Universität Regensburg
Mathematik

On the Support of Minimizers of Causal Variational Principles

Felix Finster and Daniela Schiefeneder

Preprint Nr. 22/2010

ON THE SUPPORT OF MINIMIZERS OF CAUSAL VARIATIONAL PRINCIPLES

FELIX FINSTER AND DANIELA SCHIEFENEDER

DECEMBER 2010

Abstract. A class of causal variational principles on a compact manifold is introduced and analyzed both numerically and analytically. It is proved under general assumptions that the support of a minimizing measure is either completely timelike, or it is singular in the sense that its interior is empty. In the examples of the circle, the sphere and certain flag manifolds, the general results are supplemented by a more detailed and explicit analysis of the minimizers. On the sphere, the minimal action is estimated from above and below.

Contents

1. Introduction

Causal variational principles were proposed in [4] as an approach for formulating relativistic quantum field theory (for surveys see [7, 9]). More recently, they were introduced in a broader mathematical context as a class of nonlinear variational principles defined on measure spaces [8]. Except for the examples and general existence results in [6, 3, 8] and the symmetry breaking effect in the discrete setting [5], almost nothing is known on the structure of the minimizers. In the present paper, we turn attention to the analysis of minimizing measures in the continuous setting. For simplicity, we restrict attention to variational principles on a compact manifold which generalize the causal variational principles in spin dimension one with two prescribed eigenvalues

Supported in part by the Deutsche Forschungsgemeinschaft.

(see [8, Chapter 1]). But our methods are developed with a view to possible extensions to the non-compact setting and to a general spin dimension.

More precisely, in Section 2 we introduce a class of causal variational principles on a compact manifold $\mathcal{F}$ and explain how this setting fits into the general context. As more specific model examples, we introduce variational principles on the circle, on the sphere, and on the flag manifold $\mathcal{F}^{1,2}(\mathbb{C}^f)$. In Section 3, we present numerical results on the sphere (see Figure 3) and discuss all the main effects which will be treated analytically later on. In Section 4, we derive general results on the structure of the minimizers. We first derive the corresponding Euler-Lagrange equations and conditions for minimality (see Lemma 4.4 and Lemma 4.5). We then prove under general assumptions that the minimizers are either generically timelike (see Definition 4.7) or the support of the minimizing measure ρ defined by

$$\operatorname{supp} \rho = \{x \in \mathcal{F} \mid \rho(U) \neq 0 \text{ for every open neighborhood } U \text{ of } x\}$$

is singular in the sense that its interior is empty (see Theorem 4.15 and Theorem 4.17). In the following sections, we apply these general results to our model examples and derive more detailed information on the minimizers. In Section 5, we consider the variational principle on the circle. After briefly discussing numerical results (see Figure 4), we prove a "phase transition" between generically timelike minimizers and minimizers with singular support and construct many minimizers in closed form (see Corollary 5.2 and Theorem 5.4). In Section 6, the variational principle on the sphere is considered. We again prove the above phase transition (see Corollary 6.1) and estimate the action from above and below (see Figure 6 and Proposition 6.3). Finally, in Section 7 we apply our general results to flag manifolds (see Theorem 7.1). Moreover, we prove that minimizers with singular support exist (see Theorem 7.2) and give an outlook on generically timelike minimizers.

2. Preliminaries

Before introducing our mathematical framework, we briefly put it in the general context. Causal variational principles can be formulated either in indefinite inner product spaces on an underlying space-time (the "space-time representation"; see [4, 6] and [8, Chapters 3 and 4]) or in terms of the so-called local correlation matrices acting on the space of occupied particle states (the "particle representation"; see [8, Chapters 1 and 2]). Here we shall always work in the *particle representation*, whereas for the connection to the space-time representation we refer to the constructions in [8, Section 3.2] and [10]. Thus as in [8, Chapters 1 and 2], we begin with a positive measure space (M, μ) normalized to $\mu(M) = 1$. Moreover, for a given integer parameter f, we consider a closed subset $\mathcal{F}$ of the Hermitian $(f \times f)$-matrices. For technical simplicity, we here assume that $\mathcal{F}$ is a *compact submanifold* of $\text{Mat}(\mathbb{C}^f)$; this covers the variational principle with prescribed eigenvalues as considered in [8, Chapter 1 and Section 2.1]. Generally speaking, in a causal variational principle one minimizes a certain action $\mathcal{S}[F]$ under variations of a measurable function $F : M \to \mathcal{F}$, imposing suitable constraints. Introducing the measure ρ on $\mathcal{F}$ by $\rho(\Omega) = \mu(F^{-1}(\Omega))$ (in other words, $\rho = F_* \mu$ is the push-forward measure), the action can be expressed by integrals over $\mathcal{F}$ (see [8, Section 1.2]),

$$\mathcal{S}[\rho] = \iint_{\mathcal{F} \times \mathcal{F}} \mathcal{L}(x, y)\, d\rho(x)\, d\rho(y)\,, \tag{2.1}$$

where the Lagrangian $\mathcal{L} \in C^{0,1}(\mathcal{F} \times \mathcal{F}, \mathbb{R}_0^+)$ is a given function. Now the only significance of the measure space (M, μ) is that it poses conditions on the possible form of the measure ρ. For example, in the discrete setting one chooses μ as the normalized counting measure on $M = \{1, \ldots, m\}$; then the support of ρ necessarily consists of at most m points. However, in the *continuous setting* under consideration here, we do not want to impose any conditions on the measure ρ, but instead ρ should be allowed to be any normalized positive regular Borel measure on $\mathcal{F}$. Then the measure space (M, μ) is no longer needed. For simplicity, we also leave out additional constraints (like the trace or identity constraints; see [8, Section 2.3]). This leads us to the following setting:

Let $\mathcal{F}$ be a smooth compact manifold (of arbitrary dimension). For a given function

$$\mathcal{D} \in C^\infty(\mathcal{F} \times \mathcal{F}, \mathbb{R}) \quad \text{being symmetric: } \mathcal{D}(x,y) = \mathcal{D}(y,x) \;\; \forall\, x, y \in \mathcal{F} \tag{2.2}$$

$$\text{and strictly positive on the diagonal: } \mathcal{D}(x,x) > 0\,, \tag{2.3}$$

we define the *Lagrangian* $\mathcal{L}$ by

$$\mathcal{L} = \max(0, \mathcal{D}) \in C^{0,1}(\mathcal{F} \times \mathcal{F}, \mathbb{R}_0^+)\,. \tag{2.4}$$

Introducing the *action* $\mathcal{S}$ by (2.1), our action principle is to

$$\text{minimize } \mathcal{S} \text{ under variations of } \rho \in \mathfrak{M}\,, \tag{2.5}$$

where $\mathfrak{M}$ denotes the set of all normalized positive regular Borel measures on $\mathcal{F}$. In view of the symmetric form of (2.1), it is no loss of generality to assume that $\mathcal{L}(x,y)$ is symmetric in x and y. Therefore, it is natural to assume that also $\mathcal{D}(x,y)$ is symmetric (2.2). If (2.3) were violated, every measure supported in the set $\{x : \mathcal{D}(x,x) \leq 0\}$ would be a minimizer. Thus the condition (2.3) rules out trivial cases.

The existence of minimizers follows immediately from abstract compactness arguments (see [8, Section 1.2]).

Theorem 2.1. *The infimum of the variational principle* (2.5) *is attained in* $\mathfrak{M}$.

Note that the minimizers will in general not be unique. Moreover, the abstract framework gives no information on how the minimizers look like.

The notion of causality can now be introduced via the sign of $\mathcal{D}$.

Definition 2.2 (causal structure).

$$\text{Two points } x, y \in \mathcal{F} \text{ are called } \left\{ \begin{array}{c} \text{timelike} \\ \text{lightlike} \\ \text{spacelike} \end{array} \right\} \text{ separated if } \left\{ \begin{array}{c} \mathcal{D}(x,y) > 0 \\ \mathcal{D}(x,y) = 0 \\ \mathcal{D}(x,y) < 0\,. \end{array} \right\}$$

We define the sets

$$\begin{array}{rcll} \mathcal{I}(x) &=& \{y \in \mathcal{F} \text{ with } \mathcal{D}(x,y) > 0\} & \text{open lightcone} \\ \mathcal{J}(x) &=& \{y \in \mathcal{F} \text{ with } \mathcal{D}(x,y) \geq 0\} & \text{closed lightcone} \\ \mathcal{K}(x) &=& \partial \mathcal{I}(x) \cap \partial\big(\mathcal{F} \setminus \mathcal{J}(x)\big) & \text{boundary of the lightcone}\,. \end{array}$$

Thus $y \in \mathcal{K}(x)$ if and only if the function $\mathcal{D}(x,.)$ changes sign in every neighborhood of y.

Our action is compatible with the causal structure in the sense that if x and y have lightlike or spacelike separation, then the Lagrangian vanishes, so that the pair (x,y) does not contribute to the action. Note that for a given minimizer ρ, we have similarly a causal structure on its support.

In order to work in more specific examples, we shall consider the following three model problems.

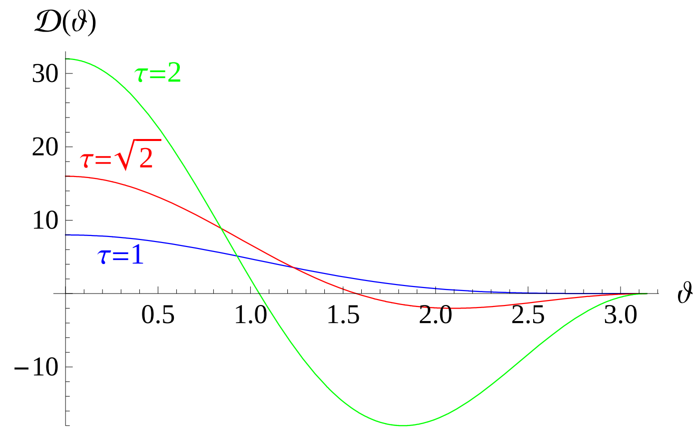

FIGURE 1. The function $\mathcal{D}$.

(a) *Variational principles on the sphere:*
We consider the setting of [8, Chapter 1] in the case $f = 2$ (see also [8, Examples 1.5, 1.6 and 2.8]). Thus for a given parameter $\tau \geq 1$, we let $\mathcal{F}$ be the space of Hermitian (2×2)-matrices whose eigenvalues are equal to $1 + \tau$ and $1 - \tau$. Writing a matrix $F \in \mathcal{F}$ as a linear combination of Pauli matrices,

$$F = \tau\, x \cdot \sigma + \mathbb{1} \qquad \text{with} \qquad x \in S^2 \subset \mathbb{R}^3 \,,$$

we can describe F by the unit vector x (here $\cdot$ denotes the scalar product in $\mathbb{R}^3$). Thus $\mathcal{F}$ can be identified with the unit sphere S^2. The function $\mathcal{D}$ is computed in [8, Example 2.8] to be

$$\mathcal{D}(x,y) = 2\tau^2 \left(1 + \langle x,y\rangle\right) \left(2 - \tau^2 \left(1 - \langle x,y\rangle\right)\right) . \tag{2.6}$$

This function depends only on the angle ϑ_{xy} between the points $x, y \in S^2$ defined by $\cos\vartheta_{xy} = \langle x,y\rangle$. Considered as function of $\vartheta \in [0,\pi]$, $\mathcal{D}$ has its maximum at $\vartheta = 0$ and is minimal if $\cos(\vartheta) = -\tau^{-2}$. Moreover, $\mathcal{D}(\pi) = 0$. Typical plots are shown in Figure 1. In the case $\tau > 1$, the function $\mathcal{D}$ has two zeros at π and

$$\vartheta_{\max} := \arccos\left(1 - \frac{2}{\tau^2}\right) . \tag{2.7}$$

In view of (2.4), the Lagrangian is positive if and only if $0 \leq \vartheta < \vartheta_{\max}$. Thus $\mathcal{I}(x)$ is an open spherical cap, and $\mathcal{J}(x)$ is its closure together with the antipodal point of x,

$$\mathcal{I}(x) = \{y : \langle x,y\rangle > 1 - \frac{2}{\tau^2}\} \,, \qquad \mathcal{J}(x) = \overline{\mathcal{I}(x)} \cup \{-x\}$$

If τ increases, the opening angle $\vartheta_{\max}$ of the lightcones gets smaller. In the degenerate case $\tau = 1$, the function $\mathcal{D}$ is decreasing, non-negative and has exactly one zero at $\vartheta = \pi$. Hence the Lagrangian $\mathcal{L}$ coincides with $\mathcal{D}$. All points on the sphere are timelike separated except for antipodal points. The lightcones are $\mathcal{I}(x) = S^2 \setminus \{-x\}$ and $\mathcal{J}(x) = S^2$.

If we regard ρ as a density on the sphere, the action (2.1) looks like the energy functional corresponding to a pair potential $\mathcal{L}$ (see for example [13]). Using physical notions, our pair potential is repelling (because $\mathcal{L}(\vartheta)$ is a decreasing function) and has short range (because $\mathcal{L}$ vanishes if $\vartheta \geq \vartheta_{\max}$).

(b) *Variational principles on the circle:*
In order to simplify the previous example, we set $\mathcal{F} = S^1$. For $\mathcal{D}$ we again choose (2.6).

(c) *Variational principles on the flag manifold* $\mathcal{F}^{1,2}(\mathbb{C}^f)$:
As in [8, Chapter 1], for a given parameter $\tau > 1$ and integer parameters $f > 2$ we let $\mathcal{F}$ be the space of Hermitian $(f \times f)$-matrices of rank two, whose nontrivial eigenvalues are equal to $1+\tau$ and $1-\tau$. Every $x \in \mathcal{F}$ is uniquely described by the corresponding eigenspaces U and V. By considering the chain $U \subset (U \cup V)$, x can be identified with an element of the flag manifold $\mathcal{F}^{1,2}(\mathbb{C}^f)$, the space of one-dimensional subspaces contained in a two-dimensional subspace of $\mathbb{C}^f$ (see [11]). It is a $(4f-6)$-dimensional compact manifold. Every $U \in \mathrm{U}(f)$ gives rise to the mapping $x \to UxU^{-1}$ on $\mathcal{F}$. This resulting group action of $\mathrm{U}(f)$ on $\mathcal{F}$ acts transitively, making $\mathcal{F}$ to a homogeneous space (see [11] for details).

For two points $x, y \in \mathcal{F}$, we denote the two non-trivial eigenvalues of the matrix product xy by $\lambda^{xy}_+, \lambda^{xy}_- \in \mathbb{C}$ and define the Lagrangian by

$$\mathcal{L}(x,y) = \frac{1}{2} \left(|\lambda^{xy}_+| - |\lambda^{xy}_-| \right)^2 .$$

This Lagrangian is $\mathrm{U}(f)$-invariant. In order to bring it into a more convenient form, we first note that by restricting to the image of y, the characteristic polynomial of xy changes only by irrelevant factors of λ,

$$\det(xy - \lambda \mathbb{1}) = \lambda^{f-2} \, \det\left((\pi_y xy - \lambda \mathbb{1})|_{\mathrm{Im}\, y} \right) ,$$

where π_y denotes the orthogonal projection to $\mathrm{Im}\, y$. It follows that λ^{xy}_+ and λ^{xy}_- are the eigenvalues of the (2×2)-matrix $\pi_y xy|_{\mathrm{Im}\, y}$. In particular,

$$\lambda^{xy}_+ \lambda^{xy}_- = \det(\pi_y xy|_{\mathrm{Im}\, y}) = \det(\pi_y x \pi_y|_{\mathrm{Im}\, y}) \, \det(y|_{\mathrm{Im}\, y}) \geq 0 ,$$

because the operator $\pi_y x \pi_y$ again has at most one positive and one negative eigenvalue. Moreover, the relation $\lambda^{xy}_+ + \lambda^{xy}_- = \mathrm{Tr}(xy) \in \mathbb{R}$ shows that the two eigenvalues are either both real and have the same sign or else form a complex conjugate pair, in which case the Lagrangian vanishes. Finally, using that $\left(\lambda^{xy}_+\right)^2 + \left(\lambda^{xy}_-\right)^2 = \mathrm{Tr}\big((xy)^2\big)$, the Lagrangian can be written in the form (2.4) with

$$\mathcal{D}(x,y) = \frac{1}{2} \left(\lambda^{xy}_+ - \lambda^{xy}_- \right)^2 = \mathrm{Tr}\big((xy)^2\big) - \frac{1}{2} \big(\mathrm{Tr}(xy)\big)^2 . \tag{2.8}$$

We finally comment on the limitations of our setting and mention possible generalizations. First, we point out that our structural results do not immediately apply in the cases when $\mathcal{F}$ is non-compact or when additional constraints are considered (see [8, Chapter 2]). However, it seems that in the non-compact case, our methods and results could be adapted to the so-called moment measures as introduced in [8, Section 2.3]. A promising strategy to handle additional constraints would be to first derive the corresponding Euler-Lagrange equations, treating the constraints with Lagrange multipliers. Then one could try to recover these Euler-Lagrange equations as those corresponding to an unconstrained variational problem on a submanifold $\mathcal{G} \subset \mathcal{F}$, where our methods could again be used. We finally point out that in the case of higher spin dimension $n > 1$, it is in general impossible to write the Lagrangian in the form (2.4) with a smooth function $\mathcal{D}$, because the Lagrangian is in general only Lipschitz continuous in the open light cone. A possible strategy would be to first show that the support of ρ lies on a submanifold $\mathcal{G} \subset \mathcal{F}$, and then to verify that by restricting $\mathcal{L}$ to $\mathcal{G} \times \mathcal{G}$, it becomes smooth in the open lightcones.

3. Motivation: Numerical Results on the Sphere

In order to motivate our general structural results, we now describe our findings in a numerical analysis of the variational principle on the sphere (see Example (a) on page 4). Clearly, in a numerical study one must work with discrete configurations. Our first attempt is to choose a finite number of points $x_1, \ldots, x_m \in S^2$ and to let ρ be the corresponding normalized *counting measure*, i.e.

$$\int_{S^2} f\, d\rho \;:=\; \frac{1}{m} \sum_{i=1}^m f(x_i) \qquad \forall f \in C^0(S^2)\,. \tag{3.1}$$

Then the action (2.1) becomes

$$\mathcal{S} = \frac{1}{m^2} \sum_{i,j=1}^m \mathcal{L}(x_i, x_j)\,. \tag{3.2}$$

By varying the points x_i for fixed m, we obtain a minimizer ρ_m. Since every normalized positive regular Borel measure can be approximated by such counting measures, we can expect that if we choose m sufficiently large, the measure ρ_m should be a good approximation of a minimizing measure $\rho \in \mathfrak{M}$ (more precisely, we even know that $\rho_m \to \rho$ as $m \to \infty$ with convergence in the weak $(C^0)^*$-topology).

If τ is sufficiently large, the opening angle of the lightcones is so small that the m points can be distributed on the sphere such that any two different points are spacelike separated. In this case, the action becomes

$$\mathcal{S} = \frac{1}{m}\, \mathcal{L}(\vartheta = 0)\,,$$

and in view of (3.2) this is indeed minimal. The question for which τ such a configuration exists leads us to the *Tammes problem*, a packing problem where the points are distributed on the sphere such that the minimal distance ϑ_m between distinct points is maximized, see [14]. More precisely, we know that the Tammes distribution is a minimizer of our action if τ is so large that $\vartheta_m > \vartheta_{\max}$. Until now, the Tammes problem is only solved if $m \leq 12$ and for $m = 24$ (for details see [2] and the references therein). For special values of m, the solutions of the Tammes problem are symmetric solids like the tetrahedron ($m = 4$), the octahedron ($m = 6$), the icosahedron ($m = 12$) and the snub cube ($m = 24$). Moreover, much research has been done on the numerical evaluation of spherical codes, mostly by N.J.A. Sloane, with the collaboration of R.H. Hardin, W.D. Smith and others, [14], containing numerical solutions of the Tammes problem for up to 130 points.

In the case $\vartheta_m < \vartheta_{\max}$, the measure ρ_m was constructed numerically using a simulated annealing algorithm[1]. In order to get optimal results, we used this algorithm iteratively, using either a Tammes distribution or previous numerical distributions as starting values. Using that $\mathcal{D}$ depends smoothly on τ, it is useful to increase or decrease τ in small steps, and to use the numerical minimizer as the starting configuration of the next step. In Figure 2, the numerically found $\mathcal{S}[\rho_m]$ is plotted for different values of m as a function of the parameter τ. The resulting plots look rather

[1]We use the *"general simulated annealing algorithm"* by J. Vandekerckhove, © 2006, `http://www.mathworks.de/matlabcentral/fileexchange/10548`.

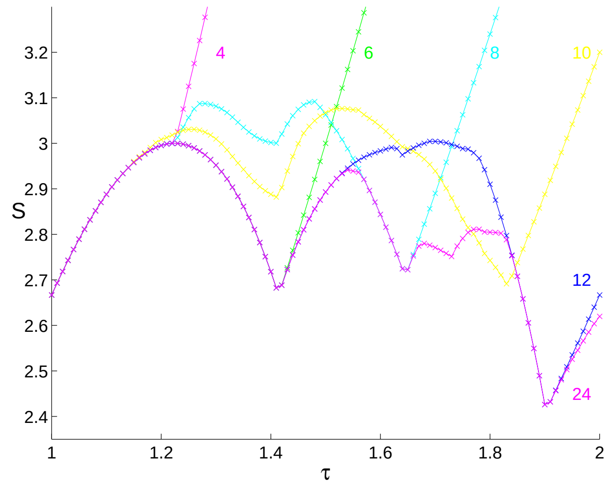

FIGURE 2. Numerical minima for the counting measure on the sphere.

complicated. The considered values for m are too small for extrapolating the limiting behavior as $m \to \infty$. Nevertheless, one observation turned out to be very helpful: Near $\tau \approx 1.2$, the plots for different values of m look the same. The reason is that some of the x_i coincide, forming "clusters" of several points. For example, in the case $m = 12$, the support of ρ only consists of six distinct points, each occupied by two x_i. A similar "clustering effect" also occurs for higher τ if m is sufficiently large.

These findings give the hope that for large m, the minimizers might be well-approximated by a measure supported at a few cluster points, with weights counting the number of points at each cluster. This was our motivation for considering a *weighted counting measure*. Thus for any fixed m, we choose points $x_1, \ldots, x_m \in S^2$ and corresponding weights $\rho_1, \ldots, \rho_m$ with

$$\rho_i \geq 0 \qquad \text{and} \qquad \sum_{i=1}^m \rho_i = 1 \, .$$

We introduce the corresponding measure ρ in generalization of (3.1) by

$$\int_{S^2} f \, d\rho := \sum_{i=1}^m \rho_i \, f(x_i) \qquad \forall f \in C^0(S^2) \, . \tag{3.3}$$

Seeking for numerical minimizers by varying both the points x_i and the weights ρ_i, we obtain the plots shown in Figure 3.

These plots suggest the following structure of the minimizers. Let us denote the the minimizing weighted counting measure for a given m by $\rho(m)$. Then for any fixed τ, the series $\mathcal{S}[\rho(m)]$ is monotone decreasing (this is obvious because every $\rho(m)$ can be realized by a weighted counting measure with $m_+ > m$ summands by choosing $m_+ - m$ weights equal to zero). The important observation is that there is an integer m_0 from where on the series stays constant, i.e.

$$\mathcal{S}[\rho(m_-)] > \mathcal{S}[\rho(m_0)] = \mathcal{S}[\rho(m_+)] \qquad \forall \, m_- < m_0 < m_+ \, .$$

This implies that the measure ρ_{m_0} is also a minimizer in the class of all Borel measures. This leads us to the following

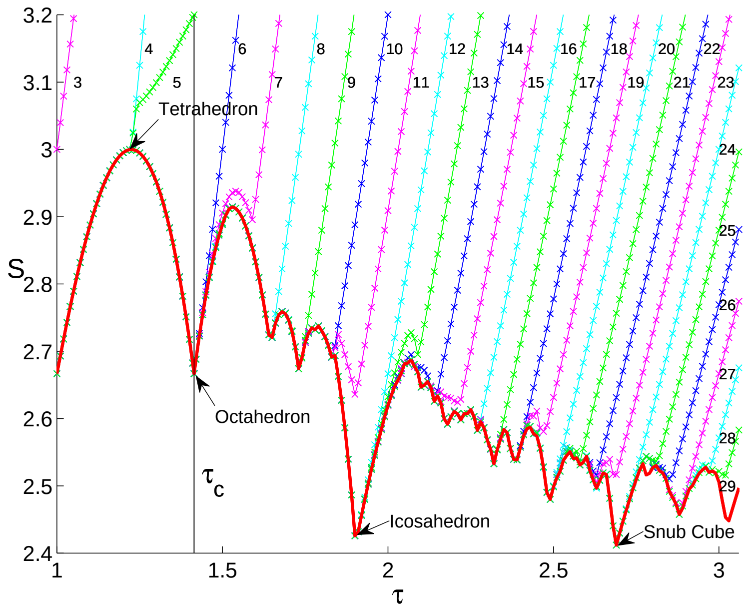

FIGURE 3. Numerical minima for the weighted counting measure on the sphere.

Conjecture 3.1. *For any $\tau \geq 1$, there is a minimizer $\rho \in \mathfrak{M}$ of the variational problem on the sphere which is a weighted counting measure supported at m_0 points.*

From Figure 3 we can read off the value of m_0 as a function of τ. Generally speaking, m_0 increases as τ gets larger. This corresponds to the fact that for increasing τ, the opening angle $\vartheta_{\max}$ of the light cones gets smaller, so that it becomes possible to distribute more points on the sphere which are all spatially separated from all the other points.

The more detailed numerical study of the minimizers showed another interesting effect. For values $\tau < \tau_c := \sqrt{2}$, we found many different minimizers of different form. They all have the property that they are *completely timelike* in the sense that all points in the support of the minimizing measure have timelike or lightlike separation from all the other points. We found minimizers supported on an arbitrarily large number of points. If on the other hand $\tau > \tau_c$, all minimizers were supported on at most $m_0(\tau)$ points, indicating that every minimizing measure $\rho \in \mathfrak{M}$ should be *discrete with finite support*. The intermediate value $\tau = \tau_c$ correspond to the opening angle $\vartheta_{\max} = \frac{\pi}{2}$ of the light cones.

Conjecture 3.2. *If $\tau < \tau_c$, every minimizer is completely timelike. If conversely $\tau > \tau_c$, every minimizing measure is discrete with finite support.*

More graphically, one can say that for $\tau > \tau_c$, our variational principle *spontaneously generates a discrete structure* on the sphere. The two regions $\tau < \tau_c$ and $\tau > \tau_c$ can also be understood as two different phases of the system, so that at $\tau = \tau_c$ we have a *phase transition* from the completely timelike phase to the discrete phase.

The above numerical results will be the guide line for our analysis. More precisely, the completely timelike phase will be analyzed in Section 4.2 using the notion of "generically timelike", whereas in Section 4.3 we will develop under which assumptions and in which sense the support of the minimizing measure is discrete or "singular". The phase transition is made precise in Theorem 4.15 and 4.17 by stating that minimizing measures are either generically timelike or singular.

4. General Structural Results

We now return to the general variational principle (2.5) with the Lagrangian of the form (2.4) and (2.2) on a general smooth compact manifold $\mathcal{F}$. Let us introduce some notation. For a given measure $\rho \in \mathfrak{M}$ we define the functions

$$\ell(x) = \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) \quad \in C^{0,1}(\mathcal{F}) \tag{4.1}$$

$$\mathsf{d}(x) = \int_{\mathcal{F}} \mathcal{D}(x,y)\, d\rho(y) \quad \in C^{\infty}(\mathcal{F})\,. \tag{4.2}$$

Moreover, we denote the Hilbert space $L^2(\mathcal{F}, d\rho)$ by $(\mathcal{H}_\rho, \langle .,. \rangle_\rho)$ and introduce the operators

$$\mathcal{L}_\rho \;:\; \mathcal{H}_\rho \to \mathcal{H}_\rho \;:\; \psi \mapsto (\mathcal{L}_\rho \psi)(x) = \int_{\mathcal{F}} \mathcal{L}(x,y)\, \psi(y)\, d\rho(y) \tag{4.3}$$

$$\mathcal{D}_\rho \;:\; \mathcal{H}_\rho \to \mathcal{H}_\rho \;: \psi \mapsto (\mathcal{D}_\rho \psi)(x) = \int_{\mathcal{F}} \mathcal{D}(x,y)\, \psi(y)\, d\rho(y)\,. \tag{4.4}$$

Lemma 4.1. *The operators $\mathcal{L}_\rho$ and $\mathcal{D}_\rho$ are self-adjoint and Hilbert-Schmidt. The eigenfunctions of $\mathcal{L}_\rho$ (and D_ρ) corresponding to the non-zero eigenvalues can be extended to Lipschitz-continuous (respectively smooth) functions on $\mathcal{F}$.*

Proof. We only consider $\mathcal{D}_\rho$, as the proof for $\mathcal{L}_\rho$ is analogous. The self-adjointness follows immediately from the fact that $\mathcal{D}(x,y)$ is symmetric. Moreover, as the kernel is smooth and $\mathcal{F}$ is compact, we know that

$$\iint_{\mathcal{F}\times\mathcal{F}} |\mathcal{D}(x,y)|^2 d\rho(x)\, d\rho(y) < \infty\,.$$

This implies that $\mathcal{D}_\rho$ is Hilbert-Schmidt (see [12, Theorem 2 in Section 16.1]).

Suppose that $\mathcal{D}_\rho \psi = \lambda \psi$ with $\lambda \neq 0$. Then the representation

$$\psi(x) = \frac{1}{\lambda} \int_{\mathcal{F}} \mathcal{D}(x,y)\, \psi(y)\, d\rho(y)$$

shows that $\psi \in C^\infty(\mathcal{F})$. □

The following notions characterize properties of $\mathcal{F}$ and the function $\mathcal{D}$ which will be needed later on.

Definition 4.2. *A measure $\mu \in \mathfrak{M}$ is a* **homogenenizer** *of $\mathcal{D}$ if $\operatorname{supp} \mu = \mathcal{F}$ and both functions*

$$\ell_\mu(x) := \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\mu(y) \qquad \text{and} \qquad \mathsf{d}_\mu(x) := \int_{\mathcal{F}} \mathcal{D}(x,y)\, d\mu(y)$$

are constant on $\mathcal{F}$. The function $\mathcal{D}$ is called **homogenizable** *if a homogenizer exists.*

In the Examples (a), (b) in Section 2, we can always choose the standard normalized volume measure as the homogenizer. More generally, in Example (c) we choose for μ the normalized Haar measure, obtained by introducing a $\mathrm{U}(f)$-invariant metric on $\mathcal{F}$ and taking the corresponding volume form (see for example [1, Section I.5]).

The next proposition gives a sufficient condition for a homogenizer to be a minimizer.

Proposition 4.3. *If $\mathcal{L}_\mu \geq 0$, the homogenizer μ is a minimizer of the variational principle* (2.5)*.*

Proof. We denote the constant function on $\mathfrak{F}$ by $1_\mathfrak{F} \equiv 1$. If μ is a homogenizer, this function is an eigenfunction of $\mathcal{L}_\mu$, which can be completed to an orthonormal eigenvector basis $(\psi_i)_{i\in\mathbb{N}_0}$ of $\mathcal{H}_\mu$ with $\psi_0 = 1_\mathfrak{F}$ and corresponding eigenvalues $\lambda_i \geq 0$.

Using an approximation argument in the $C^0(\mathcal{F})^*$-topology, it suffices to show that

$$\mathcal{S}[\mu] \leq \mathcal{S}[\psi\mu]$$

for any $\psi \in C^0(\mathfrak{F})$ with $\psi \geq 0$ and $\langle \psi, 1_\mathfrak{F}\rangle_\mu = 1$. We write ψ in the eigenvector basis ψ_i,

$$\psi = \sum_{i=0}^\infty c_i\, \psi_i\,.$$

The condition $\langle \psi, 1_\mathfrak{F}\rangle_\mu = 1$ implies that $c_0 = 1$. Thus

$$\mathcal{S}[\psi\mu] = \langle \psi, \mathcal{L}_\mu \psi\rangle_\mu = \lambda_0 + \sum_{i=1}^\infty |c_i|^2 \lambda_i \geq \lambda_0 = \mathcal{S}[\mu]\,.$$

□

4.1. **The Euler-Lagrange Equations.** Let us assume that ρ is a minimizer of the variational principle (2.5),

$$\mathcal{S}[\rho] = \inf_{\tilde{\rho}\in\mathfrak{M}} \mathcal{S}[\tilde{\rho}] =: \mathcal{S}_{\min}\,.$$

We now derive consequences of the minimality. In the first lemma, we consider first variations of ρ to obtain the Euler-Lagrange equations corresponding to our variational principle. The second lemma, on the other hand, accounts for a nonlinear effect.

Lemma 4.4. *(The Euler-Lagrange equations)*

$$\ell|_{\operatorname{supp}\rho} \equiv \inf_\mathfrak{F} \ell = \mathcal{S}_{\min}\,.$$

Proof. Comparing (2.1) with (4.1), one sees that

$$\mathcal{S}_{\min} = \int_\mathfrak{F} \ell\, d\rho\,. \tag{4.5}$$

Since ℓ is continuous and $\mathfrak{F}$ is compact, there clearly is $y \in \mathfrak{F}$ with

$$\ell(y) = \inf_\mathfrak{F} \ell\,. \tag{4.6}$$

We consider for $t \in [0,1]$ the family of measures

$$\tilde{\rho}_t = (1-t)\,\rho + t\,\delta_y \;\in \mathfrak{M}\,,$$

where δ_y denotes the Dirac measure at y. Substituting this formula in (2.1) and differentiating, we obtain for the first variation the formula

$$\delta\mathcal{S} := \lim_{t\searrow 0} \frac{\mathcal{S}[\tilde{\rho}_t] - \mathcal{S}[\tilde{\rho}_0]}{t} = -2\mathcal{S}_{\min} + 2\ell(y)\,.$$

Since ρ is a minimizer, $\delta\mathcal{S}$ is non-negative. Combining this result with (4.5) and (4.6), we obtain the relations

$$\inf_\mathfrak{F} \ell = \ell(y) \;\geq\; \mathcal{S}_{\min} = \int_\mathfrak{F} \ell\, d\rho\,.$$

It follows that ℓ is constant on the support of ρ, giving the result. □

Lemma 4.5. *The operator $\mathcal{L}_\rho$ is non-negative.*

Proof. Lemma (4.4) yields that for any $x \in \operatorname{supp} \rho$,

$$(\mathcal{L}_\rho 1_{\mathcal{F}})(x) = \int_{\mathcal{F}} \mathcal{L}(x,y)\, d\rho(y) = \ell(x) = \mathcal{S}_{\min} \, 1_{\mathcal{F}}(x) \,,$$

showing that the constant function $1_{\mathcal{F}}$ is an eigenvector corresponding to the eigenvalue $\mathcal{S}_{\min} \geq 0$.

Assume that the lemma is wrong. Then, as $\mathcal{L}_\rho$ is a compact and self-adjoint operator (see Lemma 4.1), there exists an eigenvector ψ corresponding to a negative eigenvalue, $\mathcal{L}_\rho \psi = \lambda \psi$ with $\lambda < 0$. We consider the family of measures

$$\tilde{\rho}_t = (1_{\mathcal{F}} + t\psi)\, \rho \,.$$

In view of Lemma 4.1, ψ is continuous and therefore bounded. Thus for sufficiently small $|t|$, the measure $\tilde{\rho}_t$ is positive. Moreover, the orthogonality of the eigenfunctions $1_{\mathcal{F}}$ and ψ implies that

$$\tilde{\rho}_t(\mathcal{F}) = \int_F 1_{\mathcal{F}}\, (1_{\mathcal{F}} + t\psi)\, d\rho = 1 + t\, \langle 1_{\mathcal{F}}, \psi \rangle_\rho = 1 \,,$$

showing that $\tilde{\rho}_t$ is again normalized. Finally, again using the orthogonality,

$$\mathcal{S}[\tilde{\rho}_t] = \langle (1_{\mathcal{F}} + t\psi), L_\rho (1_{\mathcal{F}} + t\psi) \rangle_\rho = \mathcal{S}_{\min} + \lambda\, t^2\, \langle \psi, \psi \rangle_\rho \,.$$

Thus $\tilde{\rho}_t$ is an admissible variation which decreases the action, a contradiction. □

An immediate consequence of this lemma is a useful positivity property of the Lagrangian when evaluated on a finite number of points in the support of ρ.

Corollary 4.6. *For a finite family $x_0, \ldots, x_N \in \operatorname{supp} \rho$ (with $N \in \mathbb{N}$), the* ***Gram matrix*** *L defined by*

$$L = \Big(\mathcal{L}(x_i, x_j) \Big)_{i,j=0,\ldots,N}$$

is symmetric and positive semi-definite.

Proof. Given $\varepsilon > 0$ and a vector $u = (u_0, \ldots, u_N) \in \mathbb{C}^{N+1}$, we set

$$\psi_\epsilon(x) = \sum_{i=0}^N \frac{u_i}{\rho(B_\varepsilon(x_i))}\, \chi_{B_\varepsilon(x_i)}(x) \in \mathcal{H}_\rho \,,$$

where B_ε is a ball of radius ε (in a given coordinate system). Lemma 4.5 implies that $\langle \psi_\varepsilon, \mathcal{L}_\rho \psi_\varepsilon \rangle \geq 0$. Taking the limit $\varepsilon \searrow 0$, it follows that

$$\langle u, Lu \rangle_{\mathbb{C}^{N+1}} = \lim_{\varepsilon \searrow 0} \langle \psi_\varepsilon, \mathcal{L}_\rho \psi_\varepsilon \rangle_\rho \geq 0 \,.$$

□

4.2. Generically Timelike Minimizers.

Definition 4.7. *A minimizing measure $\rho \in \mathfrak{M}$ is called* **generically timelike** *if the following conditions hold:*

(i) $\mathcal{D}(x,y) \geq 0$ *for all $x, y \in \operatorname{supp} \rho$.*
(ii) *The function d defined by (4.2) is constant on $\mathcal{F}$.*

This constant can easily be computed.

Lemma 4.8. *Suppose that ρ is a generically timelike minimizer. Then*

$$\mathsf{d}(x) = \mathcal{S}_{\min} \qquad \text{for all } x \in \mathcal{F} \,.$$

Proof. Since $\mathcal{L}$ and $\mathcal{D}$ coincide on the support of ρ, we know that

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

Carrying out one integral using (4.2), we obtain

$$\mathcal{S}_{\min} = \int_{\mathcal{F}} \mathsf{d}(x)\, d\rho(x)\,,$$

giving the result. □

In the remainder of this subsection, we assume that $\mathcal{D}$ is homogenizable (see Definition 4.2) and denote the homogenizer by $\mu \in \mathfrak{M}$.

Lemma 4.9. *If $\mathcal{D}_\mu$ has only a finite number of negative eigenvalues, the kernel $\mathcal{D}(x,y)$ has the representation*

$$\mathcal{D}(x,y) = \nu_0 + \sum_{n=1}^{N} \nu_n\, \phi_n(x)\, \overline{\phi_n(y)} \tag{4.7}$$

with $N \in \mathbb{N} \cup \{\infty\}$, $\nu_n \in \mathbb{R}$, $\nu_n \neq 0$, and $\phi_n \in C^\infty(\mathcal{F})$, where in the case $N=\infty$ the series converges uniformly.

Proof. By definition of the homogenizer, the function $1_{\mathcal{F}} \equiv 1$ is an eigenfunction of the operator $\mathcal{D}_\mu$. Denoting the corresponding eigenvalue by ν_0, we obtain the spectral representation (4.7).

If $\mathcal{D}_\mu$ is positive semi-definite, the uniform convergence is an immediate generalization of Mercer's theorem (see [12, Theorem 11 in Chapter 30], where we replace the interval $[0,1]$ by the compact space $\mathcal{F}$, and the Lebesgue measure by the measure μ). In the case when $\mathcal{D}_\mu$ has a finite number of negative eigenvalues, we apply Mercer's theorem similarly to the operator with kernel $\mathcal{D}(x,y) - \sum_{i=1}^K \lambda_i\, \psi_i(x)\overline{\psi_i(y)}$, where $\lambda_1, \ldots \lambda_K$ are the negative eigenvalues with corresponding eigenfunctions ψ_i. By construction, this operator is positive semi-definite, and in view of Lemma 4.1 its kernel is continuous. □

Lemma 4.10. *Suppose that ρ is a generically timelike minimizer and that the operator $\mathcal{D}_\mu$ has only a finite number of negative eigenvalues. Then*

$$\mathcal{S}[\rho] = \nu_0 \qquad \text{and} \qquad \int_{\mathcal{F}} \phi_n(y)\, d\rho(y) = 0 \quad \text{for all } n \in \{1, \ldots, N\}\,.$$

Proof. Using the decomposition of the kernel (4.7) and the uniform convergence, we obtain

$$\mathsf{d}(x) = \nu_0 + \sum_{n=1}^{N} \nu_n\, \phi_n(x) \int_{\mathcal{F}} \overline{\phi_n(y)}\, d\rho(y)\,.$$

Applying Lemma 4.8 gives the claim. □

Proposition 4.11. *Suppose that $\mathcal{D}_\mu$ is a positive semi-definite operator on $\mathcal{H}_\mu$. Then*

$$\mathcal{S}_{\min} \geq \nu_0\,.$$

In the case of equality, every minimizer is generically timelike.

Proof. If $\mathcal{D}_\mu$ is positive semi-definite, all the parameters ν_n in (4.7) are positive. It follows that for every measure $\tilde{\rho} \in \mathfrak{M}$,

$$\mathcal{S}[\tilde{\rho}] = \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d\tilde{\rho}(x)\, d\tilde{\rho}(y) \geq \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{D}(x,y)\, d\tilde{\rho}(x)\, d\tilde{\rho}(y) \geq \nu_0\, \tilde{\rho}(\mathcal{F})^2 = \nu_0\,. \quad (4.8)$$

Let us assume that equality holds. It then follows from (4.8) that $\mathcal{L}$ and $\mathcal{D}$ coincide on the support of $\tilde{\rho}$ and thus $\mathcal{D}(x,y) \geq 0$ for all $x, y \in \operatorname{supp}\tilde{\rho}$. Moreover, we find from (4.7) that

$$\nu_0 = \nu_0 + \sum_{n=1}^N \left| \int_{\mathcal{F}} \overline{\phi_n(y)}\, d\tilde{\rho} \right|^2\,,$$

and thus

$$\int_{\mathcal{F}} \overline{\phi_n(y)}\, d\tilde{\rho} = 0 \qquad \text{for all } n \geq 1\,.$$

It follows that $\mathsf{d}_{\tilde{\rho}}$ is a constant. We conclude that $\tilde{\rho}$ is generically timelike. □

This proposition can be used to construct generically timelike minimizers.

Corollary 4.12. *Suppose that $\mathcal{D}_\mu$ is a positive semi-definite operator on $\mathcal{H}_\mu$. Assume that the function $f \in \mathcal{H}_\mu$ has the following properties:*

(a) $\mathcal{D}(x,y) = \mathcal{L}(x,y)$ *for all* $x, y \in \operatorname{supp} f$.

(b) $\displaystyle\int_{\mathcal{F}} f(x)\, d\mu(x) = 1$ *and* $\displaystyle\int_{\mathcal{F}} f(x)\, \phi_n(x)\, d\mu(x) = 0$ *for all* $n \in \{1, \ldots, N\}$.

Then the measure $d\rho = f\, d\mu$ is a generically timelike minimizer.

Proof. The assumption (a) implies that

$$\mathcal{S}[\rho] = \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{D}(x,y)\, d\rho(x)\, d\rho(y)\,.$$

Using the decomposition (4.7) and the relations (b), we find that $\mathcal{S}[\rho] = \nu_0$. We now apply Proposition 4.11. □

We conclude this section by stating obstructions for the existence of generically timelike minimizers.

Proposition 4.13. *Assume that one of the following conditions hold:*

(I) *The operator $\mathcal{D}_\mu$ has only a finite number of negative eigenvalues, and the eigenvalue ν_0 in the decomposition* (4.7) *is not positive.*

(II) *For every $x \in \mathcal{F}$ there is a point $y \in \mathcal{F}$ with $\mathcal{J}(x) \cap \mathcal{J}(y) = \varnothing$ ("condition of disjoint lightcones").*

(III) *For every $x \in \mathcal{F}$ there is a point $-x \notin \overline{\mathcal{I}(x)}$ with $\mathcal{J}(x) = \overline{\mathcal{I}(x)} \cup \{-x\}$ and $\overline{\mathcal{I}(x)} \cap \overline{\mathcal{I}(-x)} = \varnothing$ ("condition of antipodal points").*

Then there are no generically timelike minimizers.

Proof. We first show that $\mathcal{S}_{\min} > 0$. Namely, choosing x in the support of a minimizing measure ρ, we know from (2.3) and the continuity of $\mathcal{D}$ that there is a neighborhood U of x and $\delta > 0$ such that $D(x,y) > \delta$ for all $y \in U$. It follows that

$$\mathcal{S}_{\min} \geq \int_{U\times U} \mathcal{L}(x,y)\, d\rho(x)\, d\rho(y) \geq \delta\, \rho(U)^2 > 0\,.$$

Case (I) is obvious in view of Lemma 4.10 and the fact that $\mathcal{S}_{\min} > 0$. To prove the remaining cases (II) and (III), we assume conversely that there exists a generically timelike minimizer $\rho \in \mathfrak{M}$. Choosing a point $x \in \operatorname{supp} \rho$, we know from property (i) in Definition 4.7 that $\operatorname{supp} \rho \subset \mathcal{J}(x)$. In case (II), we choose $y \in \mathcal{F}$ with $\mathcal{J}(x) \cap \mathcal{J}(y) = \varnothing$ to obtain

$$\mathsf{d}(y) = \int_{\mathcal{J}(x)} \mathcal{D}(y,z)\, d\rho(z) \leq 0 < \mathcal{S}_{\min} ,$$

in contradiction to Lemma 4.8.

In case (III), we know that $\operatorname{supp} \rho \subset \mathcal{J}(x) = \overline{\mathcal{I}(x)} \cup \{-x\}$. If $-x \notin \operatorname{supp} \rho$, the estimate

$$\mathsf{d}(-x) = \int_{\mathcal{J}(x)} \mathcal{D}(-x,z)\, d\rho(z) = \int_{\overline{\mathcal{I}(x)}} \mathcal{D}(-x,z)\, d\rho(z) \overset{(*)}{\leq} 0 < \mathcal{S}_{\min}$$

again gives a contradiction, where in (*) we used that $\overline{\mathcal{I}(x)} \cap \overline{\mathcal{I}(-x)} = \varnothing$. If conversely $-x \in \operatorname{supp} \rho$, then $\operatorname{supp} \rho \subset \mathcal{J}(x) \cap \mathcal{J}(-x) = \{x\} \cup \{-x\}$ (where we again used that $\overline{\mathcal{I}(x)} \cap \overline{\mathcal{I}(-x)} = \varnothing$). Hence the integral in (4.2) reduces to a sum over two points,

$$\mathsf{d}(y) = \rho(\{x\})\, \mathcal{D}(y,x) + \rho(\{-x\})\, \mathcal{D}(y,-x) . \tag{4.9}$$

In view of our assumption (2.3), we know that $x \in \mathcal{I}(x)$. On the other hand, the relation $\overline{\mathcal{I}(x)} \cap \overline{\mathcal{I}(-x)} = \varnothing$ shows that $-x \notin \mathcal{I}(x)$. Hence there is a point $y \in \partial\mathcal{I}(x)$. It follows that $\mathcal{D}(y,x) = 0$ (because $y \in \partial\mathcal{I}(x)$) and also $\mathcal{D}(y,-x) \leq 0$ (because $y \in \overline{\mathcal{I}(x)}$ and thus $y \notin \overline{\mathcal{I}(-x)}$). Using these inequalities in (4.9), we again find that $\mathsf{d}(y) \leq 0$, a contradiction. ☐

It is an interesting question how the support of a generically timelike minimizer ρ may look like. The next proposition (which will not be used later on) quantifies that $\operatorname{supp} \rho$ must be sufficiently "spread out".

Proposition 4.14. *Assume that ρ is a generically timelike minimizer and that the operator $\mathcal{D}_\mu$ has only a finite number of negative eigenvalues. Then every real function $\psi \in \mathcal{D}_\mu(\mathcal{H}_\mu)$ with*

$$\int_{\mathcal{F}} \psi(x)\, d\mu(x) = 0 \tag{4.10}$$

changes its sign on the support of ρ (here μ is again the homogenizer of Definition 4.2).

Proof. We return to the spectral decomposition (4.7) of the operator $\mathcal{D}_\mu$. Since the eigenfunctions ϕ_n are orthogonal in $\mathcal{H}_\mu$, we know that

$$\int_{\mathcal{F}} \phi_n\, d\mu = 0 \qquad \text{for all } n \geq 1 .$$

Representing ψ in an eigenvector basis of $\mathcal{D}_\mu$ and using (4.10), we find

$$\psi = \sum_{n=1}^N \kappa_n\, \phi_n$$

with complex coefficients κ_n. Integrating with respect to ρ, we can apply Lemma 4.10 to obtain

$$\int_{\mathcal{F}} \psi(x)\, d\rho(x) = \sum_{n=1}^N \kappa_n \int_{\mathcal{F}} \phi_n(x)\, d\rho(x) = 0 .$$

Hence ψ changes its sign on the support of ρ. ☐

4.3. **Minimizers with Singular Support.** We now state results on the support of a minimizing measure.

Theorem 4.15. *Let $\mathcal{F}$ be a smooth compact manifold. Assume that $\mathcal{D}(x,y)$ is symmetric (2.2) and equal to one on the diagonal, $\mathcal{D}(x,x) \equiv 1$. Furthermore, we assume that for every $x \in \mathcal{F}$ and $y \in \mathcal{K}(x)$, there is a smooth curve c joining the points x and y, along which $\mathcal{D}(.,y)$ has a non-zero derivative at x, i.e.*

$$\frac{d}{dt} \mathcal{D}\big(c(t), y\big)\Big|_{t=0} \neq 0 \,, \tag{4.11}$$

where we parametrized the curve such that $c(0) = x$. Then the following statements are true:

(A) *If $\mathcal{F}$, $\mathcal{D}$ are real analytic, then a minimizing measure ρ is either generically timelike or $\overset{\circ}{\operatorname{supp}}\, \rho = \varnothing$.*

(B) *If $\mathcal{D}$ is smooth and if there is a differential operator Δ on $C^\infty(\mathcal{F})$ which vanishes on the constant functions such that*

$$\Delta_x \mathcal{D}(x,y) < 0 \qquad \text{for all } y \in \mathcal{I}(x) \,, \tag{4.12}$$

then $\overset{\circ}{\operatorname{supp}}\, \rho = \varnothing$.

A typical example for Δ is the Laplacian corresponding to a Riemannian metric on $\mathcal{F}$. Note that the condition (4.11) implies that for every $y \in \mathcal{F}$, the set $\{x \,|\, y \in \mathcal{K}(x)\}$ is a smooth hypersurface, which the curve c intersects transversely (in the applications of Section 5 and 6, this set will coincide with $\mathcal{K}(y)$, but this does not need to be true in general).

The condition (4.11) can be removed if instead we make the following symmetry assumption.

Definition 4.16. *The function $\mathcal{D}$ is called **locally translation symmetric** at x with respect to a curve $c(t)$ with $c(0) = x$ if there is $\varepsilon > 0$ and a function $f \in C^\infty((-2\varepsilon, 2\varepsilon))$ such that the curve c is defined on the interval $(-\varepsilon, \varepsilon)$ and*

$$\mathcal{D}(c(t), c(t')) = f(t - t') \qquad \text{for all } t, t' \in (-\varepsilon, \varepsilon) \,.$$

Theorem 4.17. *Let $\mathcal{F}$ be a smooth compact manifold. Assume that $\mathcal{D}(x,y)$ is symmetric (2.2) and strictly positive on the diagonal (2.3). Furthermore, we assume that for every $x \in \mathcal{F}$ and $y \in \mathcal{K}(x)$, there is a smooth curve c joining the points x and y such that $\mathcal{D}$ is locally translation symmetric at x with respect to c, and such that the function $\mathcal{D}(c(t), y)$ changes sign at $t = 0$ (where we again parametrize the curve such that $c(0) = x$). Then statement (A) of Theorem 4.15 holds, provided that the curve c is analytic in a neighborhood of $t = 0$. Assume furthermore that there is $p \in \mathbb{N}$ with*

$$\frac{d^p}{dt^p} \mathcal{D}\big(c(t), y\big)\Big|_{t=0} \neq 0 \,. \tag{4.13}$$

Then statement (B) of Theorem 4.15 again holds.

In the smooth setting, the above theorems involve quite strong additional assumptions (see (4.11), (4.12) and (4.13)). The following counter example shows that some conditions of this type are necessary for the statements of these theorems to be true[2].

[2]We would like to thank Robert Seiringer for pointing out a similar example to us.

Example 4.18. Let $f, g \in C^\infty_0([-\pi, \pi])$ be non-negative even functions with

$$\operatorname{supp} f \subset \left[-\tfrac{\pi}{8}, \tfrac{\pi}{8}\right] , \qquad \operatorname{supp} g \subset \left(-\pi, -\tfrac{\pi}{2}\right] \cup \left[\tfrac{\pi}{2}, \pi\right) .$$

We introduce the function $\mathcal{D} \in C^\infty(S^2 \times S^2)$ by

$$\mathcal{D}(x,y) = -g\big(\operatorname{dist}(x,y)\big) + \int_{S^2} f\big(\operatorname{dist}(x,z)\big)\, f\big(\operatorname{dist}(z,y)\big)\, d\mu(z) , \tag{4.14}$$

where $d\mu$ is the standard volume measure, and dist denotes the geodesic distance (taking values in $[0, \pi]$). Note that the two summands in (4.14) have disjoint supports and thus the corresponding Lagrangian (2.4) simply is

$$\mathcal{L}(x,y) = \int_{S^2} f\big(\operatorname{dist}(x,z)\big)\, f\big(\operatorname{dist}(z,y)\big)\, d\mu(z) , \tag{4.15}$$

We again consider $\mathcal{D}(x, y)$ and $\mathcal{L}(x, y)$ as the integral kernels of corresponding operators $\mathcal{D}_\mu$ and $\mathcal{L}_\mu$ on the Hilbert space $\mathcal{H}_\mu = L^2(S^2, d\mu)$.

First, it is obvious that $\mathcal{D}(x, y)$ is symmetric and constant on the diagonal. Next, it is clear by symmetry that the measure μ is a homogenizer (see Definition 4.2). Moreover, writing $\mathcal{L}_\mu$ as $\mathcal{L}_\mu = f_\mu^2$, where f_μ is the operator with integral kernel f, one sees that the operator $\mathcal{L}_\mu$ is non-negative. Thus by Proposition 4.3, the measure μ is minimizing. If the function g is non-trivial, there are points x, y which are spacelike separated, so that this minimizer is not generically timelike. Also, its support obviously has a non-vanishing interior. We have thus found a minimizing measure which violates statement (A) of Theorem 4.15. ◊

The remainder of this section is devoted to the proof of the above theorems. We begin with a simple but very useful consideration. Suppose that for given $x \in \mathcal{F}$, the boundary of the light cone $\mathcal{K}(x)$ does not intersect the support of ρ. As the support of ρ is compact, there is neighborhood U of x such that

$$\mathcal{K}(z) \cap \operatorname{supp} \rho = \varnothing \qquad \text{for all } z \in U .$$

Thus introducing the measure $\hat{\rho} = \chi_{\mathcal{I}(x)}\, \rho$, the function ℓ can for all $z \in U$ be represented by

$$\ell(z) = \int_{\mathcal{F}} \mathcal{L}(z, \xi)\, d\hat{\rho}(\xi) = \int_{\mathcal{F}} \mathcal{D}(z, \xi)\, d\hat{\rho}(\xi) . \tag{4.16}$$

This identity can be used both in the smooth and in the analytic case.

Lemma 4.19. *If* (4.12) *holds, then for every* $x \in \operatorname{supp} \rho$ *the set* $\mathcal{K}(x) \cap \operatorname{supp} \rho$ *is nonempty.*

Proof. Applying the Laplacian to (4.16) gives

$$\Delta_x \ell(x) = \int_{\mathcal{F}} \Delta_x \mathcal{D}(x, z)\, d\hat{\rho}(z) < 0 ,$$

where in the last step we used (4.12) and the fact that $x \in \operatorname{supp} \tilde{\rho}$. This is a contradiction to Lemma 4.4. □

Lemma 4.20. *Suppose that* $\mathcal{F}$ *and* $\mathcal{D}$ *are real analytic. Assume that there exists a point* $x \in \overset{\circ}{\operatorname{supp}}\, \rho$ *such that* $\mathcal{K}(x) \cap \operatorname{supp} \rho = \varnothing$*. Then* ρ *is generically timelike and* $\operatorname{supp} \rho \subset \mathcal{I}(x)$*.*

Proof. We introduce on $\mathcal{F}$ the function
$$\hat{\mathsf{d}}(y) = \int_{\mathcal{F}} \mathcal{D}(y,z)\, d\hat{\rho}(y)\,.$$
Then $\hat{\mathsf{d}}$ is real analytic and, according to (4.16), it coincides on U with the function ℓ. Since $x \in \overset{\circ}{\text{supp}}\,\rho$, the Euler-Lagrange equations in Lemma (4.4) yield that $\ell \equiv \mathcal{S}_{\min}$ in a neighborhood of x. Hence $\hat{\mathsf{d}} \equiv \mathcal{S}_{\min}$ in a neighborhood of x, and the real analyticity implies that
$$\hat{\mathsf{d}} \equiv \mathcal{S}_{\min} \qquad \text{on } \mathcal{F}\,.$$
It follows that
$$\begin{aligned} \mathcal{S}_{\min} &= \int_{\mathcal{F}} \hat{\mathsf{d}}(x) d\rho(x) = \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{D}(x,y)\, d\hat{\rho}(x)\, d\rho(y) \\ &\leq \iint_{\mathcal{F}\times\mathcal{F}} \mathcal{L}(x,y)\, d\hat{\rho}(x)\, d\rho(y) = \int_{\mathcal{F}} \ell(x)\, d\hat{\rho}(x) = \mathcal{S}_{\min}\, \hat{\rho}(\mathcal{F})\,, \end{aligned} \tag{4.17}$$
and thus $\hat{\rho}(\mathcal{F}) = 1$. Since $\hat{\rho} \leq \rho$ and ρ is normalized, we conclude that $\rho = \hat{\rho}$. Thus $\mathsf{d} \equiv \hat{\mathsf{d}} \equiv \mathcal{S}_{\min}$. Moreover, the inequality in (4.17) becomes an equality, showing that $\mathcal{L} \equiv \mathcal{D}$ on the support of ρ. Thus ρ is indeed generically timelike. □

To complete the proof of Theorems 4.15 and 4.17, it remains to show the following statement:
$$\mathcal{K}(x) \cap \text{supp}\,\rho = \varnothing \qquad \text{for all } x \in \overset{\circ}{\text{supp}}\,\rho\,. \tag{4.18}$$
We proceed indirectly and assume that there is a point $y \in \mathcal{K}(x) \cap \text{supp}\,\rho$. Our strategy is to choose points $x_0, \ldots, x_k$ in a neighborhood of x such that $\mathcal{L}$ restricted to the set $\{x_0, \ldots, x_k, y\}$ is not positive semi-definite, in contradiction to Corollary 4.6. The points $x_0, \ldots, x_k$ will all lie on a fixed smooth curve c which joins x and y chosen as in the statement of the theorems. We parametrize c such that $c(0) = x$ and $c(1) = y$, and by extending the curve we can arrange (possibly by decreasing ε) that the curve is defined on the interval $(-k\varepsilon, 1]$. By the assumptions in Theorems 4.15 and 4.17, we know that $\mathcal{D}(c(t), y)$ changes sign at $t = 0$. Depending on the sign of $\mathcal{D}(c(\varepsilon), 0)$, we introduce the equidistant "chain" of points
$$\begin{cases} x_0 = c(\varepsilon),\;\; x_1 = c(0),\;\; x_2 = c(-\varepsilon),\; \ldots\;, x_k = c(-(k-1)\varepsilon) & \text{if } \mathcal{D}(c(\varepsilon), 0) > 0 \\ x_0 = c(-\varepsilon),\;\; x_1 = c(0),\;\; x_2 = c(\varepsilon),\; \ldots\;, x_k = c((k-1)\varepsilon) & \text{if } \mathcal{D}(c(\varepsilon), 0) < 0. \end{cases} \tag{4.19}$$
(thus y has timelike separation from x_0, lightlike separation from $x_1 = x$, and spacelike separation from $x_2, \ldots, x_k$). Then by construction, $x_0 \in \mathcal{I}(y)$, whereas all the other points of the chain are spacelike or lightlike separated from y.

For the proof of Theorem 4.15, it suffices to consider a chain of three points.

Lemma 4.21. *Assume that $\mathcal{D}(x,y)$ is symmetric* (2.2) *and equal to one on the diagonal, $\mathcal{D}(x,x) \equiv 1$. Then for x_0, x_1, x_2 as given by* (4.19) *in the case $k = 2$, there is a real constant a_1 such that for all sufficiently small ε,*
$$\mathcal{D}(x_i, x_j) = 1 + a_1\, |i-j|^2\, \varepsilon^2 + \mathcal{O}(\varepsilon^3) \qquad \textit{for all } i,j \in \{0,1,2\}\,. \tag{4.20}$$

Proof. We set $f(t,t') = \mathcal{D}(c(t), c(t'))$ for $t, t' \in (-2\varepsilon, 2\varepsilon)$. Using that $\mathcal{D}$ is symmetric and that $\mathcal{D}(x,x) \equiv 1$, we know hat
$$0 = \frac{d}{dt} f(t_0, t_0) = 2\, \frac{d}{dt} f(t_0, t)\big|_{t=t_0}\,.$$

Thus the linear term in a Taylor expansion vanishes,

$$f(t_0, t) = 1 + \frac{1}{2}\, g(t_0)\, (t-t_0)^2 + \mathcal{O}\big(|t-t_0|^3\big)\;,$$

where we set

$$g(t_0) = \frac{d^2}{dt^2} f(t_0, t)\Big|_{t=t_0}\;.$$

As the function g is smooth, we can again expand it in a Taylor series,

$$g(t_0) = g(0) + \mathcal{O}(t_0)\;.$$

We thus obtain

$$f(t_0, t) = 1 + \frac{1}{2}\, g(0)\, (t-t_0)^2 + \mathcal{O}\big(|t_0|\, |t-t_0|^2\big) + \mathcal{O}\big(|t-t_0|^3\big)\;.$$

Setting $a_1 = 2g(0)$ and using that $|t|, |t_0| \leq 2\varepsilon$, the result follows. □

Lemma 4.22. *Under the assumptions of Theorem 4.15, the statement* (4.18) *holds.*

Proof. Assume conversely that for $x \in \overset{\circ}{\operatorname{supp}}\, \rho$ there is a point $y \in \operatorname{supp} \rho \cap \mathcal{K}(x)$. We choose the chain $x_0, x_1 = x, x_2$ as in Lemma 4.21. We use the notation of Corollary 4.6 in case $N = 3$, setting $x_3 = y$. Choosing the vector $u \in \mathbb{C}^4$ as $u = (1, -2, 1, 0)$, we can apply Lemma 4.21 to obtain

$$\langle u, Lu \rangle_{\mathbb{C}^4} = 6 - 4\, \mathcal{D}(x_0, x_1) + 2\, \mathcal{D}(x_0, x_2) - 4\, \mathcal{D}(x_1, x_2) = \mathcal{O}(\varepsilon^3)\;.$$

Furthermore, using (4.11), we know that

$$\mathcal{D}(x_0, y) = b\, \varepsilon + \mathcal{O}(\varepsilon^2)$$

with $b \neq 0$. Thus, choosing $u = (\alpha, -2\alpha, \alpha, \beta)$ with $\alpha, \beta \in \mathbb{R}$, it is

$$\langle u, \mathcal{L}u \rangle_{\mathbb{C}^4} = \left\langle \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \mathcal{O}(\varepsilon^3) & b\varepsilon + \mathcal{O}(\varepsilon^2) \\ b\varepsilon + \mathcal{O}(\varepsilon^2) & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right\rangle_{\mathbb{C}^2}\;.$$

For sufficiently small ε, the matrix in this equation has a negative determinant, in contradiction to Corollary 4.6. □

This completes the proof of Theorem 4.15.

In order to finish the proof of Theorem 4.17, we first remark that combining the symmetry of $\mathcal{D}$ with the assumption that $\mathcal{D}$ is locally translation symmetric at x with respect to c, we know that $\mathcal{D}(c(t), c(t')) = f(|t-t'|)$. A Taylor expansion of f yields the following simplification and generalization of Lemma 4.21,

$$\mathcal{D}(c(t), c(t')) = 1 + \sum_{i=1}^{K} a_i\, (t-t')^{2i} + \mathcal{O}\Big((t-t')^{2(K+1)} \Big), \tag{4.21}$$

where the real coefficients a_i only depend on c.

Lemma 4.23. *Under the assumptions of Theorem 4.17, the statement* (4.18) *holds.*

Proof. Let us first verify that in the real analytic case, there is a p such that (4.13) holds. Namely, assuming the contrary, all the t-derivatives of the function $\mathcal{D}(c(t), y)$ vanish. As the function $\mathcal{D}(c(t), y)$ is real analytic in a neighborhood of $t = 0$ (as the composition of analytic functions is analytic), it follows that this function is locally constant. This contradicts the fact that $\mathcal{D}(c(t), y)$ changes sign at $t = 0$.

Assume conversely that for $x \in \overset{\circ}{\text{supp}}\, \rho$ there is a point $y \in \text{supp}\, \rho \cap \mathcal{K}(x)$. We choose the chain $x_0, x_1 = x, x_2, \ldots, x_k$ as in (4.19) with $k = p + 1$. We use the notation of Corollary 4.6 in case $N = k$. Then the Gram matrix L becomes

$$L = \big(f(\varepsilon |i-j|) \big)_{i,j=0,\ldots,k} = \begin{pmatrix} 1 & f(\varepsilon) & \cdots & f(k\varepsilon) \\ f(\varepsilon) & 1 & & \\ \vdots & & \ddots & \\ f(k\varepsilon) & \cdots & & 1 \end{pmatrix} .$$

Using the expansion (4.21) for $K = k - 1$, we obtain

$$\begin{aligned} L = E &+ a_1 \varepsilon^2 \big(|i-j|^2 \big) + a_2 \varepsilon^4 \big(|i-j|^4 \big) \\ &+ \ldots + a_{k-1} \varepsilon^{2(k-1)} \big(|i-j|^{2(k-1)} \big) + \mathcal{O}\big(\varepsilon^{2k}\big) \,, \end{aligned} \tag{4.22}$$

where E denotes the matrix where all the matrix entries (also the off-diagonal entries) are equal to one, and $(|i-j|^q)$ is the matrix whose element (i,j) has the value $|i-j|^q$.

Let us construct a vector $v \in \mathbb{C}^{k+1}$ such that the expectation value $\langle v, Lv \rangle$ is $\mathcal{O}(\varepsilon^{2k})$. To this end, we take for $v = (v_i)_{i=0}^k \in \mathbb{C}^{k+1}$ a non-trivial solution of the k linear equations

$$\sum_{i=0}^k v_i = 0, \quad \sum_{i=0}^k i v_i = 0, \quad \sum_{i=0}^k i^2 v_i = 0, \quad \ldots, \sum_{i=0}^k i^{k-1} v_i = 0 \,. \tag{4.23}$$

Then $\langle v, Ev \rangle = 0$ and for all $l \in \{1, \ldots, k-1\}$

$$\begin{aligned} \langle v, (|i-j|^{2l}) v \rangle &= \sum_{i,j=0}^k v_i v_j |i-j|^{2l} = \sum_{i,j=0}^k v_i v_j \sum_{\nu=0}^{2l} \binom{2l}{\nu} i^\nu j^{2l-\nu} = \\ &= \sum_{i,j=1}^k v_i v_j \left(i^{2l} + 2l\, i^{2l-1} j + \ldots + \binom{2l}{l} i^l j^l + \ldots + j^{2l} \right) . \end{aligned}$$

Each summand involves a power of i and a power of j, where always one of these powers is smaller than k. Thus all summands vanish according to (4.23). The solution v can always be normalized by $v_0 = 1$, because setting v_0 to zero, the system of equations (4.23) can be rewritten with the square Vandermonde matrix which has a trivial kernel. In view of the expansion (4.22), we conclude that $\langle v, Lv \rangle = \mathcal{O}(\varepsilon^{2k})$.

We next consider the setting of Corollary 4.6 in case $N = k + 1$ and $x_{k+1} = y$. Using (4.13) together with the fact that the points y and x_0 are timelike separated, we find that

$$\mathcal{L}(x_0, y) = b\, \varepsilon^p + \mathcal{O}(\varepsilon^{p+1}) \tag{4.24}$$

for $b \neq 0$. We choose the vector $u \in \mathbb{C}^{k+2}$ as $u = (\alpha v_0, \ldots, \alpha v_k, \beta)$ with $\alpha, \beta \in \mathbb{R}$,

$$\langle u, \mathcal{L} u \rangle_{\mathbb{C}^4} = \left\langle \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \mathcal{O}(\varepsilon^{2k}) & b\varepsilon^p + \mathcal{O}(\varepsilon^{p+1}) \\ b\varepsilon^p + \mathcal{O}(\varepsilon^{p+1}) & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right\rangle_{\mathbb{C}^2} ,$$

where we combined (4.24) with our normalization $v_0 = 1$, and used that y is not timelike separated from $x_1, \ldots, x_k$. For sufficiently small ε, the matrix in this equation has a negative determinant, in contradiction to Corollary 4.6. ☐

This completes the proof of Theorem 4.17.

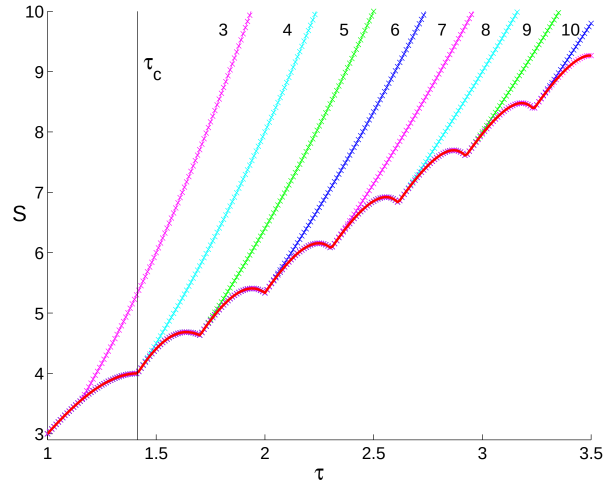

FIGURE 4. Numerical minima for the weighted counting measure on the circle.

5. The Variational Principles on the Circle

As a simple starting point for a more detailed analysis, we now consider the variational principles on the circle (see Example (b) on page 5). We first discuss numerical results, which again show the "critical behavior" discussed in Section 3 for the variational principle on S^2. Applying the previous structural results, we will prove this critical behavior and show under generic assumptions that the minimizing measure is supported at a finite number of points. Moreover, we will give many minimizers in closed form.

The numerical solution methods and results are similar as on S^2, as we now describe. We again consider the weighted counting measure (3.3). As the starting configuration we choose in analogy of the Tammes distribution on S^2 a uniform distribution of m points on the circle,

$$X_m = \{x_k = e^{i(k-1)\vartheta_m}, k = 1, \ldots, m\} , \qquad \vartheta_m = \frac{2\pi}{m} , \tag{5.1}$$

with uniform weights $\rho_k = 1/m$. Minimizing as in Section 3 with a simulated annealing algorithm, we obtain the result shown in Figure 4. The numerical results indicate that the minimizing measure is supported at a finite number of points m_0. This number can be stated explicitly by

$$m_0 = \min \left\{ n \in \mathbb{N} \; : \; n \geq \frac{2\pi}{\vartheta_{\max}} \right\} , \tag{5.2}$$

where $\vartheta_{\max}$, as given by (2.7), denotes the opening angle of the lightcone. The number m_0 increases with τ, with discontinuous "jumps" at the values

$$\tau_m := \sqrt{\frac{2}{1 - \cos(\vartheta_m)}} . \tag{5.3}$$

Besides the discrete nature of the minimizers, the numerical results reveal that at $\tau = \tau_c = \sqrt{2}$ (corresponding to $\vartheta_{\max} = \frac{\pi}{2}$), the structure of the minimizers changes completely. Just as in Section 3, this effect can be understood as a phase transition. More precisely, if $\tau \leq \tau_c$, every minimizer is generically timelike. If we further decrease τ

(i.e.. for every fixed $\tau < \tau_3$), we even found a large number of minimizing measures, supported at different numbers of points with strikingly different positions. However, if $\tau > \sqrt{2}$, the minimizer is unique (up to rotations on S^1), is supported at m_0 points, and is not generically timelike.

In the remainder of this section, we make this picture rigorous. First, the operator D_μ can be diagonalized explicitly by plane waves $\phi_n(x) = e^{in\vartheta_x}$ (where $n \in \mathbb{Z}$, and ϑ_x is the angle). This gives rise to the decomposition

$$\mathcal{D}(x,y) = \nu_0 + \sum_{n=1}^{2} \nu_n \left(e^{in(\vartheta_x - \vartheta_y)} + e^{-in(\vartheta_x - \vartheta_y)} \right),$$

where

$$\nu_0 = \iint_{S^1 \times S^1} \mathcal{D}(x,y)\, d\mu(x)\, d\mu(y) = 4\tau^2 - \tau^4 . \tag{5.4}$$

and similarly $\nu_1 = 2\tau^2$ and $\nu_2 = \frac{1}{2}\tau^4$. In the case $\tau \leq 2$ all eigenvalues ν_0, ν_1 and ν_2 are non-negative, and we can apply Proposition 4.11 to obtain

$$\mathcal{S}_{\min} \geq \nu_0 .$$

For sufficiently small τ, the uniform distribution of points on the circle (5.1) gives a family of generically timelike minimizers.

Lemma 5.1. *If $m \geq 3$ and τ is so small that $\mathcal{L}(x,y) = \mathcal{D}(x,y)$ for all $x, y \in X_m$, then $\rho = \frac{1}{m} \sum_{i=1}^m \delta_{x_i}$ is a generically timelike minimizer. Every other minimizer is also generically timelike.*

Proof. A straightforward calculation using the identities

$$\sum_{k=0}^{m-1} e^{ik\vartheta_m} = 0 \qquad \text{and} \qquad \sum_{k=0}^{m-1} \left(e^{ik\vartheta_m}\right)^2 = 0$$

yields for any $x \in S^1$,

$$\begin{aligned} \mathsf{d}(x) &= \frac{1}{m} 2\tau^2 \sum_{k=0}^{m-1} \left(2 + 2\langle x, x_k\rangle - \tau^2 + \tau^2 \langle x, x_k\rangle^2\right) \\ &= \frac{1}{m} 2\tau^2 \left(2m - m\tau^2 + \frac{m}{2}\,\tau^2\right) = \nu_0 . \end{aligned}$$

In particular, one sees that $\mathcal{S}[\rho] = \nu_0$.

The assumption $\mathcal{L}(x,y) = \mathcal{D}(x,y)$ for all $x, y \in X_m$ can only be satisfied if $\tau < 2$. Thus in view of (5.4), the operator D_μ is positive semi-definite. We finally apply Proposition 4.11. ∎

Applying this lemma in the case $m = 4$ gives the following result.

Corollary 5.2. *If $\tau \leq \tau_c$, every minimizer is generically timelike.*

More general classes of generically timelike minimizers can be constructed explicitly with the help of Corollary 4.12. In particular, one can find minimizing measures which are not discrete. For the details we refer to the analogous measure on S^2 given in Example 6.2.

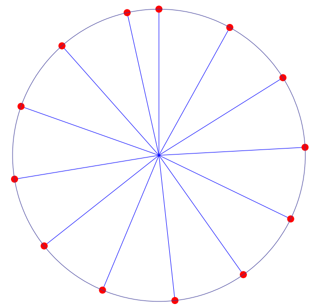

FIGURE 5. A minimizer for $\tau = 4$.

Having explored the case $\tau \leq \tau_c$, we proceed with the case $\tau > \tau_c$. As already stated, the closed lightcones are given by

$$\mathcal{J}(x) = \Big\{ y \;:\; \langle x, y \rangle \geq 1 - \frac{2}{\tau^2} = \cos(\vartheta_{\max}) \Big\} \cup \{-x\} \,.$$

Therefore if $\tau > \sqrt{2} = \tau_c$ (or equivalently $\vartheta_{\max} < \frac{\pi}{2}$), the condition of antipodal points (see Proposition 4.13) is satisfied. Thus there are no generically timelike minimizers. As the condition (4.11) is obvious, we can apply Theorem 4.15 (A) and conclude that

$$\text{if } \tau > \tau_c, \text{ every minimizing measure is discrete} \,. \tag{5.5}$$

Using results and methods from Section 4.3, we we will be able to explicitly construct all minimizers under the additional technical assumption that

$$\tau > \tau_d := \sqrt{3 + \sqrt{10}} \,.$$

We first introduce a descriptive notation:

Definition 5.3. *A* ***chain*** *of length k is a sequence $x_1, \ldots, x_k \in S^1$ of pairwise distinct points such that $\langle x_i, x_{i+1} \rangle = \cos(\vartheta_{\max})$ for all $i = 1, \ldots, k-1$.*

Theorem 5.4. *If $\tau > \tau_d$, the support of every minimizer ρ is a chain $\{x_1, \ldots, x_{m_0}\}$ (with m_0 as given by (5.2)). The minimal action is*

$$\mathcal{S}_{\min} = \frac{\mathcal{L}(0)(\mathcal{L}(0) + \mathcal{L}(\gamma))}{(m_0 - 2)(\mathcal{L}(0) + \mathcal{L}(\gamma)) + 2\mathcal{L}(0)} \,, \tag{5.6}$$

where $\gamma = \arccos(\langle x_1, x_{m_0} \rangle) \in (0, \vartheta_{\max}]$. The minimizing measure is unique up to rotations on S^1.

An example for the support of the minimizing measure is shown in Figure 5. Up to rotations, the points of the chain can be written as

$$x_k = e^{i(k-1)\vartheta_{\max}}, \quad k = 1, \ldots, m_0. \tag{5.7}$$

In the special cases $\tau = \tau_m$, the minimizer is the measure with equal weights supported on the uniform distribution X_m. In the general case, the weights will not all be the same, as will be specified below.

For the proof of Theorem 5.4 we proceed in several steps.

Lemma 5.5. *If $\tau > \sqrt{6}$, the minimal action is attained for a measure supported on a chain $x_1, \ldots, x_k$. In the case $k = m_0$, every minimizing measure is a chain.*

Proof. Let ρ be a minimizing measure. We first note that every chain K in the support of ρ must have finite length, because otherwise $\vartheta_{\max}/\pi$ would have to be irrational. As a consequence, K would be a dense set of S^1, in contradiction to the discreteness of ρ (see (5.5)). Let us assume that the support of ρ is not a chain.

We let $K \subset \operatorname{supp} \rho$ be a chain, which is maximal in the sense that it cannot be extended. Set $L = \operatorname{supp} \rho \setminus K$. We consider variations of ρ where we rotate K by a small angle ϑ, leaving the weights on K as well as $\rho|_L$ unchanged. The fact that K cannot be extended implies that that these variations are smooth in ϑ at $\vartheta = 0$. The minimality of ρ implies that

$$\delta \mathcal{S} = 0 \qquad \text{and} \qquad \delta^2 \mathcal{S} = \sum_{x \in K, y \in L} 2\, \rho(x)\, \rho(y)\, \delta^2 \mathcal{L}(x,y) \geq 0\,. \tag{5.8}$$

On the other hand, differentiating (2.6), one finds that the function $\mathcal{D}$ restricted to $[0, \vartheta_{\max}]$ is concave,

$$\mathcal{D}''(\vartheta) = -4\tau^2 (\cos(\vartheta) + \tau^2 \cos(2\vartheta)) < 0 \qquad (\text{if } \tau > \sqrt{6})\,. \tag{5.9}$$

Comparing with (5.8), we conclude that $\mathcal{L}(x,y)$ vanishes for all $x \in K$ and $y \in L$. In the case that $\#K = m_0$, this implies that $L = \varnothing$, a contradiction. In the remaining case $\#K < m_0$, we can subdivide the circle into two disjoint arcs A_K and A_L such that $K \subset A_K$ and $L \subset A_L$. The opening angle of A_K can be chosen larger than $\vartheta_{\max}$ times the length of K, giving an a-priori upper bound on the length of K.

By further rotating K, we can arrange that the chain K can be extended by a point in L, without changing the action. If the extended chain equals the support of ρ, the proof is finished. Otherwise, we repeat the above argument with K replaced by its extension. In view of our a-priori bound on the length of K, this process ends after a finite number of steps. □

Lemma 5.6. *Suppose that ρ is a minimizing measure supported on a chain. If $\tau > \sqrt{3 + \sqrt{10}}$, the length of this chain is at most m_0.*

Proof. For all $\gamma \in (0, \vartheta_{\max})$ an elementary calculation shows that

$$\mathcal{L}(\gamma)^2 + \mathcal{L}(\vartheta_{\max} - \gamma)^2 > \mathcal{L}(0)^2. \tag{5.10}$$

In the case $\tau = \tau_{m_0}$ there is nothing to prove. Thus we can assume that $\tau \neq \tau_{m_0}$. For a chain $x_1, \ldots, x_k$ with $k > m_0$, the Gram matrix corresponding to the points x_1, x_{m_0+1}, x_2 has the form

$$\begin{pmatrix} \mathcal{L}(0) & \mathcal{L}(\vartheta_{\max} - \gamma) & 0 \\ \mathcal{L}(\vartheta_{\max} - \gamma) & \mathcal{L}(0) & \mathcal{L}(\gamma) \\ 0 & \mathcal{L}(\gamma) & \mathcal{L}(0) \end{pmatrix}. \tag{5.11}$$

Using (5.10), its determinant is negative, in contradiction to Corollary 4.6. □

From the last two lemmas we conclude that every minimizer ρ is supported on one chain of length at most m_0. Parametrizing the points as in (5.7), the only contributions to the action come from $\mathcal{L}(x_l, x_l)$ and $\mathcal{L}(x_1, x_{m_0})$. Using Lagrange multipliers, the optimal weights $\rho_i = \rho(x_i)$ are calculated to be

$$\rho_1 = \rho_{m_0} = \frac{\lambda}{\mathcal{L}(0) + \mathcal{L}(\gamma)} \qquad \text{and} \qquad \rho_i = \frac{\lambda}{\mathcal{L}(0)} \quad \text{for } i = 2, \ldots, m_0 - 1\,, \tag{5.12}$$

where we set

$$\lambda = \frac{\mathcal{L}(0)\,(\mathcal{L}(0)+\mathcal{L}(\gamma))}{(m_0-2)(\mathcal{L}(0)+\mathcal{L}(\gamma))+2\mathcal{L}(0)}\,.$$

The corresponding action is computed to be $\mathcal{S}[\rho] = \lambda$, giving the formula in (5.6). Using this explicit value of the action, we obtain the following

Lemma 5.7. *Suppose that ρ is a minimizing measure supported on a chain. Then the length of this chain is at least m_0.*

Proof. For a chain of length $n < m_0$, the only contribution to the action come from $\mathcal{L}(x_l, x_l)$, $l = 1, \ldots, n$. The corresponding optimal weights are computed by $\rho_i = 1/n$. The resulting action is

$$\mathcal{S} = \sum_{i=1}^{n} \frac{1}{n^2}\mathcal{L}(x_i, x_i) = \frac{1}{n}\mathcal{L}(0)\,.$$

This is easily verified to be strictly larger than the value of the action in (5.6). □

This completes the proof of Theorem 5.4.

We finally remark that if τ lies in the interval $(\sqrt{2}, \sqrt{3+\sqrt{10}})$ where Theorem 5.4 does not apply, the numerics show that the minimizing ρ is again the measure supported on the chain of length m_0, with one exception: If τ is in the interval $(1.61988, \tau_5)$ with $\tau_5 = \sqrt{2+\frac{2}{\sqrt{5}}}$, a chain of length $m_0+1 = 6$ gives a lower action than the chain of length 5. In this case, the Gram matrix (5.11) is indeed positive definite, so that the argument in Lemma 5.7 fails.

6. The Variational Principles on the Sphere

We now come to the analysis of the variational principles on the sphere (see Example (a) on page 4). Applying Theorem 4.15 (A) with the curve c chosen as the grand circle joining x and y, we immediately obtain that every minimizing measure ρ on S^2 is either generically timelike or $\overset{\circ}{\text{supp}}\,\rho = \varnothing$. The numerics in Section 3 indicated that these two cases are separated by a "phase transition" at $\tau = \tau_c = \sqrt{2}$. We will now prove that this phase transition really occurs. Moreover, we will develop methods for estimating the minimal action from above and below. Many of these methods apply just as well to the general setting introduced in Section 2 (see (2.1)–(2.5)).

6.1. Generically Timelike Minimizers.

We first decompose $\mathcal{D}$ in spherical harmonics. A short calculation yields in analogy to (4.7) the decomposition

$$\mathcal{D}(x,y) = \nu_0 + 4\pi \sum_{l=1}^{2} \nu_l \sum_{m=-l}^{l} Y_l^m(x)\overline{Y_l^m(y)}\,,$$

where the eigenvalues are given by

$$\nu_0 = 4\,\tau^2 - \frac{4}{3}\,\tau^4\,, \qquad \nu_1 = \frac{4}{3}\,\tau^2\,, \qquad \nu_2 = \frac{4}{15}\,\tau^4\,. \tag{6.1}$$

In particular, the operator $\mathcal{D}_\mu$ is positive semi-definite if $\tau \leq \sqrt{3}$.

If $\tau \leq \tau_c$, there is a large family of minimizers, as we now discuss. The simplest example is the *octahedron*: Denoting the unit vectors in $\mathbb{R}^3$ by e_1, e_2, e_3, we consider the measure ρ supported at $\pm e_i$ with equal weights $\frac{1}{6}$. Obviously, the condition (i) in Definition 4.7 is satisfied. Moreover, for any $x \in S^2$ one calculates

$$\mathsf{d}(x) = \frac{1}{6} \sum_{y \in \operatorname{supp} \rho} 2\tau^2 \left(2 + 2\langle x, y\rangle - \tau^2 + \tau^2 \langle x, y\rangle^2\right) =$$
$$= \frac{1}{3} \tau^2 \left(12 - 6\tau^2 + 2\tau^2 (x_1^2 + x_2^2 + x_3^2)\right) = \nu_0 \,.$$

Thus Proposition 4.11 yields that ρ is a generically timelike minimizer. Moreover, from Proposition 4.11 we conclude that every minimizer is generically timelike. If conversely $\tau > \tau_c$, the condition of antipodal points is fulfilled, and thus Proposition 4.13 shows that no generically timelike minimizers exist. We have thus proved the following result.

Corollary 6.1. *If $\tau \leq \tau_c$, every minimizing measure ρ on S^2 is generically timelike, and the minimal action is equal to ν_0 as given by* (6.1)*. If conversely $\tau > \tau_c$, every minimizing measure ρ is not generically timelike and* $\overset{\circ}{\operatorname{supp}}\, \rho = \varnothing$.

Using Corollary 4.12, one can also construct minimizers which are not discrete, as is illustrated by the following example.

Example 6.2. We introduce the function $f \in L^2(S^2)$ by

$$f(\vartheta, \varphi) = \begin{cases} \frac{5}{3} & \text{if } \vartheta \in [0, \arccos(0.8)], \\ \frac{35}{9} & \text{if } \vartheta \in [\arccos(0.4), \arccos(0.2)] \\ \frac{40}{9} & \text{if } \vartheta \in [\arccos(-0.5), \arccos(-0.7)], \\ 0 & \text{otherwise.} \end{cases}$$

Then if $\tau < 1.00157$, a straightforward calculation shows that f has the properties (a) and (b) of Corollary 4.12. Thus the measure $d\rho = f d\mu$ is a minimizing generically timelike measure with $\overset{\circ}{\operatorname{supp}}\, \rho \neq \varnothing$. ♢

6.2. **Estimates of the Action.** As not even the solution of the Tammes problem is explicitly known, we cannot expect to find explicit minimizers for general τ. Therefore, we need good estimates of the action from above and below. We now explain different methods for getting estimates, which are all compiled in Figure 6.

Estimates from above can be obtained simply by computing the action for suitable test measures. For example, the action of the normalized volume measure is

$$\mathcal{S}[\mu] = \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^{\vartheta_{\max}} d\vartheta \, \sin\vartheta \, \mathcal{D}(\vartheta) = 4 - \frac{4}{3\tau^2} \geq \mathcal{S}_{\min} \,.$$

As one sees in Figure 6, this estimate is good if τ is close to one. Another example is to take the measure supported at the Tammes distribution for K points, with equal weights. We denote the corresponding action by $\mathcal{S}_T^K$. We then obtain the estimate

$$\mathcal{S}_{\min} \leq \mathcal{S}_T := \min_K \mathcal{S}_T^K \,.$$

One method is to compute $\mathcal{S}_T$ numerically using the tables in [14]. This gives quite good results (see Figure 6), with the obvious disadvantage that the estimate is not given in closed form. Moreover, the Tammes distribution is useful for analyzing the

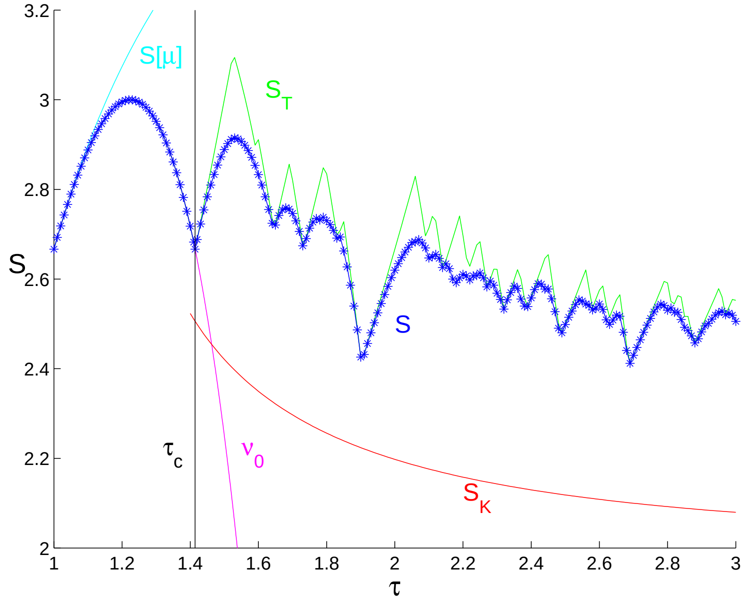

FIGURE 6. Estimates of the action on S^2: Upper bounds obtained from the volume measure $S[\mu]$ and from the Tammes distribution S_T, lower bounds by ν_0 and by the heat kernel estimate S_K.

asymptotics for large τ. To this end, for every Tammes-distribution X_K we introduce τ_K as the minimal value of τ for which all distinct points in X_K are spacelike separated. In analogy to (5.3), the value of τ_K is given by

$$\tau_K = \sqrt{\frac{2}{1-\cos(\vartheta_K)}}\,,$$

where ϑ_K now denotes the minimal angle between the points of the Tammes distribution,

$$\vartheta_K = \max_{x_1,\ldots,x_K \in S^2} \min_{i \neq j} \arccos(\langle x_i, x_j \rangle)\,.$$

Using an estimate by W. Habicht and B.L. van der Waerden for the solution ϑ_K (see [13, page 6]), we obtain

$$4\left(\left(\frac{8\pi}{\sqrt{3}K}\right)^{1/2} - \frac{C}{K^{2/3}}\right)^{-2} \;\geq\; \tau_K^2 \geq 4\,\frac{\sqrt{3}K}{8\pi} \tag{6.2}$$

for some constant $C>0$. For given $\tau > 1$ we choose $K \in \mathbb{N}$ such that $\tau_{K-1} \leq \tau < \tau_K$. Then

$$\mathcal{S}_{\min} \leq \mathcal{S}_T^{K-1} = \frac{8\tau^2}{K-1} < \frac{8\tau_K^2}{K-1} \leq 32\,\frac{K}{K-1}\left(\left(\frac{8\pi}{\sqrt{3}}\right)^{1/2} - \frac{C}{K^{1/6}}\right)^{-2}.$$

In the limit $\tau \to \infty$, we know that $K \to \infty$, and thus

$$\limsup_{\tau\to\infty} \mathcal{S}_{\min} \geq \frac{4\sqrt{3}}{\pi}\,.$$

Constructing a lower bound is more difficult. From (6.1) it is obvious that the operator $\mathcal{D}_\mu$ is positive semi-definite if $\tau \leq \sqrt{3}$. Thus we can apply Proposition 4.11

to obtain

$$\mathcal{S}_{\min} \geq \nu_0 \qquad \text{if } \tau \leq \sqrt{3}\,.$$

If $\tau \leq \sqrt{2}$, this lower bound is even equal to $\mathcal{S}_{\min}$ according to Corollary 6.1. As shown in Figure 6, the estimate is no longer optimal if $\tau > \sqrt{2}$.

Another method to obtain lower bounds is based on the following observation:

Proposition 6.3. *Assume that K_μ is an integral operator on $\mathcal{H}_\mu$ with integral kernel $K \in C^0(S^2 \times S^2, \mathbb{R})$ with the following properties:*

(a) *$K(x,y) \leq \mathcal{L}(x,y)$ for all $x, y \in S^2$.*

(b) *The operator K_μ is positive semi-definite.*

Then the minimal action satisfies the estimate

$$\mathcal{S}_{\min} \geq \iint_{S^2 \times S^2} K(x,y)\, d\mu(x)\, d\mu(y)\,. \tag{6.3}$$

Proof. For any $\rho \in \mathfrak{M}$, assumption (a) gives rise to the estimate

$$\mathcal{S}[\rho] = \iint_{S^2\times S^2} \mathcal{L}(x,y) d\rho(x) d\rho(y) \geq \iint_{S^2\times S^2} K(x,y) d\rho(x) d\rho(y)\,.$$

Next, using property (b), we can apply Proposition 4.3 to conclude that the volume measure μ is a minimizer of the variational principle corresponding to K, i.e.

$$\iint_{S^2\times S^2} K(x,y)\, d\rho(x)\, d\rho(y) \geq \iint_{S^2\times S^2} K(x,y)\, d\mu(x)\, d\mu(y)\,.$$

Combining these inequalities gives the result. □

In order to construct a suitable kernel, we first consider the heat kernel h_t on S^2,

$$h_t(x,y) = \left(e^{t\Delta_{S^2}}\right)(x,y) = 4\pi \sum_{l=0}^\infty e^{-t\,l(l+1)} \sum_{m=-l}^{l} Y_l^m(x) \overline{Y_l^m(y)}\,.$$

The heat kernel has the advantage that condition (b) is satisfied, but condition (a) is violated. This leads us to choosing K as the difference of two heat kernels,

$$K(x,y) = \lambda \left(h_{t_1}(x,y) - \delta h_{t_2}(x,y)\right)\,. \tag{6.4}$$

For given $t_1 < t_2$, we choose δ and λ such that $K(x,x) = 1$ and $K(\vartheta_{\max}) = 0$, i.e.

$$\delta = \frac{h_{t_1}(\vartheta_{\max})}{h_{t_2}(\vartheta_{\max})} < 1 \qquad \text{and} \qquad \lambda = \frac{\mathcal{L}(0)}{h_{t_1}(0) - \delta\, h_{t_2}(0)} > 0\,.$$

By direct inspection one verifies that condition (a) is satisfied (see Figure 7 for a typical example). The eigenvalues of the operator K_μ are computed to be

$$\lambda\left(e^{-t_1\, l(l+1)} - \delta\, e^{-t_2\, l(l+1)}\right),$$

showing that the operator K_μ is indeed positive semi-definite. Thus we can apply Proposition 6.3. Using that

$$\iint_{S^2\times S^2} h_t(x,y)\, d\mu(x)\, d\mu(y) = \iint_{S^2\times S^2} 4\pi\, Y_0^0(x) \overline{Y_0^0(y)}\, d\mu(x)\, d\mu(y) = 1\,,$$

we obtain the **heat kernel estimate**

$$\mathcal{S}_{\min} \geq S_K = \lambda\,(1-\delta)\,.$$

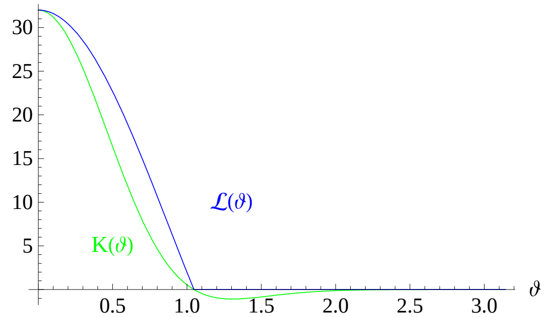

FIGURE 7. The Lagrangian $\mathcal{L}$ and the function K in the heat kernel estimate for $\tau = 2$.

In this estimate, we are still free to choose the parameters t_1 and t_2. By adjusting these parameters, one gets the lower bound shown in Figure 6. Thus the heat kernel estimate differs from the minimal action only by an error of about 20%, and describes the qualitative dependence on τ quite well. But of course, it does not take into account the discreteness of the minimizers.

7. The Variational Principles on the Flag Manifold $\mathcal{F}^{1,2}(\mathbb{C}^f)$

We finally make a few comments on the variational principles on the flag manifold $\mathcal{F}^{1,2}(\mathbb{C}^f)$ (see Example (c) on page 5). We first apply our main Theorems 4.15 and 4.17 to obtain the following general result.

Theorem 7.1. *Every minimizer ρ on $\mathcal{F}^{1,2}$ is either generically timelike or* $\overset{\circ}{\text{supp}}\, \rho = \varnothing$.

Proof. As a homogeneous space, the flag manifold $\mathcal{F}^{1,2}(\mathbb{C}^f)$ has a real analytic structure (see [11, Chapter II, §4]). Then the function $\mathcal{D}$ is obviously real analytic. Moreover, it is symmetric and constant on the diagonal. In order to apply Theorem 4.15, for given $y \in \mathcal{K}(x)$ we must find a curve c joining x and y which satisfies (4.11). Alternatively, in order to apply Theorem 4.17, our task is to construct a curve $c(t)$ with $c(0) = x$ and $c(1) = y$ which is analytic in a neighborhood of $t = 0$, such that the function $\mathcal{D}(c(t), y)$ changes sign at $t = 0$. Since in the case $\tau = 1$, the sets $\mathcal{K}(x)$ are all empty, we may assume that $\tau > 1$.

We denote the range of x by $I \subset \mathbb{C}^f$ and the orthogonal projection to I by π_I. Choosing an orthonormal basis (e_1, e_2) of I, the matrix $x|_I$ can be represented with Pauli matrices by

$$x|_I = \mathbb{1} + \tau\, \vec{u}\vec{\sigma} \qquad \text{with } \vec{u} \in S^2 \,.$$

Similarly, the operator $\tilde{y} := \pi_I y \pi_I$ has the representation

$$\tilde{y}|_I = \rho \mathbb{1} + \kappa\, \vec{v}\vec{\sigma} \qquad \text{with } \vec{v} \in S^2 \,,$$

where the real parameters ρ and κ satisfy the inequalities

$$1 - \tau \le \rho - \kappa \le 0 \le \rho + \kappa \le 1 + \tau \,.$$

Using (2.8), the function $\mathcal{D}$ is computed by

$$\mathcal{D}(x, y) = 2\Big((\rho\tau + \kappa \cos\vartheta)^2 - \kappa^2\, (\tau^2 - 1)\, \sin^2\vartheta \Big),$$

where ϑ denotes the angle between $\vec{u}$ and $\vec{v}$. The operator $\tilde{y}$ has rank two if and only if $\kappa > |\rho|$. A short calculation shows that in this case, $\mathcal{D}$ only has transverse zeros. Thus we can choose a direction $\dot{c}(0)$ where the condition (4.11) is satisfied. Choosing a smooth curve starting in this direction which joins x and y, we can apply Theorem 4.15 (A) to conclude the proof in this case.

It remains to consider the situation when $\tilde{y}$ has rank at most one. This leads us to several cases. We begin with the case when $y|_I$ vanishes. In this case, we may restrict attention to the four-dimensional subspace $U = \text{Im}\, x \oplus \text{Im}\, y$. In a suitable basis $(e_1, \ldots, e_4)$ of this subspace, the operators x and y have the matrix representations

$$x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes (\mathbb{1} + \tau\, \vec{u}\vec{\sigma})\,, \qquad y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes (\mathbb{1} + \tau\, \vec{v}\vec{\sigma})\,,$$

where again $\vec{u}, \vec{v} \in S^2$. A unitary transformation of the basis vectors e_1 and e_2 describes a rotation of the vector $\vec{u}$ in $\mathbb{R}^3$. By a suitable transformation of this type, we can arrange that the angle between $\vec{u}$ and $\vec{v}$ equals $\vartheta_{\max}$ (see (2.7)). We now define the curve $c : [0, \pi] \to \mathcal{F}^{1,2}$ by

$$c(t) = \begin{pmatrix} \cos(t)^2 & \sin(t)\cos(t) \\ \sin(t)\cos(t) & \sin(t)^2 \end{pmatrix} \otimes (\mathbb{1} + \tau\, \vec{w}(t)\, \vec{\sigma}), \tag{7.1}$$

where $\vec{w} : [0, \pi] \to S^2$ is the geodesic on S^2 with $\vec{w}(0) = \vec{u}$ and $\vec{w}(\pi) = \vec{v}$. The curve c is a real analytic function with $c(0) = x$ and $c(\pi) = y$, which is obviously translation symmetric. Furthermore, one computes

$$\mathcal{D}(c(t), y) = \sin(t)^4\, \mathcal{D}_{S^2}(\vec{w}(t), \vec{v}),$$

where $\mathcal{D}_{S^2}$ is the corresponding function on the unit sphere (2.6). As $\mathcal{D}_{S^2}(\vartheta)$ changes sign at $\vartheta_{\max}$, the function $\mathcal{D}(c(t), y)$ changes sign at $t = 0$. Thus Theorem 4.17 (A) applies, completing the proof in the case $y|_I = 0$.

We next consider the case that $\tilde{y}$ has rank one. We choose the basis (e_1, e_2) of I such that $\tilde{y}$ is diagonal,

$$\tilde{y} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{with } a \neq 0.$$

An elementary consideration shows that we can extend the basis of I to an orthonormal system (e_1, e_2, e_3) such that on the subspace $J := \langle \{e_1, e_2, e_3\} \rangle$, the operator $\hat{y} := \pi_J y \pi_J$ has the form

$$\hat{y}|_{\langle \{e_1, e_2, e_3\} \rangle} = \begin{pmatrix} a & 0 & \overline{b} \\ 0 & 0 & 0 \\ b & 0 & c \end{pmatrix} \qquad \text{with } a \neq 0 \text{ and } ac \neq |b|^2\,. \tag{7.2}$$

We let U be the unitary transformation

$$U(t)|_J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix} \quad \text{and} \quad U(t)|_{J^\perp} = \mathbb{1}\,. \tag{7.3}$$

Setting $y(t) = U(t)\, y\, U(t)^{-1}$, the matrix $\tilde{y}$ becomes

$$\tilde{y}(t) = \begin{pmatrix} 1 & 0 \\ 0 & \sin t \end{pmatrix} (\rho\, \mathbb{1} + \kappa\, \vec{v}\vec{\sigma}) \begin{pmatrix} 1 & 0 \\ 0 & \sin t \end{pmatrix}\,,$$

where ρ and κ are new parameters with

$$\kappa > |\rho| \qquad \text{and} \qquad \rho + \kappa\, v_3 = a \neq 0 \tag{7.4}$$

and $\vec{v} \in S^2$ is again a unit vector. The function $\mathcal{D}$ is now computed by

$$\begin{aligned}\mathcal{D}(x,y(t)) &= \frac{1}{2}\,\mathrm{Tr}\big(x|_I\,\tilde{y}(t)\big)^2 - 2\det\big(x|_I\,\tilde{y}(t)\big) \\ &= \frac{1}{2}\Big(\mathrm{Tr}\big(x|_I\,\tilde{y}(t)\big)^2 - 4\,(\tau^2-1)\,(\kappa^2-\rho^2)\,\sin^2 t\Big). \qquad (7.5)\end{aligned}$$

In order to simplify the trace, we transform the phase of e_3. This changes the phase of b in (7.2), thus describing a rotation of the vector $\vec{v}$ in the $(1,2)$-plane. This makes it possible to arrange that the vectors (u_1,u_2) and (v_1,v_2) are orthogonal in $\mathbb{R}^2$. We thus obtain

$$\mathrm{Tr}\big(x|_I\,\tilde{y}(t)\big) = (1+\tau u_3)(\rho+\kappa v_3) + (1-\tau u_3)(\rho-\kappa v_3)\,\sin^2 t\,.$$

We now have two subcases:

(1) $v_3 \neq \pm 1$: We vary the vectors $\vec{u}$ and $\vec{v}$ as functions of t such that the above orthogonality relations remain valid and

$$u_3 = \cos(\vartheta+\alpha t)\,, \qquad v_3 = \cos(\varphi+\beta t)$$

with free "velocities" α and β. Since $\mathcal{L}(x,y)=0$ at $t=0$, we know that

$$\cos\vartheta = -\frac{1}{\tau}\,, \qquad \sin\vartheta = \frac{\sqrt{\tau^2-1}}{\tau} \neq 0\,. \qquad (7.6)$$

A Taylor expansion yields

$$\begin{aligned}\mathrm{Tr}\big(x|_I\,\tilde{y}(t)\big) &= -t\,\alpha\tau\,(\rho+\kappa v_3)\,\sin\vartheta \qquad (7.7)\\ &+\frac{t^2}{2}\Big((4+\alpha^2)\,\rho + (-4+\alpha^2)\kappa\cos\varphi + 2\alpha\beta\kappa\tau\,\sin\vartheta\,\sin\varphi\Big) + \mathcal{O}(t^3)\,. \qquad (7.8)\end{aligned}$$

As the factor $(\rho+\kappa v_3)$ is non-zero in view of (7.4), the linear term (7.7) does not vanish whenever $\alpha \neq 0$. By suitably adjusting α, we can arrange that the square of this linear term compensates the last term in (7.5) (which is also non-zero in view of our assumption $\kappa > \rho$). Next, we know from (7.6) and our assumptions that the term $\sim \alpha\beta$ in (7.8) is non-zero. Thus by a suitable choice of β, we can give the quadratic term (7.8) any value we want. Taking the square, in (7.5) we get a contribution $\sim t^3$. Thus the function $\mathcal{D}$ changes sign. Transforming to a suitable basis where y is a fixed matrix, we obtain a curve $x(t)$ which is locally translation symmetric. Extending this curve to a smooth curve c which joins the point y, we can apply Theorem 4.17 (A).

(2) $v_3 = \pm 1$: We know that the matrix $\tilde{y}$ is diagonal,

$$\tilde{y}(t) = \begin{pmatrix} \rho\pm\kappa & 0 \\ 0 & (\rho\mp\kappa)\sin^2 t \end{pmatrix}\,. \qquad (7.9)$$

Now we keep v fixed, while we choose the curve $u(t)$ to be a great circle which is inclined to the $(1,3)$-plane by an angle $\gamma \neq 0$, i.e.

$$u_3 = \cos(\vartheta+\alpha t)\,\cos\gamma\,.$$

Repeating the above calculation leading to (7.7) and (7.8), one sees that we again get a non-zero contribution to $\mathcal{D}$ of the order $\sim t^3$. Thus $\mathcal{D}$ again changes sign, making it possible to apply Theorem 4.17 (A).

It remains to consider the case when $\tilde{y}$ vanishes but $y|_I \neq 0$. A short consideration shows that $y|_I$ cannot have rank two. Thus we can choose the orthonormal basis (e_1, e_2) of I such that $ye_1 \neq 0$ and $ye_2 = 0$. By suitably extending this orthonormal system by e_3 and e_4, we can arrange that the operator y is invariant on the subspace $\langle\{e_1, e_2, e_3, e_4\}\rangle$ and has the matrix representation

$$y|_{\langle\{e_1,e_2,e_3,e_4\}\rangle} = \begin{pmatrix} 0 & 0 & \overline{a} & 0 \\ 0 & 0 & 0 & 0 \\ a & 0 & c & \overline{b} \\ 0 & 0 & b & 0 \end{pmatrix} .$$

If $b \neq 0$, we can again work with the curve (7.1). If on the other hand $b = 0$, the operator y is invariant on $\langle\{e_1, e_2, e_3\}\rangle$ and has the canonical form

$$y|_{\langle\{e_1,e_2,e_3\}\rangle} = \begin{pmatrix} 0 & 0 & \sqrt{\tau^2-1} \\ 0 & 0 & 0 \\ \sqrt{\tau^2-1} & 0 & 2 \end{pmatrix} .$$

Transforming y by the unitary matrix

$$V(\tau) \begin{pmatrix} e_1 \\ e_3 \end{pmatrix} = \begin{pmatrix} \cos\tau & \sin\tau \\ -\sin\tau & \cos\tau \end{pmatrix} \begin{pmatrix} e_1 \\ e_3 \end{pmatrix} ,$$

we can arrange that y is again of the form (7.2), but now with coefficients depending on τ. Setting $t = \tau^2$, we can again use the construction after (7.2). This completes the proof. ☐

For sufficiently large τ, we can rule out one of the cases in Theorem 7.1, showing that the minimizing measures do have a singular support.

Theorem 7.2. *There are no generically timelike minimizers if*

$$\tau^2 > \frac{3f + 2\sqrt{3\,(f^2-1)}}{(2+f)} .$$

The method of proof is to apply Proposition 4.13 (I). In the next two lemmas we verify the necessary assumptions and compute ν_0.

Lemma 7.3. *The operator $\mathcal{D}_\mu$ has rank at most $3f^4$.*

Proof. We extend the method used in the proof of [8, Lemma 1.10]. A point $x \in \mathcal{F}$ is a Hermitian $f \times f$-matrix of rank two, with non-trivial eigenvalues $1+\tau$ and $1-\tau$. Thus, we can represent x in in bra/ket notation as

$$x = |u(x)\rangle\langle u(x)| - |v(x)\rangle\langle v(x)| ,$$

where $u(x)$ and $v(x)$ are the eigenvectors of x, normalized such that

$$\langle u(x)|u(x)\rangle = \tau + 1 \qquad \text{and} \qquad \langle v(x)|v(x)\rangle = \tau - 1 .$$

A short calculation shows that the non-trivial eigenvalues of the matrix product xy coincide with the eigenvalues of the 2×2-matrix product

$$A_{xy} := \begin{pmatrix} \langle u(x)|u(y)\rangle & -\langle u(x)|v(y)\rangle \\ \langle v(x)|u(y)\rangle & -\langle v(x)|v(y)\rangle \end{pmatrix} \begin{pmatrix} \langle u(y)|u(x)\rangle & -\langle u(y)|v(x)\rangle \\ \langle v(y)|u(x)\rangle & -\langle v(y)|v(x)\rangle \end{pmatrix} .$$

Using (2.8), we can thus write the function $\mathcal{D}$ as

$$\mathcal{D}(x,y) = \mathrm{Tr}\Big[\Big(A_{xy} - \frac{1}{2}\,\mathrm{Tr}(A_{xy})\Big)^2\Big].$$

This makes it possible to recover $\mathcal{D}(x,y)$ as the "expectation value"

$$\mathcal{D}(x,y) = \left\langle \begin{pmatrix} u\otimes u^*\otimes u\otimes u^* \\ u\otimes u^*\otimes v\otimes v^* \\ v\otimes v^*\otimes v\otimes v^* \end{pmatrix}\Big|_x , B \begin{pmatrix} u\otimes u^*\otimes u\otimes u^* \\ u\otimes u^*\otimes v\otimes v^* \\ v\otimes v^*\otimes v\otimes v^* \end{pmatrix}\Big|_y \right\rangle_{\mathbb{C}^{3f^4}}$$

of a suitable matrix B, whose 3×3 block entries are of the form

$$B_{ij} = b_{ij} + \delta_{i,2}\delta_{j,2}\,(c_1\rho_1 + c_2\rho_2 + c_3\rho_3) \quad \text{with} \quad b_{ij}, c_i \in \mathbb{C},$$

and the operators ρ_i permute the factors of the tensor product,

$$\begin{aligned}
\rho_1(u\otimes u^*\otimes v\otimes v^*) &= v\otimes v^*\otimes u\otimes u^* \\
\rho_2(u\otimes u^*\otimes v\otimes v^*) &= u\otimes v^*\otimes v\otimes u^* \\
\rho_3(u\otimes u^*\otimes v\otimes v^*) &= v\otimes u^*\otimes u\otimes v^* \,.
\end{aligned}$$

Hence introducing the operator

$$K \;:\; L^2(\mathcal{F}, d\mu_L) \rightarrow \mathbb{C}^{3f^4} \;:\; \psi \mapsto \int_{\mathcal{F}} \begin{pmatrix} u\otimes u^*\otimes u\otimes u^* \\ u\otimes u^*\otimes v\otimes v^* \\ v\otimes v^*\otimes v\otimes v^* \end{pmatrix}\Big|_x \psi(x)\, d\mu_L(x)\,,$$

we find that $\mathcal{D}_\mu = K^* B K$. This gives the claim. □

In view of this lemma, we may decompose $\mathcal{D}$ in the form (4.7).

Lemma 7.4. *The eigenvalue ν_0 in the decomposition* (4.7) *is given by*

$$\nu_0 = \frac{2(3f + 6f\tau^2 - (2+f)\tau^4 - 6)}{f(f^2-1)}\,. \tag{7.10}$$

Proof. Now it is now most convenient to represent the elements in $\mathcal{F}$ as

$$(1+\tau)\,|u\rangle\langle u| + (1-\tau)\,|v\rangle\langle v|\,, \tag{7.11}$$

where the vectors $u, v \in \mathbb{C}^f$ are orthonormal. Then the normalized volume measure μ on $\mathcal{F}$ can be written as

$$d\mu = \frac{1}{\text{vol}(\mathcal{F})}\,\delta\big(\text{Re}\,\langle u,v\rangle\big)\,\delta\big(\text{Im}\,\langle u,v\rangle\big)\,\delta\big(\|u\|^2-1\big)\,\delta\big(\|v\|^2-1\big)\,du\,dv,$$

where du and dv denotes the Lebesgue measure on $\mathbb{C}^f$. The total volume is computed to be

$$\begin{aligned}
\text{vol}(\mathcal{F}) &= \iint_{\mathbb{C}^f\times\mathbb{C}^f} \delta(\text{Re}\,\langle u,v\rangle)\delta(\text{Im}\,\langle u,v\rangle)\delta(\|u\|^2-1)\delta(\|v\|^2-1)du\,dv = \\
&= \frac{1}{4}\text{vol}(S^{2f-1})\text{vol}(S^{2f-3})\,.
\end{aligned}$$

To simplify the calculations, we fix x and choose an eigenvector basis of x. Then $x = \text{diag}((1+\tau),(1-\tau),0,\ldots,0)$, whereas y is again represented in the form (7.11). Then the eigenvalues of the product xy only depend on the components u_1, u_2, v_1, v_2. More precisely, using (2.8), we obtain

$$\begin{aligned}
\mathcal{D}(x,y) \;&= \frac{1}{2}\left[(1+\tau)^2|u_1|^2 + (1-\tau^2)(|v_1|^2 - |u_2|^2) - (1-\tau)^2|v_2|^2\right]^2 \\
&\quad + 2(1-\tau^2)\big|(1+\tau)u_1\overline{u_2} + (1-\tau)v_1\overline{v_2}\big|^2 =: f(u,v).
\end{aligned}$$

Our task is to compute the integral $\nu_0 = \int_{\mathcal{F}} f(u,v)\, d\mu$. In the case $f \geq 4$, one uses the symmetries to reduce to a lower-dimensional integral,

$$\begin{aligned}\nu_0 \;=\; c \int_0^\infty du_1 \int_0^\infty du_2 \int_0^\infty du_3 \int_0^\infty dv_1 \int_{\mathbb{C}} dv_2 \int_{\mathbb{C}} dv_3 \int_0^\infty dv_4 \\ \times\, \delta\big(\|u\|^2-1\big)\, \delta\big(\|v\|^2-1\big)\, \delta\left(\mathrm{Re}\left(\langle u,v\rangle\right)\right) \delta\left(\mathrm{Im}\left(\langle u,v\rangle\right)\right)) f(u,v)\, u_1 u_2 u_3^{2f-5} v_1 v_4^{2f-7},\end{aligned}$$

where c is the constant

$$c = \frac{1}{\mathrm{vol}(\mathcal{F})}\, \mathrm{vol}(S^{2f-5})\, \mathrm{vol}(S^{2f-7})\, (2\pi)^3\,.$$

Now carrying out all integrals gives the claim. The proof in the case $f=3$ is similar. ☐

The remaining question is whether generically timelike minimizers exist for small τ. In the special case $\tau = 1$, the operator $\mathcal{D}_\mu = \mathcal{L}_\mu$ is positive semi-definite (see [8, Lemma 1.10]), so that Proposition 4.3 or similarly Proposition 4.11 yields that the standard volume measure is a generically timelike minimizer. However, if $\tau > 1$, these propositions can no longer be used, because the operator D_μ fails to be positive semi-definite:

Lemma 7.5. *If $f \geq 3$ and $\tau > 1$, the operator $\mathcal{D}_\mu$ has negative eigenvalues.*

Proof. Since $\mathrm{supp}\,\mu = \mathcal{F}$, it suffices to find two points $x_1, x_2 \in \mathcal{F}$ such that the corresponding Gram matrix $\mathcal{D}(x_i, x_j)$ is not positive semi-definite. For given $\varepsilon \in (0,1)$ we choose the four vectors

$$u_1 = e_1\,, \quad v_1 = e_2 \quad \text{and} \quad u_2 = e_1\,, \quad v_2 = \sqrt{\varepsilon}\, e_2 + \sqrt{1-\varepsilon}\, e_3$$

(where e_i are the standard basis vectors of $\mathbb{C}^f$). Taking the representation (7.11), we obtain two points $x_1, x_2 \in \mathcal{F}$. The corresponding Gram matrix is computed to be

$$\begin{pmatrix} 8\tau^2 & \frac{1}{2}\left(-\varepsilon(\tau-1)^2 + (\tau+1)^2\right)^2 \\ \frac{1}{2}\left(-\varepsilon(\tau-1)^2 + (\tau+1)^2\right)^2 & 8\tau^2 \end{pmatrix}.$$

The determinant of this matrix is negative for small $\varepsilon > 0$. ☐

In this situation, Proposition 4.14 still gives some information on the possible support of generically timelike minimizers. However, it remains an open problem whether and under which conditions generically timelike minimizers exist.

References

[1] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics, vol. 98, Springer-Verlag, New York, 1985.

[2] H.T. Croft, K.J. Falconer, and R.K. Guy, *Unsolved problems in geometry*, Problem Books in Mathematics, Springer-Verlag, New York, 1994, Corrected reprint of the 1991 original.

[3] A. Diethert, F. Finster, and D. Schiefeneder, *Fermion systems in discrete space-time exemplifying the spontaneous generation of a causal structure*, arXiv:0710.4420 [math-ph], Int. J. Mod. Phys. A **23** (2008), no. 27/28, 4579–4620.

[4] F. Finster, *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.

[5] ______, *Fermion systems in discrete space-time—outer symmetries and spontaneous symmetry breaking*, arXiv:math-ph/0601039, Adv. Theor. Math. Phys. **11** (2007), no. 1, 91–146.

[6] ______, *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.

[7] ______, *From discrete space-time to Minkowski space: Basic mechanisms, methods and perspectives*, arXiv:0712.0685 [math-ph], Quantum Field Theory (B. Fauser, J. Tolksdorf, and E. Zeidler, eds.), Birkhäuser Verlag, 2009, pp. 235–259.

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

[9] ______, *A formulation of quantum field theory realizing a sea of interacting Dirac particles*, arXiv:0911.2102 [hep-th] (2010).

[10] F. Finster, A. Grotz, and D. Schiefeneder, *Causal fermion systems: A quantum space-time emerging from an action principle*, in preparation (2011).

[11] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics, vol. 80, Academic Press Inc., New York, 1978.

[12] P.D. Lax, *Functional analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2002.

[13] E.B. Saff and A.B.J. Kuijlaars, *Distributing many points on a sphere*, Math. Intelligencer **19** (1997), no. 1, 5–11.

[14] N.J.A. Sloane, *Tables of sphere packings and spherical codes*, IEEE Trans. Inform. Theory **27** (1981), no. 3, 327–338. www.research.att.com/~njas/packings/dim3/

Fakultät für Mathematik, Universität Regensburg, D-93040 Regensburg, Germany
E-mail address: Felix.Finster@mathematik.uni-regensburg.de
E-mail address: Daniela.Schiefeneder@mathematik.uni-regensburg.de