x) \;=\; \sum_{y\in M} P(x,y)\,u(y)\,.$$

This relation resembles the representation of an operator with an integral kernel. Therefore, we call $P(x,y)$ the *discrete kernel* of the fermionic projector. The discrete kernel can be used for expressing general operator products; for example,

$$P\,E_x\,P\,E_y \;=\; \sum_{z\in M} P(z,x)\,P(x,y)\,.$$

2 A Variational Principle

We want to form a positive quantity which depends on the form of the fermionic projector relative to the space-time projectors. Since scalar invariants (like the trace or the determinant) can be introduced only for operators which map a vector space into itself, we first define the *closed chain* A_{xy} by

$$A_{xy} \;=\; P(x,y)\,P(y,x) \;=\; E_x\,P\,E_y\,P\,E_x \;:\; E_x(H) \;\mapsto\; E_x(H)\,. \tag{4}$$

We often omit the subscripts 'xy'. Let $\lambda_1,\ldots,\lambda_{2n}$ be the zeros of the characteristic polynomial of A, counted with multiplicities. We define the *spectral weight* $|A|$ by

$$|A| \;=\; \sum_{j=1}^{2n} |\lambda_j|\,.$$

We introduce the *Lagrangian* $\mathcal{L}[A]$ by

$$\mathcal{L}[A] \;=\; |A^2| - \frac{1}{2n}\,|A|^2 \tag{5}$$

and form the *action* $\mathcal{S}$ by summing over the discrete space-time points,

$$\mathcal{S}[P] \;=\; \sum_{x,y\in M} \mathcal{L}[A_{xy}]\,. \tag{6}$$

Our variational principle is to

$$\text{minimize } \mathcal{S}[P] \text{ by varying } P\,, \tag{7}$$

keeping the number of particles f as well as discrete space-time $(H, <.|.>, (E_x)_{x\in M})$ fixed.

Before moving on to simple examples in the next section, we now give a few general explanations. First of all, let us verify that the action is non-negative. According to the Schwarz inequality,

$$|A| \;=\; \sum_{j=1}^{2n} |\lambda_j| \;\leq\; \left(\sum_{j=1}^{2n} 1\right)^{\frac{1}{2}} \left(\sum_{j=1}^{2n} |\lambda_j|^2\right)^{\frac{1}{2}} \;=\; \sqrt{2n}\,|A^2|^{\frac{1}{2}}\,,$$

and squaring both sides we find that $\mathcal{L}[A] \geq 0$. Our Lagrangian can also be written in the form

$$\mathcal{L}[A] \;=\; \frac{1}{4n} \sum_{i,j=1}^{n} \left(|\lambda_i| - |\lambda_j|\right)^2 \;, \tag{8}$$

as is easily verified by multiplying out the square in (8). This shows that the Lagrangian vanishes only if the $|\lambda_j|$ are all equal. Thus one can say qualitatively that our variational principle tries to achieve that the zeros of the characteristic polynomial of A all have the same absolute value.

The Lagrangian $\mathcal{L}[A_{xy}]$ is symmetric in its two arguments x and y, as the following consideration shows. For any two quadratic matrices B and C, we choose ε not in the spectrum of C and set $C^\varepsilon = C - \varepsilon \mathbb{1}$. Taking the determinant of the relation $C^\varepsilon(BC^\varepsilon - \lambda) = (C^\varepsilon B - \lambda)C^\varepsilon$, we can use that the determinant is multiplicative and that $\det C^\varepsilon \neq 0$ to obtain the equation $\det(BC^\varepsilon - \lambda) = \det(C^\varepsilon B - \lambda)$. Since both determinants are continuous in ε, this equation holds even for all $\varepsilon \in \mathbb{R}$, proving the elementary identity

$$\det(BC - \lambda \mathbb{1}) \;=\; \det(CB - \lambda \mathbb{1})\,.$$

Applying this identity to the closed chain,

$$\begin{aligned} \det(A_{xy} - \lambda \mathbb{1}) \;&=\; \det(P(x,y)\,P(y,x) - \lambda \mathbb{1}) \\ &=\; \det(P(y,x)\,P(x,y) - \lambda \mathbb{1}) \;=\; \det(A_{yx} - \lambda \mathbb{1})\,, \end{aligned}$$

we conclude that the operators A_{xy} and A_{yx} have the same characteristic polynomial, and thus

$$\mathcal{L}[A_{xy}] \;=\; \mathcal{L}[A_{yx}] \qquad \forall\, x, y \in M\,. \tag{9}$$

It is a simple but important observation that a joint unitary transformation of all projectors,

$$E_x \;\to\; U E_x U^{-1}\,, \qquad P \;\to\; U P U^{-1} \qquad \text{with } U \text{ unitary} \tag{10}$$

keeps the action unchanged, because

$$\begin{aligned} P(x,y) \;&\to\; U\,P(x,y)\,U^{-1}\,, \qquad A_{xy} \;\to\; U A_{xy} U^{-1} \\ \det(A_{xy} - \lambda \mathbb{1}) \;&\to\; \det\big(U (A_{xy} - \lambda \mathbb{1})\, U^{-1}\big) \;=\; \det(A_{xy} - \lambda \mathbb{1})\,, \end{aligned}$$

and so the λ_j stay the same. Such unitary transformations can also be used to vary the fermionic projector. However, since we want to keep discrete space-time fixed, we are only allowed to consider unitary transformations which do not change the space-time projectors,

$$E_x \;=\; U E_x U^{-1} \qquad \forall\, x \in M\,. \tag{11}$$

Then (10) reduces to the transformation of the fermionic projector

$$P \;\to\; U P U^{-1}\,. \tag{12}$$

Unitary transformations of the form (11, 12) are called *gauge transformations*. The conditions (11) mean that U maps the subspaces $E_x(H)$ into itself. Hence U splits into a direct sum of unitary transformations

$$E_x U \;:\; E_x(H) \;\mapsto\; E_x(H)\,, \tag{13}$$

which act "locally" on the subspaces associated to the individual space-time points. Obviously, the gauge transformations form a group, referred to as the *gauge group* $\mathcal{G}$. Localizing the gauge transformations according to (13), we obtain at any space-time point x the so-called *local gauge group*. The local gauge group is the group of isometries of $E_x(H)$ and can thus be identified with the group $U(n,n)$.

One may ask why the space-time projectors are to be kept fixed in our variational principle (7). More generally, one could vary both P and the $(E_x)_{x \in M}$, fixing only the integer parameters f and n. Recall that the space-time projectors are equivalently described by the orthogonal decomposition (2) together with the condition that the subspaces $E_x(H)$ should all have signature (n,n). For two different sets of space-time projectors, we can find a unitary transformation which maps the corresponding subspaces $E_x(H)$ onto each other. Then the transition from one set of space-time projectors to the other is described by the unitary transformation $E_x \to U E_x U^{-1}$. Since such unitary transformations leave the action unchanged if also the fermionic projector is transformed according to (10), it is no loss in generality to fix the space-time projectors throughout.

It is instructive to consider our framework in a concrete basis of H. Then our inner product can be represented in the form

$$<u \,|\, v> \;=\; (u \,|\, Sv) \,,$$

where $(.|.)$ is the canonical scalar product on $\mathbb{C}^{2mn}$. Here S is a Hermitian matrix (meaning that $(u \,|\, Sv) = (Su \,|\, v) \;\forall\, u,v \in H$), referred to as the *signature matrix*. By choosing the basis of H appropriately, we can arrange that S is diagonal with eigenvalues equal to ± 1. In particular, S is unitary and $S^2 = \mathbb{1}$. The signature matrix is useful for calculations. For example,

$$<u \,|\, Av> \;=\; (u \,|\, SAv) \;=\; (A^\dagger S u \,|\, v) \;=\; (SA^\dagger S u \,|\, Sv) \;=\; <SA^\dagger S u \,|\, v> \,,$$

where the dagger denotes transposition and complex conjugation. Thus the adjoint can be expressed by

$$A^* \;=\; SA^\dagger S \,.$$

A matrix is symmetric if and only if the matrix SA is Hermitian. As one already sees in the two-dimensional example

$$S \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad A \;=\; \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \tag{14}$$

a symmetric matrix in an indefinite inner product space need not be diagonalizable. This explains why after (4) we had to speak of "zeros of the characteristic polynomial" and not of "eigenvalues." Note that the matrix A in (14) is nilpotent and thus $|A| = 0$. This shows that the spectral weight is not a matrix norm, not even on symmetric operators. We remark that it seems impossible to introduce any other basis independent matrix norm; in particular, the analogue of the Hilbert-Schmidt norm $(\mathrm{Tr}(A^* A))^{\frac{1}{2}}$ vanishes in the example (14). Even if a symmetric matrix is diagonalizable, its eigenvalues are in general not real, as can be seen in the example

$$S \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad A \;=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{15}$$

At least, the calculation

$$\overline{\det(A - \lambda \mathbb{1})} \;=\; \det\Big(A^\dagger - \overline{\lambda}\Big) \;=\; \det\Big(S(A^\dagger - \overline{\lambda})S\Big) \;=\; \det\big(A^* - \overline{\lambda}\big) \;=\; \det\big(A - \overline{\lambda}\mathbb{1}\big)$$

shows that the characteristic polynomial of a symmetric matrix A has real coefficients. In other words, the non-real λ_j always appear in complex conjugate pairs.

3 Simple Examples

In order to get a first impression of our variational principle, we now consider a few simple examples. We begin with the case $m = 1$ of one space-time point. In this case, the only space-time projector E is the identity, and the sum over the space-time points in (6) drops out. Thus

$$\mathcal{S} \;=\; |A^2| - \frac{1}{2n}\,|A|^2$$

with $A = P^2$. Using that P is idempotent and that its only non-vanishing eigenvalue is one with multiplicity f, we find

$$\mathcal{S} \;=\; |P| - \frac{1}{2n}\,|P|^2 \;=\; f - \frac{f^2}{2n}\,.$$

Hence the action is unchanged if the fermionic projector is varied. This can also be understood from the fact that with only one space-time point, the condition (11) is trivial, and therefore any variation of P can be realized as a gauge transformation (12). The situation becomes more interesting with two space-time points, as the next example shows.

Example 3.1 Choose $M = \{1, 2\}$ with spin dimension $(1, 1)$ and $f = 1$. Then H is 4-dimensional, and by choosing a suitable basis we can arrange that

$$S \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad E_1 \;=\; \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 \;=\; \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \tag{16}$$

where for $E_{1\!/\!2}$ we used a block matrix notation (thus every matrix entry stands for a 2×2-matrix). Again in this block matrix notation, the gauge transformations (11) are of the form

$$U \;=\; \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}, \tag{17}$$

where $(U_x)_{x \in M}$ are two independent "local" unitary transformations on $E_x(H)$ of the form

$$U_x \;=\; \begin{pmatrix} \cosh\vartheta & \sinh\vartheta \\ \sinh\vartheta & \cosh\vartheta \end{pmatrix} \qquad \text{with } \vartheta \in \mathbb{R}.$$

Thus the local gauge group is $U(1, 1)$, and the gauge transformations (17) are elements of the gauge group $\mathcal{G} = U(1, 1) \otimes U(1, 1)$.

Since we consider a system of one particle ($f = 1$), the fermionic projector P must be a projector on a one-dimensional, negative definite subspace. It is convenient to write P using bra/ket-notation as

$$P \;=\; -|\,u{>}{<}u\,| \qquad \text{with} \qquad {<}u\,|\,u{>} \;=\; -1\,. \tag{18}$$

A possible choice is

$$u \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and thus} \qquad P \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{19}$$

A short calculation yields that $|A_{11}| = |A_{11}^2| = 1$, and all other A_{ij} vanish. Thus

$$\mathcal{S} \;=\; \mathcal{L}(A_{11}) \;=\; |A_{11}^2| - \frac{1}{2}\,|A_{11}|^2 \;=\; \frac{1}{2}\,.$$

It turns out that the above P is not a minimizer. Namely, choosing

$$u \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \qquad P \;=\; \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \tag{20}$$

we get a smaller value for the action,

$$\begin{aligned} |A_{ij}|^2 &\;=\; |A_{ij}^2| \;=\; \frac{1}{16} \quad \text{for all } i,j \in M \\ \mathcal{S} &\;=\; 4\,\mathcal{L}(A_{11}) \;=\; 4\left(\frac{1}{16} - \frac{1}{2\cdot 16}\right) \;=\; \frac{1}{8}\,. \end{aligned}$$

Let us verify that this indeed the minimum. Representing a general P in the form (18), we can use the gauge freedom (17) to arrange that u is of the form $u = (0, \cos\varphi, 0, \sin\varphi)$ with $\varphi \in [0, 2\pi)$. A short calculation yields that

$$\begin{aligned} |A_{11}|^2 &\;=\; |A_{ij}^2| \;=\; \cos^8\varphi\,, \quad |A_{22}|^2 \;=\; |A_{22}^2| \;=\; \sin^8\varphi \\ |A_{12}|^2 &\;=\; |A_{12}^2| \;=\; |A_{21}|^2 \;=\; |A_{21}^2| \;=\; \sin^4\varphi\,\cos^4\varphi \\ \mathcal{S} &\;=\; \sum_{i,j\in M} |A_{ij}^2| - \frac{1}{2}\,|A_{ij}|^2 \;=\; \frac{1}{2}\left(\cos^4\varphi + \sin^4\varphi\right)^2 \;=\; \frac{1}{2}\left(2\sin^4\varphi - 2\sin^2\varphi + 1\right)^2\,, \end{aligned}$$

and the last function really attains its minimum when $\sin^2\varphi = 1/2$. ♦

In the above example, our variational principle has, up to gauge transformations, a unique minimum. The fact that the configuration (19) where the particle is localized at the first space-time point is not optimal can be understood qualitatitively by saying that our variational principle "tends to spread out particles in space-time." We will quantify this observation later (see Lemma 6.1); it will be important in our analysis. We also point out that the local gauge group is *non-compact*, and that the set of gauge-equivalent minima UPU^{-1} with P and U according to (20, 17) form an *unbounded* family of matrices. This explains why minimizers cannot be constructed with simple compactness arguments.

We next consider a system of two particles.

Example 3.2 Choose $M = \{1, 2\}$ with spin dimension $(1, 1)$ and $f = 2$. Thus the discrete space-time is the same as in Example 3.1; it is again described by (16). However, P now maps onto a two-dimensional negative subspace of H. Since the subspaces $E_x(H)$ have at most one-dimensional negative subspaces, one may expect that the spaces $E_x(H) \cap P(H)$, $x \in M$, should both be one-dimensional. This is indeed the case, as the following consideration shows. Since $P(H) \subset H$ is a negative subspace of maximal dimension (recall that H has signature $(2, 2)$), its orthogonal complement $(\mathbb{1} - P)(H)$ is a positive subspace. This allows us to introduce on H a scalar product $(.|.)_P$ by

$$(u \,|\, v)_P \;=\; <u \,|\, (\mathbb{1} - P)\, v> \;-\; <u \,|\, P\, v>. \tag{21}$$

For clarity, we denote the corresponding Hilbert space by $(\mathcal{H}, (.|.)_P)$. Since P is symmetric in H and commutes with itself, it is obvious from (21) that the operator P is self-adjoint in $\mathcal{H}$. Furthermore, we let $\mathcal{E}_x$ be the projectors on the subspaces $\mathcal{H}_x := E_x(H) \subset \mathcal{H}$ (thus the $\mathcal{E}_x$ are self-adjoint with respect to $(.|.)_P$; note that they are in general different from the E_x, which are symmetric with respect to $<.|.>$). Then the operator $P_1 := \mathcal{E}_1 P \mathcal{E}_1$ is self-adjoint on the Hilbert space $(\mathcal{H}_1, (.|.)_P)$, and we can diagonalize it. For an eigenvalue λ and corresponding eigenvector $u \in \mathcal{H}_1$, we write the eigenvalue equation as follows,

$$((P - \lambda \mathbb{1})\, u \,|\, v)_P \;=\; 0 \qquad \forall\, v \in \mathcal{H}_1\,. \tag{22}$$

In the case $\lambda = 1$, we evaluate this equation for $v = u$ and use that $(\mathbb{1} - P)^2 = \mathbb{1} - P$,

$$0 \;=\; ((P - \mathbb{1})^2\, u \,|\, u)_P \;=\; ((P - \mathbb{1})\, u \,|\, (P - \lambda)\, u)_P$$

and thus $Pu = u \in \mathcal{H}_1$. Hence we have found one vector in $E_1(H) \cap P(H)$. If the other eigenvalue of P_1 were also equal to one, we could repeat the above argument to conclude that $P(H)$ spans $E_1(H)$, which is a contradiction because $E_1(H)$ is not negative. If on the other hand we have an eigenvalue $\lambda \neq 1$, we can rewrite (22) with (21) and use that $(\mathbb{1} - P)\, v = 0$ to obtain

$$0 \;=\; ((P - \lambda \mathbb{1})\, u \,|\, v)_P \;=\; (\lambda - 1) <Pu \,|\, v> \qquad \forall\, v \in \mathcal{H}_1\,.$$

This means that $<Pu \,|\, v> = 0$ for all $v \in E_1(H)$, implying that $Pu \in E_2(H)$. Thus we have found a vector in $E_2(H) \cap P(H)$. Now the other eigenvalue of P_1 must be equal to one, because otherwise $P(H)$ would span $E_2(H)$, and we would again get a contradiction. We conclude that there are vectors u_1 and u_2 with

$$0 \neq u_x \;\in\; E_x(H) \cap P(H)\,,$$

and since $P(H)$ has dimension two, these two vectors clearly span $P(H)$.

Using the gauge freedom (17), we can arrange that u_1 is a multiple of the vector $(0, 1, 0, 0)$, and u_2 is proportional to $(0, 0, 0, 1)$. Then P must be of the form

$$P \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ * & * & * & * \\ 0 & 0 & 0 & 0 \\ * & * & * & * \end{pmatrix},$$

where the stars stand for any complex matrix entries. Using furthermore that P is symmetric, idempotent and has two-dimensional range, we conclude that P must be equal to the matrix

$$P \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{23}$$

This means that the two particles must be localized, one at each space-time point. The fermionic projector is unique up to gauge transformations. ♦

The argument of this example shows more generally that our variational principle becomes trivial when the number of particles is equal to the maximal dimension $m \cdot n$ of negative subspaces of H. This is intuitively clear, because in this case there is (up to gauge transformations) only one possible configuration: at each space-time point we must localize n particles.

The most interesting case is when the number of particles is large, but still much smaller than the number of space-time points.

Example 3.3 Choose $1 \ll f \ll m$ with spin dimension $(1,1)$. In this case, we can represent discrete space-time by the following matrices,

$$S = \begin{pmatrix} 1 & 0 & & & \\ 0 & -1 & & & \\ & & 1 & 0 & \\ & & 0 & -1 & \\ & & & & \ddots \end{pmatrix}$$

and

$$E_1 = \begin{pmatrix} \mathbb{1} & & \\ & 0 & \\ & & \ddots \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & & \\ & \mathbb{1} & \\ & & \ddots \end{pmatrix}, \ldots, E_m .$$

One possibility to choose P is to localize each of the f particles similar to (23) at one of the space-time points. However, the resulting value for the action

$$\mathcal{S} = \frac{f}{2}$$

is certainly not minimal. It is better if, in analogy to (20), each particle is evenly spread over m/f space-time points (we here assume for simplicity that m/f is an integer). A short calculation yields

$$\mathcal{S} = \frac{f}{2} \left(\frac{f}{m} \right)^2 .$$

There is no reason why this configuration should be optimal. It is completely unknown how the minimizer looks like in general. ♦

The case of physical interest is spin dimension $(2,2)$ (or more generally $(2N, 2N)$ with $N \geq 1$), because in this case the vectors of H can be identified with the Dirac wave functions of relativistic quantum mechanics. We expect the general structure of the minima to be very complicated. Our qualitative picture is that, depending on the value of the so-called *filling factor* $f/(nm)$, the minimizers should induce relations between the discrete space-time points which for large m should correspond to different geometric configurations of the space-time points. To give a specific example for such a "relation between the discrete space-time points", we finally remark that two points $x, y \in M$ may be called *timelike* or *spacelike* separated if the matrix A_{xy} has a purely real spectrum or has also complex spectral points, respectively. As explained in [2, §5.6], this relation should, in a suitable limit in which discrete space-time goes over to a continuum space-time, give the causal structure of a Lorentzian manifold.

4 Positive Operators, Lower Bounds for the Lagrangian

We now introduce a concept which will be an important ingredient to our existence proof.

Def. 4.1 *A symmetric operator A on an indefinite inner product space of signature (p, q) is said to be* **positive** *if*

$$<u \,|\, A u> \;\geq\; 0 \qquad \forall\, u \in H\,.$$

Expressed with the signature matrix, we can say that A is positive if and only if the matrix SA is positive semi-definite on $\mathbb{C}^{p+q}$ endowed with the standard Euclidean scalar product. To avoid confusion, we point out that the statements "A is positive" and "the image of A is positive" are completely different. In the two-dimensional examples

$$S \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad A_1 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 \;=\; \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \tag{24}$$

the operator A_1 is positive, although its image has signature $(1, 1)$. The operator A_2 is also positive, but its image is negative. The last example also shows that the trace of a positive operator can be negative. At least, the *projector* on a positive subspace is a positive operator, as the following argument shows.

Lemma 4.2 *The operator $(-P)$ (with P the fermionic projector) is a positive operator.*

Proof. Using that P is idempotent, symmetric, and maps to a negative subspace, we obtain for all $u \in H$,

$$<u \,|\, (-P)\, u> \;=\; -<u \,|\, P^2\, u> \;=\; -<Pu \,|\, Pu> \;\geq\; 0\,. \qquad \blacksquare$$

We now collect a few elementary but useful properties of positive operators.

Lemma 4.3 *Suppose that A is a positive operator on H. Then*

(i) *If Q is a projector in H, the operator QAQ is again positive.*

(ii) *For all $u, v \in H$,*

$$|<u \,|\, A v>| \;\leq\; \sqrt{<u \,|\, A u>}\; \sqrt{<v \,|\, A v>}\,. \tag{25}$$

Proof. Part **(i)** is obvious because $<u \,|\, QAQ\, u> \,=\, <Qu \,|\, A\, Qu> \,\geq 0$. Part **(ii)** can be regarded as the Schwarz inequality for the positive semi-definite inner product $(.|.)_A := <.|A.>$. The proof is almost as simple as in scalar product spaces: First note that for all $a, b \in H$,

$$0 \;\leq\; (a - b \,|\, a - b)_A \;=\; (a \,|\, a)_A \;+\; (b \,|\, b)_A \;-\; 2\,\mathrm{Re}\,(a \,|\, b)_A\,.$$

By changing the phase of the vector a, we can arrange that $(a \,|\, b)_A \geq 0$. Thus

$$2\,|(a \,|\, b)_A| \;\leq\; (a \,|\, a)_A \;+\; (b \,|\, b)_A \qquad \forall\, a, b \in H\,. \tag{26}$$

Suppose that $(u \,|\, u)_A = 0$. Then applying (26) for $a = u/\varepsilon$ and $b = \epsilon u$ with a parameter $\epsilon > 0$ gives

$$2\,|(u \,|\, v)_A| \;\leq\; \varepsilon^2\, (v \,|\, v)_A\,,$$

and letting $\varepsilon \to 0$, we see that (25) is trivially satisfied. The same argument applies if $(v \,|\, v)_A = 0$. In the remaining case $(u \,|\, u)_A \neq 0$ and $(v \,|\, v)_A \neq 0$, we apply (26) with

$$a \;=\; \left(\frac{(v \,|\, v)_A}{(u \,|\, u)_A} \right)^{\frac{1}{4}} u\,, \qquad b \;=\; \left(\frac{(u \,|\, u)_A}{(v \,|\, v)_A} \right)^{\frac{1}{4}} v\,. \qquad \blacksquare$$

Compared to the situation for general symmetric operators as explained after (14), positive operators have nice spectral properties, as the following approximation argument shows.

Lemma 4.4 *A positive operator A on an indefinite inner product space of signature (p, q) has a purely real spectrum. The zeros $(\lambda_j)_{j=1,\ldots,2n}$ if its characteristic polynomial (again counted with multiplicities) can be ordered as follows,*

$$\lambda_1 \leq \cdots \leq \lambda_q \;\leq\; 0 \;\leq\; \lambda_{q+1} \leq \cdots \leq \lambda_{p+q} \,.$$

Proof. We choose a matrix representation with signature matrix S and set $A^\varepsilon = A + \varepsilon S$. Clearly, the matrices A^ε converge to A as $\varepsilon \to 0$. Since the spectrum is continuous in ε, it suffices to prove the lemma for the matrix A^ε and any $\varepsilon > 0$.

The matrix A^ε is symmetric and strictly positive in the sense that for all $u \neq 0$,

$$<u \,|\, A^\varepsilon u> \;=\; <u \,|\, A\, u> \;+\; <u \,|\, S\, u> \;\geq\; <u \,|\, S\, u> \;=\; (u \,|\, u) \;>\; 0 \,.$$

Hence we can introduce a scalar product by

$$(u \,|\, v)_{A^\varepsilon} \;:=\; <u \,|\, A^\varepsilon v> .$$

Since the operator A^ε is symmetric and commutes with itself, it is clearly self-adjoint in the Hilbert space $(H, (.|.)_{A^\varepsilon})$. Thus we can choose an eigenvector basis $(u_j)_{j=1,\ldots,p+q}$. The corresponding eigenvalues λ_j satisfy the identity

$$\lambda_j <u_j \,|\, u_j> \;=\; <u_j \,|\, A^\varepsilon u_j> \;=\; (u_j \,|\, u_j)_{A^\varepsilon} \;>\; 0 \,.$$

Thus p of the eigenvalues are positive, whereas the other q eigenvalues are negative. ■

Note that a positive operator is in general not diagonalizable, as the example (14) shows.

The above lemmas can be used to get lower estimates of our Lagrangian and the action, which shed some light on the mathematical behavior of our variational principle. It is obvious from (3) that the local trace is non-zero at least at some $x \in M$. The next lemma shows that the Lagrangian of A_{xx} can be bounded below by expressions involving the local trace at x.

Proposition 4.5 *Let P be a symmetric operator on $(H, <.|.>)$ such that $(-P)$ is positive. Then, using again the notation (4, 5),*

$$\mathcal{L}[A_{xx}] \;\geq\; \frac{|\mathrm{Tr}(E_x P)|^4}{64\, n^2} \tag{27}$$

$$\mathcal{L}[A_{xx}] \;\geq\; |\mathrm{Tr}(E_x P)|^2 \;\inf \sigma\Big(A_{xx}|_{E_x(H)} \Big) \,. \tag{28}$$

Proof. According to Lemma 4.3 **(i)**, the operator $(-P(x,x)) \;:\; E_x(H) \mapsto E_x(H)$ is positive. Lemma 4.4 tells us that the zeros of the characteristic polynomial of $P(x,x)$, which we denote by $(\mu_j)_{j=1,\ldots,2n}$, are all real and have the ordering

$$\mu_1 \leq \cdots \leq \mu_n \;\leq\; 0 \;\leq\; \mu_{n+1} \leq \cdots \leq \mu_{2n} \,. \tag{29}$$

This allows us to write the local trace as follows,

$$\mathrm{Tr}(E_x P) \;=\; \sum_{j=1}^{2n} \mu_j \;=\; \sum_{i=n+1}^{2n} |\mu_i| - \sum_{j=1}^{n} |\mu_j| \;=\; \frac{1}{n} \sum_{i=n+1}^{2n} \sum_{j=1}^{n} \big(|\mu_i| - |\mu_j|\big) \,, \tag{30}$$

where the last equality is obvious if one notices that when for example adding up the $|\mu_i|$, the sum over j can be carried out giving a factor n. We now take absolute values and increase the right side by taking more summands,

$$|\mathrm{Tr}(E_x P)| \;\leq\; \frac{1}{n} \sum_{i,j=1}^{2n} \Big| |\mu_i| - |\mu_j| \Big|. \tag{31}$$

Now we can proceed with Hölder's inequality to obtain

$$|\mathrm{Tr}(E_x P)| \;\leq\; \frac{1}{n} \left(4n^2\right)^{\frac{1}{2}} \left(\sum_{i,j=1}^{2n} \Big| |\mu_i| - |\mu_j| \Big|^2 \right)^{\frac{1}{2}} \;=\; 2 \left(\sum_{i,j=1}^{2n} \Big| |\mu_i| - |\mu_j| \Big|^2 \right)^{\frac{1}{2}} \tag{32}$$

$$|\mathrm{Tr}(E_x P)| \;\leq\; \frac{1}{n} \left(4n^2\right)^{\frac{3}{4}} \left(\sum_{i,j=1}^{2n} \Big| |\mu_i| - |\mu_j| \Big|^4 \right)^{\frac{1}{4}} \;=\; \sqrt{8n} \left(\sum_{i,j=1}^{2n} \Big| |\mu_i| - |\mu_j| \Big|^4 \right)^{\frac{1}{4}}. \tag{33}$$

Since the matrix A_{xx} is the square of $P(x,x)$, the zeros of its characteristic polynomial, again denoted by $(\lambda_j)_{j=1,\ldots,2n}$, satisfy the relations

$$0 \;\leq\; \lambda_j \;=\; |\mu_j|^2 \qquad \forall\, j = 1, \ldots, 2n\,. \tag{34}$$

Using the formula (8) for the Lagrangian, we then obtain

$$\mathcal{L}(A_{xx}) \;=\; \sum_{i,j=1}^{2n} (\lambda_i - \lambda_j)^2 \;=\; \sum_{i,j=1}^{2n} (|\mu_i| - |\mu_j|)^2 \, (|\mu_i| + |\mu_j|)^2\,. \tag{35}$$

The last expression can be bounded from below in two ways. Either we use the inequality

$$(|\mu_i| + |\mu_j|)^2 \;\geq\; (|\mu_i| - |\mu_j|)^2$$

and apply (33) to obtain (27). Or we use the estimate

$$\sum_{i,j=1}^{2n} (|\mu_i| - |\mu_j|)^2 \, (|\mu_i| + |\mu_j|)^2 \;\geq\; 4 \, \min_j |\mu_j|^2 \sum_{i,j=1}^{2n} (|\mu_i| - |\mu_j|)^2$$

together with (34) and (32), giving (28). ∎

The inequality (27) immediately gives a positive lower bound for the action.

Corollary 4.6 *In the setting of Section 1, the action (6) satisfies the inequality*

$$\mathcal{S}[P] \;\geq\; \frac{f^4}{64\, n^2\, m^3}\,.$$

Proof. We first apply Hölder's inequality in (3),

$$f \;\leq\; m^{\frac{3}{4}} \left(\sum_{x \in M} |\mathrm{Tr}(E_x P)|^4 \right)^{\frac{1}{4}}\,. \tag{36}$$

Dropping the contributions for $x \neq y$ in (6), we obtain the lower bound

$$\mathcal{S}[P] \;=\; \sum_{x,y \in M} \mathcal{L}[A_{xy}] \;\geq\; \sum_{x \in M} \mathcal{L}[A_{xx}]\,, \tag{37}$$

and using (27) and (36) gives the claim. ■

We point out that for the estimates of Proposition 4.5 it is crucial that the maximal dimensions of the positive and negative definite subspaces of $E_x(H)$ coincide. If we considered more general discrete space-times with spin dimension (p, q), then in the case $p \neq q$ the last transformation in (30) would no longer be valid, and the statements of Lemma 4.5 and Corollary 4.6 would break down. This can most easily be seen in the extreme example of spin dimension $(0, q)$, where by "localizing" q particles similar to (19) at the space-time point x we could arrange that $P(x, x) = \mathbb{1}_{|E_x(H)}$. Then A_{xx} would be the identity, and $\mathcal{L}[A_{xx}]$ would vanish, although the local trace $\mathrm{Tr}(E_x P)$ would be equal to q. By localizing all particles in this way at individual space-time points, we could construct minimizers of the action which are not particularly interesting. This consideration is the reason why in this paper we only consider systems of spin dimension (n, n). We feel that, apart from their physical significance, these systems are the ones for which the minimizers of our variational principle should have the most interesting mathematical structure.

5 Statement of the Main Results

For the estimates of Proposition 4.5 it was essential that the operator $(-P)$ was positive, but P did not necessarily need to be a projector. This is the motivation for considering our variational principle on a more general class of operators, which we now introduce.

Def. 5.1 *An operator P on an inner product space $(H, <.|.>)$ is said to be* **of class** $\mathcal{P}^f$ *if*

(i) *The operator $(-P)$ is positive.*

(ii) *The operator P has trace f and rank at most f.*

Clearly, the fermionic projector as introduced in Section 1 is of class $\mathcal{P}^f$. The next theorem shows that our variational principle behaves nicely on $\mathcal{P}^f$.

Theorem 5.2 *The action (6) attains its minimum in $\mathcal{P}^f$, i.e. there is $P \in \mathcal{P}^f$ with*

$$\mathcal{S}[P] \;=\; \inf_{Q \in \mathcal{P}^f} S[Q]\,.$$

The theorem makes no statement on uniqueness, and indeed we do not see a reason why the minimizer should be unique. For the proof we will use the direct method of the calculus of variations. By starting with different minimal sequences, our method allows to construct all minimizers.

It is not known whether the minimizer from Theorem 5.2 is a projector. Also, one might want to consider the variational principle under various constraints. In this respect, the following theorem is very useful.

Theorem 5.3 *Suppose that $(P_k)_{k \in \mathbb{N}}$ is a sequence of operators of class $\mathcal{P}^f$ such that the following two conditions are satisfied,*

(A) *The series $\mathcal{S}[P_k]$ is bounded.*

(B) *The local trace is bounded away from zero in the sense that for suitable $\delta > 0$,*

$$|\mathrm{Tr}(E_x P_k)| \;\geq\; \delta \qquad \forall\, k \in \mathbb{N},\; x \in M\,.$$

Then there is a subsequence (P_{k_l}) and a sequence of gauge transformations $U_l \in \mathcal{G}$ such that the gauge-transformed operators have a limit

$$P \;:=\; \lim_{l\to\infty} U_l P_{k_l} U_l^{-1}\,,$$

and the actions converge,

$$\mathcal{S}[P] \;=\; \lim_{l\to\infty} \mathcal{S}[P_{k_l}]\,.$$

Thus the only obstruction is that in a minimal sequence the local trace (as defined after (3)) must not go to zero at any space-time point. It is an open problem whether this condition is only a technical condition needed in our proof, or whether it is really necessary for the theorem to hold. Also, it is not clear whether condition **(B)** will be satisfied for a general minimal sequence of fermionic projectors (see also the remark at the end of Section 6).

Using the ansatz

$$\Phi(t,r,\vartheta,\varphi) = e^{-i\omega t - ik\varphi}\, R(r)\, \Theta(\vartheta), \tag{38}$$

the wave equation can be separated into an angular and a radial ODE,

In order to give an example for a variational principle with constraints, we finally consider homogeneous operators.

Def. 5.4 *A fermionic projector P is called* **homogeneous** *if for any $x_0, x_1 \in M$ there is a permutation $\sigma : M \mapsto M$ with $\sigma(x_0) = x_1$ and a gauge transformation $U \in \mathcal{G}$ such that*

$$P(\sigma(x), \sigma(y)) \;=\; U\, P(x,y)\, U^{-1} \qquad \forall\, x, y \in M\,.$$

We remark that this definition generalizes the usual notion of "homogeneity" as defined via a symmetry group K acting transitively on space-time. Namely, in this case we take for any $x_0, x_1 \in M$ a group element $g \in K$ with $g(x_0) = x_1$ and set $\sigma(x) = g(x)$, together with unitary maps $U_x : E_x(H) \to E_{g(x)}(H)$ which identify the corresponding spinor spaces. Homogeneous operators seem of physical interest because the vacuum should be described by a homogeneous fermionic projector.

Corollary 5.5 *Varying P in the class of homogeneous fermionic projectors, the action (6) attains its minimum.*

Proof. Let P be a homogeneous fermionic projector. Then, with σ and U as in Definition 5.4,

$$\mathrm{Tr}(E_{x_1} P) \;=\; \mathrm{Tr}(P(x_1, x_1)) \;=\; \mathrm{Tr}(U\, P(x_0, x_0)\, U^{-1}) \;=\; \mathrm{Tr}(P(x_0, x_0))\,.$$

Thus the local trace is the same at all space-time points, and from (3) we conclude that

$$\mathrm{Tr}(E_x P) \;=\; \frac{f}{m} \qquad \forall\, x \in M\,.$$

Hence we can apply Theorem 5.3 with $\varepsilon = f/m$. ∎

6 A Lower Bound for the Local Trace

In this section we shall analyze how the infimum of our action depends on the number of space-time points. Thus for fixed spin dimension (n, n) and a fixed number of particles f, we consider for any $m \in \mathbb{N}$ a discrete space-time $(H, <.|.>, (E_x)_{x \in M})$ with $m = \# M$ (note that this discrete space-time is unique up to isomorphisms). We define

$$\left. \begin{array}{rcl} I(f,m) &=& \inf\{\mathcal{S}[P] \,|\, P \in \mathcal{P}^f\} \\ J(f,m) &=& \inf\{\mathcal{S}[P] \,|\, P \text{ fermionic projector}\} \end{array} \right\} . \tag{39}$$

In the case $f > mn$, when the set of fermionic projectors is empty, we set $J(f,m) = \infty$. The functions I and J are strictly positive by Corollary 4.6. Also, it is obvious that $I(f,m) \leq J(f,m)$. Apart from simple examples as considered in Section 3, nothing is known about the values of $I(f,m)$ and $J(f,m)$. In particular, it would be interesting to know whether $I(f,m)$ is always strictly smaller than $J(f,m)$.

Our next lemma shows that the functions $I(f,m)$ and $J(f,m)$ are strictly decreasing in the parameter m. This can be understood from the fact that if m is increased, the particles can spread out over more space-time points, making the infimum of the action smaller.

Lemma 6.1 *The functions I and J defined by (39) satisfy the inequalities*

$$I(f, m+1) \;\leq\; \left(1 - \frac{3}{4m}\right) I(f,m) \,, \qquad J(f, m+1) \;\leq\; \left(1 - \frac{3}{4m}\right) J(f,m) \,. \tag{40}$$

Proof. Let P be an operator of class P^f in a discrete space-time $(H, <.|.>, (E_x)_{x \in M})$ with $M = \{1, \ldots, m\}$. Introducing a discrete space-time $(\hat{H}, <.|.>, \hat{M})$ where $\hat{M} = \{0, \ldots, m\}$ consists of one more space-time point, there is a unitary transformation U from H to the subspace $K = \oplus_{x=1}^m \hat{E}_x(\hat{H})$ of $\hat{H}$ which maps the space-time projectors E_x to the $\hat{E}_x$ in the sense that $E_x = U^{-1} \hat{E}_x U$ for all $x = 1, \ldots, m$. In other words, we can identify $(H, <.|.>, (E_x)_{x \in M})$ with the discrete space-time $(K, <.|.>, (\hat{E}_x)_{x \in M})$. Using this identification, the operator P maps K to itself, and extending it by zero to $\hat{E}_0(\hat{H})$, we obtain an operator

$$P \,:\, \hat{H} \mapsto \hat{H} \qquad \text{with} \qquad E_0\, P \;=\; 0 \;=\; P\, E_0 \,.$$

Since $P(x,y)$ vanishes when $x = 0$ or $y = 0$, the action of P is given by

$$\mathcal{S}[P] \;=\; \sum_{x,y \in M} \mathcal{L}[A_{xy}] \,,$$

and this also shows that our reinterpretation of P did not change its action.

Our method is to construct a unitary transformation $V \,:\, \hat{H} \mapsto \hat{H}$ such that the action of the operator

$$\hat{P} \;:=\; V\, P\, V^{-1} \tag{41}$$

is strictly smaller than that of P. First, in

$$\mathcal{S}[P] \;=\; \sum_{x \in M} \left(\sum_{y \in M} \mathcal{L}[A_{xy}] \right)$$

we choose a point $x \in M$ for which the inner sum is maximal. Then

$$\sum_{y \in M} \mathcal{L}[A_{xy}] \;\geq\; \frac{\mathcal{S}[P]}{m} \,. \tag{42}$$

We choose V such that it is the identity on the subspaces $\hat{E}_y(\hat{H})$ for $y \not\in \{0, x\}$, whereas on the subspace $\hat{E}_0(\hat{H}) \oplus \hat{E}_x(\hat{H})$ it has in block matrix notation the form

$$V \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix} , \qquad V^{-1} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} .$$

A short calculation shows that the discrete kernels of P and $\hat{P}$ are related by

$$\left\{ \begin{array}{lllll} \hat{P}(y,z) & = & P(y,z) & & \text{if } y,z \not\in \{0,x\} \\ \hat{P}(y,z) & = & \frac{1}{\sqrt{2}}\, P(x,z)\,,\; \hat{P}(z,y) & = \frac{1}{\sqrt{2}}\, P(z,x) & \text{if } y \in \{0,x\} \text{ and } z \not\in \{0,x\} \\ \hat{P}(z,y) & = & \frac{1}{2}\, P(x,x) & & \text{if } y,z \in \{0,x\}\,. \end{array} \right.$$

Using that the Lagrangian is homogeneous in P of degree four, we obtain with the obvious notation $\hat{A}_{xy} = \hat{P}(x,y)\,\hat{P}(y,x)$ that

$$\begin{aligned} \mathcal{S}[\hat{P}] &= \sum_{y,z \not\in \{0,x\}} \mathcal{L}[\hat{A}_{y,z}] \;+\; 4 \sum_{y \not\in \{0,x\}} \mathcal{L}[\hat{A}_{xy}] \;+\; 4\, \mathcal{L}[\hat{A}_{xx}] \\ &= \sum_{y,z \not\in \{0,x\}} \mathcal{L}[A_{y,z}] \;+\; \sum_{y \not\in \{0,x\}} \mathcal{L}[A_{xy}] \;+\; \frac{1}{4}\, \mathcal{L}[A_{xx}] \end{aligned}$$

and thus

$$\mathcal{S}[P] - \mathcal{S}[\hat{P}] \;=\; \sum_{y \not\in \{0,x\}} \mathcal{L}[A_{xy}] \;+\; \frac{3}{4}\, \mathcal{L}[A_{xx}] \;\geq\; \frac{3}{4} \sum_{y \in M} \mathcal{L}[A_{xy}] \,.$$

Now we can put in (42) to obtain the inequality

$$\mathcal{S}[\hat{P}] \;\leq\; \left(1 - \frac{3}{4m}\right) \mathcal{S}[P] \,.$$

Consider a minimal sequence $P_k \in \mathcal{P}^f(H)$. Then

$$I(m+1, f) \;\leq\; \mathcal{S}[\hat{P}_k] \;\leq\; \left(1 - \frac{3}{4m}\right) \mathcal{S}[P_k] \;\stackrel{k \to \infty}{\longrightarrow}\; \left(1 - \frac{3}{4m}\right) I(m,f) \,,$$

proving the left inequality in (40). Similarly, if we let P_k be a minimal sequence of projectors, then the corresponding operators $\hat{P}_k$ are also projectors (because (41) is a unitary transformation), and we obtain the right inequality in (40). ■

Proposition 6.2 *Let $P_k \in \mathcal{P}^f$ be a minimal sequence for the action (6), i.e.*

$$\lim_{k \to \infty} \mathcal{S}[P_k] \;=\; I(f,m) \,. \tag{43}$$

Then there is $\delta > 0$ such that

$$\mathrm{Tr}(E_x P_k) \;\geq\; \delta \qquad \forall\, k \in \mathbb{N},\; x \in M \,.$$

Proof. We argue by contradiction. Assume that there is $x \in M$ and a subsequence of (P_k) (again denoted by $(P_k)_{k\in\mathbb{N}}$) such that $\lim_{k\to\infty} \mathrm{Tr}(E_x P_k) \leq 0$. Then we must clearly have more than one space-time point, because otherwise $\mathrm{Tr}(E_x P) = \mathrm{Tr}(P) = f > 0$. We introduce the projector $F = \mathbb{1} - E_x$ and define for large k the series of operators Q_k by

$$Q_k \;=\; c_k\, F P_k F \qquad \text{with} \qquad c_k \;:=\; \frac{f}{\mathrm{Tr}(F P_k)}\,. \tag{44}$$

Since $\mathrm{Tr}(FP_k) = \mathrm{Tr}(P_k) - \mathrm{Tr}(E_x P_k) \to f$, we know that

$$\lim_{k\to\infty} c_k \;\leq\; 1\,. \tag{45}$$

According to Lemma 4.3 **(i)**, the operators $(-Q_k)$ are positive, and we normalized them such that $\mathrm{Tr}\, Q_k = f$. Therefore, the operators Q_k are again of class $\mathcal{P}^f$. Since they vanish identically on $E_x(H)$, we can regard them as operators in a discrete space-time consisting of $m-1$ space-time points, and thus

$$\mathcal{S}[Q_k] \;\geq\; I(m-1, f)\,. \tag{46}$$

Using that the Lagrangian is homogeneous of degree four, we obtain furthermore

$$\mathcal{S}[Q_k] \;=\; c_k^4\, \mathcal{S}[F P_k F] \;\leq\; c_k^4\, \mathcal{S}[P_k]\,, \tag{47}$$

where in the last step we used that the Lagrangians of P_k and FP_kF coincide away of the space-time point x; more precisely,

$$\mathcal{S}[P_k] - \mathcal{S}[F P_k F] \;=\; \mathcal{L}[A_{xx}] + 2 \sum_{y \neq x} \mathcal{L}[A_{xy}] \;\geq\; 0\,.$$

Taking in (47) the limit $k \to \infty$ and using (45, 43), we obtain in view of (46) that

$$I(m-1, f) \;\leq\; \lim_{k\to\infty} \mathcal{S}[Q_k] \;\leq\; \lim_{k\to\infty} \mathcal{S}[P_k] \;=\; I(m, f)\,,$$

in contradiction to Lemma 6.1. ∎

This proposition ensures that for a minimal sequence in $\mathcal{P}^f$, condition **(B)** of Theorem 5.3 is satisfied. Thus Theorem 5.2 follows from Theorem 5.3. We point out that, unfortunately, the above argument does not apply to a minimal sequence of projectors, because the property to be idempotent gets lost when the operators are restricted similar to (44) to a subspace of H.

7 Gauge Fixing, Rescaling

We now enter the proof of Theorem 5.3. Thus let $(P_k)_{k\in\mathbb{N}}$ be a sequence of operators of class $\mathcal{P}^f$ which satisfy conditions **(A)** and **(B)**. We again choose a basis in H and let $(.|.)$ be the canonical scalar product on $\mathbb{C}^{2nm}$. We let $\|.\|$ be the corresponding Hilbert-Schmidt norm, $\|A\| := (\mathrm{Tr}(A^\dagger A))^{\frac{1}{2}}$.

Our first task is to treat the non-compact gauge group $\mathcal{G}$ (as defined after (13)). We denote the equivalence class of gauge-equivalent operators by $\langle . \rangle_{\mathcal{G}}$, i.e.

$$\langle P \rangle_{\mathcal{G}} \;=\; \{ U P U^{-1} \,|\, U \in \mathcal{G} \}\,.$$

We consider for any fixed $k \in \mathbb{N}$ the variational principle

$$\text{minimize} \quad \{\|Q\| \mid Q \in \langle P_k \rangle_{\mathcal{G}}\}\,. \tag{48}$$

If $(Q_l)_{l \in \mathbb{N}}$ is a minimal sequence of this variational principle, the Hilbert-Schmidt norms of the Q_l are uniformly bounded. Thus we can use a compactness argument to select a convergent subsequence. We conclude that the variational principle (48) attains its minimum. We choose for each k a minimizer and denote it by $\hat{P}_k$. We point out that the above construction of the $\hat{P}_k$ involves the norm $\|.\|$ and thus depends on the choice of our basis of H. This will be no problem in what follows because the minimizers obtained by choosing different norms will be gauge equivalent. We refer to our method of arbitrarily choosing one representative of each gauge equivalence class as *gauge fixing*; it can be understood in analogy to the gauge fixing used in electrodynamics or in general relativity.

In the case that the sequence of operators $(\hat{P}_k)$ has a subsequence of bounded Hilbert-Schmidt norm, we can by compactness choose a subsequence which converges to an operator P. Since our action is obviously continuous, this P would be the desired minimizer. Thus it remains to consider the case when the Hilbert-Schmidt norm is unbounded for any subsequence of $(\hat{P}_k)$; in other words, that

$$\|\hat{P}_k\| \;\to\; \infty\,. \tag{49}$$

We introduce new operators R_k by rescaling the $\hat{P}_k$,

$$R_k \;=\; \alpha_k\, \hat{P}_k \qquad \text{with} \qquad \alpha_k \;:=\; \frac{1}{\|\hat{P}_k\|} \;\stackrel{k\to\infty}{\longrightarrow}\; 0\,.$$

Then obviously $\|R_k\| \equiv 1$, and thus we can, again after choosing a subsequence, assume that the R_k converge,

$$R_k \;\to\; R\,.$$

It is clear from their construction that the operators R_k and R have the following properties: The operators $(-R_k)$ and $(-R)$ are positive and normalized,

$$\|R_k\| \;=\; 1 \;=\; \|R\|\,. \tag{50}$$

Their traces are given by

$$\mathrm{Tr}(R_k) \;=\; \alpha_k\, f\,, \qquad \mathrm{Tr}(R) \;=\; 0\,, \tag{51}$$

and their action is computed to be

$$\mathcal{S}[R_k] \;=\; \alpha_k^4\, \mathcal{S}[P_k]\,, \qquad \mathcal{S}[R] \;=\; 0\,. \tag{52}$$

Finally, the local trace of R_k is bounded from below,

$$|\mathrm{Tr}(E_x R_k)| \;\geq\; \delta\, \alpha_k \qquad \forall\, k \in \mathbb{N},\; x \in M\,. \tag{53}$$

8 Existence of Minimizers

Our goal is to show that the properties (50–53) contradict the fact that the $\hat{P}_k$ are minimizers of (48) (this then implies that the case (49) cannot occur, completing the proof of Theorem 5.3). For any $x \in M$, the operator $T := -R(x,x)$ is positive according to

Lemma 4.3 **(i)**. From Lemma 4.4 we conclude that the zeros $(\mu_j)_{j=1,\ldots,2n}$ of its characteristic polynomial are all real and ordered as in (29). Since $\mathcal{L}[T^2] = 0$, the absolute values of the μ_j must all be equal, and thus there is a parameter $\mu \geq 0$ such that

$$\mu_1 = \ldots \mu_n = -\mu \qquad \text{and} \qquad \mu_{n+1} = \ldots \mu_{2n} = \mu\,.$$

In the case $\mu > 0$, we know by the continuity of the spectrum that for large k,

$$\inf \sigma\Big((T_k)^2|_{E_x(H)} \Big) \;\geq\; \frac{\mu^2}{2}$$

(with $T_k := -R_k(x,x)$). Combining the lower bound (28) with (52, 53), we obtain

$$\frac{\mu^2}{2}\, \delta^2\, \alpha_k^2 \;\leq\; \mathcal{L}[T_k^2] \;\leq\; \mathcal{S}[R_k] \;=\; \alpha_k^4\, \mathcal{S}[P_k]\,,$$

and taking the limit $k \to \infty$, we obtain a contradiction to the fact that the $\mathcal{S}[P_k]$ are uniformly bounded away from zero according to Corollary 5.5.

It remains to consider the case $\mu = 0$ where the operator T is nilpotent. As in the proof of Lemma 4.4, we approximate T by the strictly positive operators $T_\varepsilon = T + \varepsilon S$. Diagonalizing the T_ε by unitary transformations U_ε on $E_x(H)$, the diagonal matrices $U_\varepsilon T_\varepsilon U_\varepsilon^{-1}$ converge to zero as $\varepsilon \to 0$. Hence for any $\Psi \in H$,

$$<\!\Psi \,|\, U_\varepsilon\, T\, U_\varepsilon\, \Psi\!> \;+\; <\!\Psi \,|\, U_\varepsilon\, S\, U_\varepsilon\, \Psi\!> \;=\; <\!\Psi \,|\, U_\varepsilon\, T_\varepsilon\, U_\varepsilon\, \Psi\!> \;\stackrel{\varepsilon \to 0}{\longrightarrow}\; 0\,.$$

Since the summands on the left are both positive, we conclude that $<\!\Psi \,|\, U_\varepsilon T U_\varepsilon\, \Psi\!> \to 0$ for all $\Psi \in H$ and thus

$$\lim_{\varepsilon \to 0} U_\varepsilon T U_\varepsilon \;=\; 0\,.$$

For given $\kappa > 0$ we choose ε such that $\|U_\varepsilon T U_\varepsilon\| < \kappa/2$ and subsequently k so large that $\|T_k - T\| < \kappa/(2\, \|U_\varepsilon\|\, \|U_\varepsilon^{-1}\|)$. Then

$$\|U_\varepsilon\, T_k\, U_\varepsilon^{-1}\| \;\leq\; \|U_\varepsilon\|\, \|T_k - T\|\, \|U_\varepsilon^{-1}\| \;+\; \|U_\varepsilon\, T\, U_\varepsilon^{-1}\| \;\leq\; \kappa\,.$$

Since κ can be chosen arbitrarily small, we conclude that there is a subsequence of the T_k (which we denote again by $(T_k)_{k \in \mathbb{N}}$) together with unitary transformations U_k such that

$$\lim_{k \to \infty} U_k\, T_k\, U_k^{-1} \;=\; 0\,.$$

Extending the U_k by the identity to the subspaces $E_y(H)$, $y \neq x$, we obtain a sequence of gauge transformations such that

$$U_k\, R_k\, U_k^{-1} \;\to\; \tilde{R}\,.$$

Since these gauge transformations act only on $E_x(H)$, it is clear that $R(y,z) = \tilde{R}(y,z)$ if $y,z \neq x$. By construction, $\tilde{R}(x,x) = 0$. The Schwarz inequality, Lemma 4.3 **(ii)**, tells us that also the entries $\tilde{R}(x,y)$ and $\tilde{R}(y,x)$ for $y \neq x$ vanish. Since we chose the operators $\hat{P}_k$ such that their Hilbert-Schmidt norm was minimal among all gauge-equivalent operators, the Hilbert-Schmidt norm of the operators R_k (which were obtained from the $\hat{P}_k$ only by rescaling) cannot be decreased by a subsequent gauge transformation, and thus $\|U_k\, R_k\, U_k^{-1}\| \geq \|R_k\|$. Taking the limit $k \to \infty$, we find that $\|\tilde{R}\| \geq \|R\|$. Since these operators coincide up to matrix elements where $\tilde{R}$ vanishes, the operators $\tilde{R}$ and R must coincide. In particular, $R(x,x) = 0$.

We conclude that the diagonal entries $R(x,x)$ of R all vanish. Again applying the Schwarz inequality, Lemma 4.3 **(ii)**, we see that the off-diagonal entries of R are also zero. Thus $R = 0$, in contradiction to (50).

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

References

[1] J. Bognar, "Indefinite Inner Product Spaces," Ergebnisse der Mathematik und ihrer Grenzgebiete, Band **78**, Springer-Verlag, New York-Heidelberg (1974)

[2] F. Finster, "The Principle of the Fermionic Projector," *International Press*, to appear (2005)

[3] I. Gohberg, P. Lancaster, L. Rodman, "Matrices and Indefinite Scalar Products," *Birkhäuser Verlag* (1983)

NWF I – Mathematik, Universität Regensburg, 93040 Regensburg, Germany,
`Felix.Finster@mathematik.uni-regensburg.de`