Damir Ferizović, BSc MSc

Uniform Point Distributions and Energy Estimates on Compact Riemannian Manifolds

DOCTORAL THESIS

to achieve the university degree of

Doktor der Naturwissenschaften

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dipl.-Ing. Dr.techn. Peter Grabner

Institute of Analysis and Number Theory

Graz, May 2021

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date, Signature

to my family

Abstract

The focus of this cumulative dissertation are three papers that the author has submitted during his PhD. After a more detailed introduction to the thesis and the research, some background in Differential Geometry and Analysis on Manifolds is given in the first part. The content of the articles is collected in the second part and in some cases additional material is given. There we find results concerning the Green and Riesz s-energies on the 3-dimensional rotation group, the asymptotic behavior of the L^2-norm of Gegenbauer polynomials and a first systematic study of multivariate kernels and its potential theory.

Kurzfassung

Der Fokus dieser kumulativen Dissertation sind die drei Arbeiten des Autors die während seiner Studiendauer eingereicht wurden. Der erste Teil dieser These gibt eine umfassendere Einleitung in dieses Werk und den Forschungsschwerpunkt. Zudem wird eine Einleitung in die Differenzialgeometrie und Analysis auf Mannigfaltigkeiten gegeben. Der zweite und letzte Teil gibt den Inhalt der Forschungsarbeiten wieder mit einigen Ergänzungen. Hier finden wir Resultate über die Green und Riesz s-Energien auf der Lie Gruppe der Rotationen im 3-dimensionalen Raum, über das asymptotische Verhalten der L^2-Norm von Gegenbauer Polynomen und die erste systematische Untersuchung von multivariaten Kernen und die dazugehörige Potentialtheorie.

Acknowledgements

I thank the Austrian Science Fund (FWF): F5503 “Quasi-Monte Carlo Methods” and FWF: W1230 “Doctoral School Discrete Mathematics”, and the Austrian Marshall Plan Foundation for their financial support.

Even as I am writing these lines after four long years, it strikes me as a bit sudden I have to admit. It feels like ages ago when I shared my office with Stefan, Gigi, Adrian, Mahadi and Sarah. They were nice company, as were Niclas and Thomas. Most of them left, but I was lucky to find friends in Besmira and Amr too – all defended their thesis already and I was following in their footsteps up until this point. Of course, Marc and Tanja joined us many evenings for which we all were grateful. I also made wonderful friends with Deniz, Muaaz and Azim; and I look forward for us to become awfully rich. Hatice and Sabine accompanied me for years now and I hope they will for decades to come. With Ryan and Giovanni I share many scientific interests, and I hope our friendship grows over the years like our sweet, sweet h-index.

I had many nice moments with them, and even in the face of time, though some memories inevitably fade or mold into merely shades of reality, what will last is a warm feeling of nostalgia. I will always look back fondly on our dinner nights together, the times we went snowboarding or surfing, or the one time when we slayed a dragon. I thank them all.

I thank Jochen Resch, Michael Muhr, Irene Pfeifer-Wilfinger and Hermine Panzenböck for their competent support during the years. I thank Jörg Thuswaldner, Johannes Wallner and Mario Weitzer for being my DK-mentors and for proof reading some of my papers.

Finally, my thanks go to the people who made this work possible. I thank them for writing letters of recommendation and for being referees at my defense.

I thank Carlos Beltrán for his invitation to come to Santander and his jolly nature. It is no understatement that the work with him made my career possible. These first results gave me something to present at conferences and seminars, not to mention the pressure it took of me.

I thank Dmitriy Bilyk for his hospitality and collaboration. The experience of working with him was a joy, as was hanging out with him.

Finally I thank Peter Grabner for hiring me even though my background was far removed from point distributions. I thank him for his patience and support.

Contents

List of Acronyms and Symbols

Throughout the text, $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ will denote the sets of natural (starting with 1), real and complex numbers respectively. The letter n will usually denote the dimension of an underlying vector space or manifold and is understood to be a natural number.

We will use the big- and little-oh notation where for functions $f, g : \mathbb{N} \to \mathbb{R}$ we write $f = O(g)$ if $\lim_{j\to\infty} \frac{f(j)}{g(j)}$ exists and is bounded. The little-oh notation $f = o(g)$ will mean that $\lim_{j\to\infty} \frac{f(j)}{g(j)} = 0$.

The following symbols will be used, usually with no definition within the text:

$(a, b]$	$=$	$\{x \in \mathbb{R} \ : \ a < x \leq b\}$ for $a, b \in \mathbb{R}$
$\{y_j\}$	$=$	$\{y_j\}_{j\in J}$ a sequence indexed by some J, which is clear from context
$\mathbb{R}_{>0}$	$=$	$\{x \in \mathbb{R} \ : \ x > 0\}$
$\mathbb{S}^d$	...	the unit sphere in $\mathbb{R}^{d+1}$
$B_\rho(x)$	$=$	$\{y \in \mathbb{R}^n \ : \ \|x - y\| < \rho\}$
$\{e_j\}$	$=$	$\{e_1, \dots, e_n\}$ the standard basis of $\mathbb{R}^n$
$\langle S \rangle$	...	the smallest subgroup containing S
$\langle x, y \rangle$	...	the inner product in $\mathbb{R}^n$
$\|x\|$	...	the norm of x, with inner product clear from context
δ_j^k	...	the Kronecker delta; equals 1 iff $j = k$ and 0 otherwise
δ_p	...	the Dirac measure, gives a set measure 1 if it contains the point p
$\mathbb{N}_0$	$=$	$\mathbb{N} \cup \{0\}$
$\mathcal{C}^k$	...	the set of k-times continuosly differentiable functions for $k \in \mathbb{N}_0$
$J(f)$	...	the Jacobi matrix of $f \in \mathcal{C}^1$
$\mathcal{C}^\infty(X)$	$=$	$\bigcap_k \mathcal{C}^k(X) \dots$ the set of smooth functions from X to $\mathbb{R}$
$\mathcal{C}_0^k$	...	compactly supported functions of $\mathcal{C}^k$, for $k \in \mathbb{N}_0 \cup \{\infty\}$
$\mathrm{supp}(f)$	$=$	closure of $\{x \in \mathrm{dom}(f) \ : \ \|f(x)\| > 0\}$
$\mathbb{1}$	...	denotes the identity in the given context (matrix, map)
(M, g)	...	a Riemannian manifold with metric g
$\langle u, v \rangle_g$	...	the inner product of two tangent vectors u, v wrt. the metric g
(U, φ)	...	a chart domain and chart of the manifold at hand
$\partial_j^{\varphi,p}$	...	the j-th derivative at p with respect to the chart (U, φ)
$\partial_1^{\mathbb{1},t}$	...	the vector field in $\mathbb{R}$ of unit length pointing to $+\infty$, at t
$\mathfrak{X}(M)$	...	set of smooth vector fields on M
$\mathfrak{X}^*(M)$	...	set of smooth covector fields on M
$\Gamma_0^\infty(M)$	...	set of compactly supported elements of $\mathfrak{X}(M)$

$\mathcal{C}_m^{(\lambda)}$	...	the Gegenbauer polynomial of index λ and degree m
σ	...	the normalized surface measure on the sphere
ω_N	...	a set of points with N elements
$\nabla_g f$	...	the gradient of the function f with respect to the metric g
$-\mathrm{div} u$	...	the divergence of the vector field u wrt. to the metric g
Δ_g	...	the Laplace-Beltrami operator
$\mathcal{G}$	...	the Green function for the Laplace-Beltrami operator with zero mean
$\mathbb{P}(\Omega)$	...	the set of probability measures on Ω
$\mathcal{M}(\Omega)$	...	the set of finite, signed measures on Ω
2^X	...	the power set of X
$\omega(\alpha^{-1}\beta)$	...	the rotation angle distance of $\alpha, \beta \in \mathrm{SO}(3)$

Part I.

Introduction

1. Introduction to the Research and the Thesis

This thesis is a cumulative dissertation and based on the three papers [BF20], [Fer21] and [Bil+21b], which appear in Part II with usually moderate changes in content and form, when compared to their versions on arXiv.org – with one notable exception. The chapter on multivariate kernels has been extended by two counter examples to some sensible questions. Part I will give the necessary background in Differential Geometry, as well on Sobolev spaces and the Laplacian on a Riemannian manifold in order to follow [BF20] with some ease. There one also finds a potentially new result on the existence of a series expansion for the Green function for compact Riemannian manifolds of dimension 3.

The undertaken research lies in the field of point distributions and energy estimates on compact Riemannian manifolds. One of the aims was to estimate various energies or notions of discrepancy for a finite point set $\omega_N = \{x_1, \dots, x_N\} \subset X$, where X is a measure space. Discrepancy is based on a family of measurable test sets $T \subset 2^X$, usually balls when X is additionally endowed with a metric, and it quantifies how well distributed points are by comparison to the normalized volume measure σ on X:

$$\mathcal{D}(\omega_N) := \sup_{A \in T} \left| \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_A(x_j) - \sigma(A) \right|.$$

Discrepancies are difficult to bound and only probabilistic methods give near optimal results. This is closely related to uniform distribution theory and low-discrepancy sequences and finds applications in Quasi-Monte-Carlo methods, for instance. A nice overview for the sphere is given in [KS97], and a direct application of the related Thomson Problem is given in [Li+18].

Another way to measure well distribution of points is via energies: For a symmetric, lower semicontinuous kernel K, we define the K-energy of ω_N as

$$E_K(\omega_N) = \sum_{j=1}^{N} \sum_{\ell \neq j}^{N} K(x_j, x_\ell) \qquad \text{and} \qquad \mathcal{E}_K(N) = \min_{\omega_N \subset X} E_K(\omega_N).$$

Calculating $\mathcal{E}_K(N)$ or finding minimizing configurations is an extremely hard problem and actual minimizers have been found in only a handful of cases for the Coulomb potential and certain powers of it on the sphere, i.e. $K(x, y) = \frac{1}{\|x-y\|}$ (see [Sch13]).

This problem becomes tractable if X is Riemannian and $K = \mathcal{G}$ is the Green function for the Laplace-Beltrami operator Δ with zero mean, as we will see. Well distribution follows by [BCC19], for if ω_N attains the minimal possible $\mathcal{G}$-energy, or Green energy, then the associated discrete measure approaches the uniform distribution in X as $N \to \infty$.

Next we give short summaries of the articles in Part II, where we also highlight important results.

Green Energy

In [BF20], Carlos Beltrán and this author deduced a closed form expression for the Green function with zero mean on the Lie group of 3-dimensional rotation matrices SO(3) via series expansion:

$$\mathcal{G}(\alpha, \beta) = \left(\pi - \omega(\alpha^{-1}\beta)\right) \cot\left(\tfrac{\omega(\alpha^{-1}\beta)}{2}\right) - 1,$$

where we used the so-called *rotation angle distance* of $\alpha, \beta \in \mathrm{SO}(3)$, defined as:

$$\omega(\alpha^{-1}\beta) = \arccos\left(\frac{\mathbf{Trace}(\alpha^{-1}\beta) - 1}{2}\right).$$

This eigenfunction expansion is justified in the case of SO(3) since the group can be covered by a single chart (Euler angles) up to a set of measure zero. For domains in $\mathbb{R}^n$, this is then well known in the theory of elliptic Partial Differential Equations and follows for instance from Fredholm Theory, see chapter 9 of [RR04] for a reference. Note that the push-forward of the Laplacian is an elliptic partial differential operator of order two, and its eigenfunctions will just be precomposed with the chart map.

We further obtained upper bounds for the Green and Riesz s-energies for SO(3), by applying the theory of determinantal point processes (dpp). A lower bound was obtained thanks to a Lemma of N. Elkies.

Theorem 1.0.1. *Let $N = \binom{2L+3}{3}$ for $L \in \mathbb{N}$. Then*

$$-3\sqrt[3]{\pi} N^{4/3} + O(N) \leq \ \mathcal{E}_{\mathcal{G}}(N) \ \leq -4\left(\tfrac{3}{4}\right)^{4/3} N^{4/3} + O(N).$$

A deterministic algorithm to produce point sets $\omega_N \subset \mathrm{SO}(3)$ for arbitrary $N \in \mathbb{N}$ is given, which did not make it into the published version. Numerical experiments indicate optimal distribution with respect to the Green energy and further investigations will follow.

Gegenbauer polynomials

One of the key ingredients in [BF20] was the asymptotic behavior of

$$\|\mathcal{C}_n^{(2)}\|_2^2 := \int_0^1 \left[\mathcal{C}_n^{(2)}(x)\right]^2 \, \mathrm{d}x,$$

where $\mathcal{C}_n^{(2)}$ is a Gegenbauer polynomial $\mathcal{C}_n^{(\lambda)}$ with index $\lambda = 2$ and degree n.

Lemma 1.0.2. *Let ψ denote the digamma function and γ the Euler-Mascheroni constant. Then the Gegenbauer polynomials satisfy for $n \geq 2$:*

$$\big\|\mathcal{C}_{n-2}^{(2)}\big\|_2^2 = \tfrac{1}{16}n^4 + \tfrac{1}{64}(4n^2-1)\big(\psi(n+\tfrac{1}{2}) + \gamma + \log(4)\big) - \tfrac{5}{32}n^2.$$

A potential first step to generalize results from SO(3) to rotation groups of higher dimension is finding the leading asymptotic terms of this kind of integrals with arbitrary index. It was certainly a surprise that no literature on the L^2-norm of Gegenbauer polynomials seemed to exist and we obtained following results.

Theorem 1.0.3. *The Gegenbauer polynomials satisfy for $\lambda > -\frac{1}{2}$, $\lambda \neq 0$, and $n > 1$:*

$$\big\|\mathcal{C}_{n-2}^{(\lambda+1)}\big\|_2^2 = \frac{n^2-2\lambda n}{2^4\lambda^3}\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 + \frac{n(2n+1)}{2^3\lambda^2}\big\|\mathcal{C}_n^{(\lambda)}\big\|_2^2 - \sum_{k=0}^{n-1}\frac{\lambda+k}{2^2\lambda^2}\big\|\mathcal{C}_k^{(\lambda)}\big\|_2^2.$$

For $\lambda > 1$, we have the following asymptotic in n:

$$\big\|\mathcal{C}_{n-2}^{(\lambda+1)}\big\|_2^2 = \frac{n^{4\lambda}}{4\lambda\Gamma(2\lambda+1)^2} + \frac{\lambda-1}{\Gamma(2\lambda+1)^2}n^{4\lambda-1} + O(n^{4\lambda-2}).$$

Potential Theory with Multivariate Kernels

This is a joint work with Dmitriy Bilyk, Alexey Glazyrin, Ryan Matzke, Josiah Park, and Oleksandr Vlasiuk. A generalization of K-energy minimizers is obtained by regarding arbitrary Borel probability measures $\mu \in \mathbb{P}(X)$ (since point sets ω_N always can be identified with Dirac masses of weight N^{-1}), thus one can seek to find

$$\inf_{\mu\in\mathbb{P}(X)} \iint K(x,y)\ \mathrm{d}\mu(x)\ \mathrm{d}\mu(y).$$

The book [BHS19] is an in-depth reference on this topic, but no theory exists on three-input kernels, which cover interesting cases like Menger curvature or Stillinger-Weber type potentials.

We developed a potential theoretic treatment of continuous and symmetric kernels K with multiple inputs on a compact metric space X, and determined conditions under which a minimizing probability measure can be identified. Let $K : X^n \to \mathbb{R}$ and $\mu \in \mathbb{P}(X)$, we define

$$I_K(\mu) = \int_X \cdots \int_X K(x_1, ..., x_n)\ d\mu(x_1)\cdots d\mu(x_n),$$

and, the potential function as

$$U_K^{\mu^{n-1}}(z) = \int_X \cdots \int_X K(x_1, ..., x_{n-1}, z)\ d\mu(x_1)\cdots d\mu(x_{n-1}).$$

If $n \geq 3$, we will say that K is (conditionally) n-positive definite, if

$$K_{z_1,\dots,z_{n-2}}(x,y) := K(x,y,z_1,\dots,z_{n-2})$$

is (conditionally) positive definite for fixed, arbitrary $z_1, \ldots, z_{n-2} \in X$. We will restrict our sample of results to kernels on the sphere which are in addition *rotationally invariant*, i.e. have the form

$$K(x_1, \ldots, x_n) = F\Big((\langle x_i, x_j \rangle)_{i,j=1}^n\Big).$$

Theorem 1.0.4. *Suppose that $K : (\mathbb{S}^{d-1})^n \to \mathbb{R}$ is continuous, symmetric, rotationally invariant, and conditionally n-positive definite on $\mathbb{S}^{d-1}$. Then σ is a minimizer of I_K over $\mathbb{P}(\mathbb{S}^{d-1})$.*

We obtain a particularly nice result for kernels of the type

$$K(x, y, z) = F(u, v, t), \qquad (\star)$$

where $u = \langle x, y \rangle$, $v = \langle y, z \rangle$ and $t = \langle z, x \rangle$.

Corollary 1.0.4.1. *Let $f : [-1, 1] \to \mathbb{R}$ be a real-analytic function with nonnegative Taylor coefficients and let $F(u, v, t) = f(uvt)$. Then, for K defined as in $(\star)$, the uniform surface measure σ minimizes the energy I_K over $\mathbb{P}(\mathbb{S}^{d-1})$.*

An additional result to the paper, where we deal with the special case of kernels of the type $K(x, y, z) = h(u)h(v)h(t)$ on the unit sphere of arbitrary dimension and h is a certain smooth function is given. This will give a counter example to at least two reasonable questions that one might ask about properties of multivariate kernels.

2. Essentials of Analysis on Manifolds

This chapter intends to collect sufficient material on Differential Geometry, Riemannian structure and the Laplace-Beltrami operator for a reader with some knowledge in these fields to follow [BF20] with ease. Further, a series expansion for compact Riemannian manifolds of dimension 3 is derived in subsection 2.3.2. A reference to the strong maximum principle that was missed out in the aforementioned paper is found in [Gri09], page 222.

2.1. Differential Geometry Redux

Here we summarize important definitions and theorems from [Lee13] – a book that this author wholeheartedly recommends. General properties of manifolds and Lie groups, in particular for SO(3) will be found in the next pages. A discussion of tensor products and manifolds with boundary has been left out, but will be used in the text. The notation differs from the book in a fair amount, and was inspired by lectures of Michael Eichmair[1], who got me interested in Differential Geometry in the first place and I thank him for that.

2.1.1. Manifolds

Definition 2.1.1. Let M denote a topological space. We say that M is a topological manifold of dimension n, or n-dimensional topological manifold if following holds:

1. M is a Hausdorff[2] space,
2. M is second countable, i.e. has a countable basis for its topology, and
3. every point in M has a neighborhood homeomorphic to some open set of $\mathbb{R}^n$.

A coordinate chart for M is always a pair (U, φ), where $U \subset M$ is open and $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ is a homeomorphism between open sets. For a fixed point p and a chart containing it, the map φ can be chosen such that $\varphi(p) = 0$, then called centered at p, and $\varphi(U)$ is an open ball or cube.

The projection onto the j-th coordinate $\mathrm{Pr}_j(\mathbf{v}) = v_j$ for $(v_1, \ldots, v_n)^t = \mathbf{v} \in \mathbb{R}^n$ is clearly continuous, hence $x_j := \mathrm{Pr}_j \circ \varphi$ is continuous, and $\varphi(q) = (x_1(q), \ldots, x_n(q))^t$. These x_j's are called local coordinates for the chart, and thus we can imagine coordinate

[1]University of Vienna.
[2]Thus two distinct points have disjoint neighborhoods, i.e. open sets containing them.

axes pulled back onto U and pretend to work in $\mathbb{R}^n$, and as long as we do it locally, nothing bad will happen.

We just note that the dimension of the manifold is invariant under homeomorphisms.

Proposition 2.1.1. *Let M be a topological manifold. Then*

- *M is locally path-connected,*
- *M is connected if and only if it is path-connected,*
- *connected and path-connected components are equivalent,*
- *M has countable many connected components, each an open set and connected topological n-manifold on its own right.*

A collection $\mathcal{X} \subset 2^M$ is called locally finite, if every point in M admits a neighborhood that intersects at most finitely many elements of $\mathcal{X}$.

Given a cover $\mathcal{U}$ of M, a refinement of $\mathcal{U}$ is another cover $\mathcal{V}$ of M, such that for each $V \in \mathcal{V}$, we find a $U \in \mathcal{U}$, such that $V \subset U$. M is said to be paracompact, iff every open cover admits an open, locally finite refinement.

Theorem 2.1.2. *A topological manifold is paracompact.*

Two charts (U, φ), (V, ψ) are said to be smoothly compatible if either $U \cap V = \emptyset$ or

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$$

is a diffeomorphism between open sets of $\mathbb{R}^n$. An atlas for M consists of charts whose domains cover M, and a smooth atlas is an atlas that has smoothly compatible charts. Compatibility defines an equivalence relation between smooth atlases, and the union of all compatible atlases is called maximal atlas or differentiable structure on M.

Definition 2.1.2. A smooth manifold M is a topological manifold with a smooth structure.

Not every topological manifold has a smooth structure, a counterexample was found by M. Kervaire in 1960.

Definition 2.1.3. A function $F : M \to N$ between smooth manifolds is called smooth if for each $p \in M$ we find chart neighborhoods (U, φ), (V, ψ) of p and $F(p)$ respectively, such that

$$F(U) \subset V \qquad \text{and} \qquad \psi \circ F \circ \varphi^{-1} : \varphi(U) \to \psi(V),$$

is smooth between opens sets of $\mathbb{R}^n$ to $\mathbb{R}^m$, where n, m are the dimensions of M and N respectively.

Note that the whole image $F(U)$ really needs to be contained in V.

Proposition 2.1.3. *Let $F : M \to N$ be a function between smooth manifolds. Then F is smooth if and only if one of either of the following conditions is satisfied:*

- *For each $p \in M$ we find chart neighborhoods (U, φ), (V, ψ) of p and $F(p)$ respectively, with $U \cap F^{-1}(V)$ open in M and the composition is smooth there*

$$\psi \circ F \circ \varphi^{-1} : \varphi(U \cap F^{-1}(V)) \to \psi(V).$$

- *If F is assumed continuous and atlases exists for M, N, such that all compositions*

$$\psi_\beta \circ F \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap F^{-1}(V_\beta)) \to \psi_\beta(V_\beta) \quad \textit{are smooth.}$$

Corollary 2.1.3.1 (Gluing Lemma for Smooth Maps)**.** *Let M, N be smooth manifolds and let $\{(U_\alpha, \varphi_\alpha)\}$ be an open cover of M. If further there is a family of smooth maps $\{F_\alpha\}$ defined on U_α mapping to N with the property that F_α, F_β when restricted to $U_\alpha \cap U_\beta$ are identical, then there is a unique smooth $F : M \to N$ that extends all $\{F_\alpha\}$.*

Smooth manifolds are said to be diffeomorphic if there is a smooth bijective map $F : M \to N$ whose inverse is smooth too. While manifolds usually have incompatible smooth structures, take $N = (\mathbb{R}, x^3)$ and $M = (\mathbb{R}, x)$ for instance where $x \circ (x^3)^{-1}$ is not smooth around the origin, they might still be diffeomorphic: $\sqrt[3]{\cdot} : M \to N$ is continuous and bijective, and $x^3 \circ \sqrt[3]{x} \circ x^{-1}$ is smooth. Indeed, all topological manifolds of dimension $n \leq 3$ have a unique smooth structure up to diffeomorphisms, which follows from results of J. Munkres 1960 and E. Moise 1977. There are also topological manifolds with many smooth structures that are not diffeomorphic, which follows by works of S. Donaldson and M. Freedman 1984 – and $\mathbb{R}^4$ is one such example, the only one among the $\mathbb{R}^n$'s.

Definition 2.1.4. Let M be a smooth manifold. A partition of unity subordinate to an open cover $\{X_\alpha\}$ of M is a family of real valued continuous functions $\{\psi_\alpha\}$ with following properties

1. $\psi_\alpha : M \to [0, 1]$,
2. $\mathrm{supp}(\psi_\alpha) \subset X_\alpha$ for each α,
3. the family $\{\mathrm{supp}(\psi_\alpha)\}$ is locally finite, and
4. $\sum_\alpha \psi_\alpha(p) = 1$ for each $p \in M$.

If the ψ's are in addition smooth, then we call this a smooth partition of unity.

Theorem 2.1.4 (Existence of Partition of Unity)**.** *On a smooth manifold there exists a smooth partition of unity subordinate to any open cover.*

The next proposition implies the existence of smooth bump functions and extensions of vector valued smooth functions.

Proposition 2.1.5. *Let M be a smooth manifold and $A \subset U \subset M$ be chosen, such that A is closed and U open. Then*

- *there is a smooth bump function for A supported in U, i.e. $\psi : M \to [0, 1]$ smooth with $A = \psi^{-1}(\{1\})$ and $\mathrm{supp}(\psi) \subset U$;*

- *for any smooth $f : A \to \mathbb{R}^k$, there exists a smooth $\tilde{f} : M \to \mathbb{R}^k$ identical to f on A and* $\operatorname{supp}(\tilde{f}) \subset U$.

Here it is important that the image lies in $\mathbb{R}^k$ and is not an arbitrary manifold N. Otherwise one first needs to establish the existence of a continuous extension.

An exhaustion function for M is a continuous function $f : M \to \mathbb{R}$ with the property that each sublevel set $f^{-1}((-\infty, c])$ for $c \in \mathbb{R}$ is compact.

Proposition 2.1.6. *Every smooth manifold admits a smooth exhaustion function with values in* $[1, \infty)$.

Theorem 2.1.7 (Level Sets of Smooth Functions)**.** *Let M be a smooth manifold and A be a closed subset. Then there exists a smooth function $f : M \to \mathbb{R}_{\geq 0}$ such that* $A = f^{-1}(\{0\})$.

2.1.2. Tangent structure

Let $\mathcal{C}^\infty(M)$ denote the vector space of smooth real valued functions defined on a smooth manifold M. A derivation at a fixed point $p \in M$ is a $\mathbb{R}$-linear map $\mathrm{w}_p : \mathcal{C}^\infty(M) \to \mathbb{R}$ which satisfies the product rule

$$\mathrm{w}_p(fg) = g(p)\mathrm{w}_p(f) + f(p)\mathrm{w}_p(g).$$

Definition 2.1.5. The tangent space of M at p is the vector space of all derivations at p, and is denoted by T_pM.

If follows that $\mathrm{w}_p(1) = \mathrm{w}_p(1 \cdot 1) = 2\mathrm{w}_p(1)$, thus derivations send constant functions to zero. Also, if $f(p) = g(p) = 0$, then $\mathrm{w}_p(fg) = 0$, i.e. functions with double roots will be send to zero too, as will be functions that are zero in a neighborhood of p.

Definition 2.1.6. The differential at p of a smooth map $F : M \to N$, is defined as a linear map

$$\begin{aligned} dF_p : T_pM &\to T_{F(p)}N \\ \mathrm{w}_p &\mapsto dF_p(\mathrm{w}_p), \end{aligned}$$

where $dF_p(\mathrm{w}_p)f = \mathrm{w}_p(f \circ F)$ for all $f \in \mathcal{C}^\infty(N)$.

Proposition 2.1.8 (Properties of Differentials)**.** *For smooth manifolds M, N and P, and smooth maps $F : M \to N$, $G : N \to P$, we have for $p \in M$*

- *dF_p is linear,*
- $d(G \circ F)_p = dF_p \circ dG_{F(p)} : T_pM \to T_{G\circ F(p)}P$,
- $d\mathbb{1}_p = \mathbb{1}_{T_pM}$,
- *if F is a diffeomorphism, the so is dF_p and* $\left(dF_p\right)^{-1} = dF^{-1}_{F(p)}$.

Proposition 2.1.9 (Tangent Space of Open Subsets)**.** *Let $U \subset M$ be open and $I : U \to M$ be the inclusion map. Then for every $q \in U$, $dI_q : T_qU \to T_qM$ is an isomorphism.*

To identify a derivation it is hence enough to know the values $w^j = \mathrm{w}_p(\mathrm{Pr}_j(\varphi))$, where (U, φ) is a centered chart neighborhood of p, with $\varphi(U)$ a ball in $\mathbb{R}^n$. For if $\mathbf{w} = (w^1, \ldots, w^n)^t$, then the map on $\mathcal{C}^\infty(U)$ defined via

$$f \mapsto \frac{d}{dt}\Big|_{t=0} f \circ \varphi^{-1}\big(t\mathbf{w}\big) = \langle \nabla\big(f \circ \varphi^{-1}\big)(0), \mathbf{w}\rangle$$

has the same action as w_p. To see this, let $g \in \mathcal{C}^\infty(U)$ and use Taylor expansion at 0:

$$g(q) = g \circ \varphi^{-1}(x) = g(p) + \langle \nabla\big(g \circ \varphi^{-1}\big)(0), \mathbf{x}\rangle + \langle \mathbf{x}, H\big(g \circ \varphi^{-1}\big)(\varepsilon)\mathbf{x}\rangle,$$

where $q = \varphi^{-1}(x)$ and H is the Hessian with $0 \le \varepsilon \le \|x\|$ chosen to satisfy the equality above. Since the term with the Hessian has a double root at 0, we apply w_p on both sides and obtain with $x^j = \mathrm{Pr}_j(\varphi)$

$$\mathrm{w}_p(g) = \langle \nabla\big(g \circ \varphi^{-1}\big)(0), \mathbf{w}\rangle.$$

Hence in particular, the dimension of T_pM is n, and a basis in the chart (U, φ) is given by $\partial_j^{\varphi,p}$, which act at p as

$$\partial_j^{\varphi,p}(f) = \partial_j\big(f \circ \varphi^{-1}\big)(\varphi(p)),$$

hence

$$\mathrm{w}_p(f) = \sum_{j=1}^{n} \mathrm{w}_p(\mathrm{Pr}_j(\varphi))\ \partial_j^{\varphi,p}(f).$$

Since dF_p is a linear map, a matrix representation is immediate once bases are chosen. Thus let T_pM have basis induced by (U, φ), and $T_{F(p)}N$ have one induced by a chart neighborhood (V, ψ) of $F(p)$ with $\psi(F) = (F_\psi^1, \ldots, F_\psi^m)^t$. Let $g \in \mathcal{C}^\infty(N)$, then

$$\begin{aligned}
dF_p(\partial_j^{\varphi,p})(g) &= \partial_j^{\varphi,p}(g \circ F) = \partial_j(g \circ \psi^{-1} \circ \psi \circ F \circ \varphi^{-1})(\varphi(p)) \\
&= \sum_{k=1}^{m} \partial_k(g \circ \psi^{-1})(\psi(F(p)))\ \partial_j(F_\psi^k \circ \varphi^{-1})(\varphi(p)) \\
&= \sum_{k=1}^{m} \partial_k^{\psi,F(p)}(g)\ \partial_j^{\varphi,p}(F_\psi^k),
\end{aligned}$$

hence

$$dF_p(\partial_j^{\varphi,p}) = \sum_{k=1}^{m} \partial_j^{\varphi,p}(\mathrm{Pr}_k(\psi \circ F)) \cdot \partial_k^{\psi,F(p)}. \tag{2.1}$$

This also helps to write down a change of coordinates formula, since the identity map $\mathbb{1} : M \to M$ with two different chart neighborhoods and suggestive notation yields

$$\partial_j^{\varphi,p} = d\mathbb{1}_p(\partial_j^{\varphi,p}) = \sum_{k=1}^{m} \partial_j^{\varphi,p}(\mathrm{Pr}_k(\psi)) \cdot \partial_k^{\psi,p} =: \Big\langle J(\psi \circ \varphi^{-1})\big|_{\varphi(p)} \cdot e_j, \vec{\partial}^{\psi,p}\Big\rangle. \tag{2.2}$$

Definition 2.1.7. The tangent bundle of M, denoted by TM, is the disjoint union of all tangent spaces

$$TM = \bigcup T_pM.$$

There is a natural projection map $\pi : TM \to M$, sending $\partial_j^{\varphi,p}$ to p.

Proposition 2.1.10. *If M is a smooth, n-dimensional manifold, then TM is in a natural way a smooth $2n$-manifold.*

A chart neighborhood $(\mathcal{U}, \Phi)$ of TM at (p, w_p) is given via charts (U, φ) of M around p with the notation as above:

$$\mathcal{U} \simeq U \times \mathbb{R}^n \qquad \text{and} \qquad \Phi(p, \mathrm{w}_p) = (\varphi(p), \mathbf{w})^t.$$

2.1.3. Whitney approximation theorems

Suppose we have a continuous function on a smooth manifold $\delta : M \to \mathbb{R}_{>0}$. Two functions $F_1, F_2 : M \to \mathbb{R}^m$ are said to be δ-close, if for all $p \in M$

$$\|F_1(p) - F_2(p)\| \leq \delta(p).$$

Theorem 2.1.11 (Whitney Approximation Theorem for Vector Functions)**.** *Suppose M is a smooth manifold with or without boundary, and $F : M \to \mathbb{R}^k$ is continuous. If $\delta : M \to \mathbb{R}_{>0}$ is continuous, then there is a smooth function $\tilde{F} : M \to \mathbb{R}^k$ with*

$$\|F(p) - \tilde{F}(p)\| \leq \delta(p)$$

for all $p \in M$. If in addition F is smooth on a closed subset of M, then $\tilde{F}$ can be chosen to agree with F there.

Theorem 2.1.12 (Whitney Approximation Theorem)**.** *Suppose M is a smooth manifold with or without boundary, and N is a smooth manifold without boundary. Let $F : M \to N$ be continuous, then it is homotopic to a smooth function.*

It follows that a smooth map $f : A \to N$ on a closed set $A \subset M$ admits a smooth extension to all of M if and only if it admits a continuous extension.

2.1.4. Lie groups

Recall that a group in the sense of algebra is a non-empty set G with a pairing $\circ : G^2 \to G$, usually written multiplicative or additive, such that there is a neutral or identity element $e \in G$, i.e. $\circ(g, e) = g$ for all $g \in G$ and each element h has an inverse h^{-1} satisfying $\circ(h, h^{-1}) = e$. We will use the multiplicative notation, i.e. $\circ(g, h) = gh$. A normal subgroup H of G is a subgroup that satisfies $gHg^{-1} = H$ for all $g \in G$.

Definition 2.1.8. A Lie group is a smooth manifold G that is a group in the algebraic sense, such that the action of taking products or inverses is smooth, i.e.

$$\cdot : G^2 \to G \qquad \text{and} \qquad \cdot^{-1} : G \to G,$$

are smooth maps between manifolds.

The left and right translations by an element g are natural diffeomorphisms acting as $L_g(h) = gh$ and $R_g(h) = hg$.

Definition 2.1.9. A smooth map $F : G \to H$ between Lie groups is a Lie group homomorphism if it is a group homomorphism. If F is a diffeomorphism that is also a group homomorphism, then we call F a Lie group isomorphism and G and H isomorphic.

A first example of a Lie group is $\mathrm{GL}(n, \mathbb{R})$, the set of real invertible $n \times n$ matrices – it is a group, and it can be thought of being a submanifold of $\mathbb{R}^{n^2}$. Products are smooth by nature, and inverses by Cramer's rule. If we regard $\mathbb{R} \setminus \{0\}$ as a multiplicative group, then taking the determinant is a smooth Lie group homomorphism between these groups.

Definition 2.1.10. Let S be a subset of the Lie group G, then we denote by $\langle S \rangle$ the smallest subgroup of G that contains S.

Proposition 2.1.13. *Let G be a Lie group and W a neighborhood of the identity. Then*

- *$\langle W \rangle$ is an open subgroup of G,*
- *if W is connected, then so is $\langle W \rangle$,*
- *if G is connected, then $\langle W \rangle = G$.*

The identity component of G is the connected component containing the identity.

Proposition 2.1.14. *Let G be a Lie group and G_e be the identity component. Then G_e is a normal subgroup of G, and it is the only open connected subgroup of G. Further, every connected component of G is diffeomorphic to G_e.*

Proposition 2.1.15. *Let $F : G \to H$ be a Lie group homomorphism. Then the kernel of F is a properly embedded Lie subgroup of G. If F is injective, then $F(G)$ can be made naturally a Lie subgroup of H.*

A left action of G onto any set M is a map $\vartheta : G \times M \to M$ such that for all $p \in M$ and $g, h \in G$ we have $\vartheta(e, p) = p$ and $\vartheta(g, \vartheta(h, p)) = \vartheta(gh, p)$. If M is a manifold, we call ϑ a smooth action and M a G-space. Similar statements and results hold for right actions.

Definition 2.1.11. The orbit of p is the set

$$\vartheta(G, p) = \{\vartheta(g, p) \ : \ g \in G\}.$$

The isotropy group or stabilizer of p is the set

$$G_p = \{g \in G \ : \ \vartheta(g, p) = p\}.$$

The action is said to be transitive if $\vartheta(G, p) = M$ for some and hence all $p \in M$. The action is said to be free if $G_p = \{e\}$ for all $p \in M$.

An example is given by $\mathrm{GL}(n, \mathbb{R})$ acting on $\mathbb{R}^n$. It is not transitive as the origin is not moved. It is not free as the origin is fixed by all elements.

Definition 2.1.12. Suppose the smooth manifolds M, N are G-spaces with smooth action given by ϑ and ζ. A function $F : M \to N$ is said to be equivariant with respect to ϑ, ζ if for all $p \in M$ and $g \in G$

$$F(\vartheta(g, p)) = \zeta(g, F(p)).$$

It is a theorem that every Lie group homomorphism has constant rank – recall that a smooth map $F : M \to N$ has constant rank if the rank of dF_p is the same for all $p \in M$.

Theorem 2.1.16 (Equivariant Rank Theorem)**.** *Let M and N be smooth manifolds and G-spaces for a Lie group G. If ϑ is transitive, and $F : M \to N$ is an equivariant smooth map, then F has constant rank.*

In particular this means that if F is surjective, it is a submersion; if F is injective, it is an immersion; and if F is bijective, then it is a diffeomorphism.

Recall that a submanifold S is embedded in M if its topology and charts are restrictions of the ones for M. A submanifold is properly embedded if the inclusion map is proper, i.e. the preimage of any compact set is compact.

Theorem 2.1.17 (Constant Rank Level Set Theorem)**.** *Let ϕ be a smooth map between smooth manifolds $M \to N$. If ϕ has constant rank r, then each level set of ϕ is a properly embedded submanifold of codimension r.*

An application is to show that the orthogonal group $\mathrm{O}(n)$ is a Lie group. It is clearly a subgroup of $\mathrm{GL}(n, \mathbb{R})$, and we can define a map $\phi : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ via $\phi(A) = A^t A$. Then $\mathrm{O}(n)$ is the level set $\phi^{-1}(\mathbb{1})$. We define a right action on $\mathrm{GL}(n, \mathbb{R})$ via $\rho(A, B) = B^t AB$ – thus $\rho(\rho(A, B), C) = C^t \rho(A, B)C = (BC)^t ABC = \rho(A, BC)$. The map ϕ is equivariant with respect to right translation in its domain and ρ in the codomain: $\phi(AB) = B^t A^t AB = B^t \phi(A)B = \rho(\phi(A), B)$. Hence by the Equivariant Rank Theorem it has constant rank, and by the Constant Rank Theorem all level sets are properly embedded submanifolds. We finish using following proposition.

Proposition 2.1.18. *Let H be a subgroup of the Lie group G, if H is also an embedded submanifold, then H is a Lie subgroup.*

It follows that $\mathrm{SO}(n)$ is a Lie group, for it is the identity component of $\mathrm{O}(n)$, and by Proposition 2.1.14 it is an open subgroup, thus each chart for $\mathrm{O}(n)$ restricted to $\mathrm{SO}(n)$ is a chart for the latter, hence being an embedded submanifold.

2.1.5. Vector fields

A (rough) vector field is a map $M \to TM$ with the property that p is sent to some derivation w_p. Thus vector fields act on functions $f \in \mathcal{C}^\infty(M)$ in the obvious way

$$Xf(p) := \mathrm{w}_p(f).$$

Definition 2.1.13. A smooth vector field is a smooth map $X : M \to TM$ such that its value at p is a derivation at p. The set of all smooth vector fields of M is denoted $\mathfrak{X}(M)$.

All our previous results on smooth functions between smooth manifolds apply here, and in particular X is smooth if and only if it is smooth in any chart (U, φ) around every $p \in M$; and in local coordinates

$$X\big|_U(p) = \sum_{j=1}^{n} X(\mathrm{Pr}_j(\varphi))(p)\ \partial_j^{\varphi,p}.$$

Evidently $\{\partial_j^{\varphi}\}$ are smooth vector fields on U, and for arbitrary X, we use the basis change formula to deduce that X is smooth if and only if the functions $X(\mathrm{Pr}_j(\varphi))$ are smooth in each chart domain. Note that smooth functions $F : M \to N$ do not yield a map $dF : \mathfrak{X}(M) \to \mathfrak{X}(N)$ in general[3], as $F(M)$ might not be all of N.

A smooth n-manifold is called parallelizable, if it admits a smooth global frame, i.e. there are smooth vector fields $E_1, \dots, E_n$ such that for all p, the derivations that these fields represent at p span T_pM. All Lie groups are parallelizable, and it was shown by Bott–Milnor and Kervaire that the only spheres that are parallelizable are $\mathbb{S}^1$, $\mathbb{S}^3$ and $\mathbb{S}^7$, but only the first two admit a Lie group structure.

Definition 2.1.14. The Lie bracket $[X, Y]$ of $X, Y \in \mathfrak{X}(M)$ is a smooth vector field of its own and defined as the derivation with action on $f \in \mathcal{C}^\infty(M)$ via

$$[X, Y]f = XYf - YXf.$$

Proposition 2.1.19 (Properties of the Lie Bracket)**.** *Let $X, Y, Z \in \mathfrak{X}(M)$ and $f, g \in \mathcal{C}^\infty(M)$, then the Lie bracket*

1. *is bilinear for the $\mathbb{R}$-module $\mathfrak{X}(M)$,*
2. *is antisymmetric,*
3. *satisfies the Jacobi identity*

$$\big[[X, Y], Z\big] + \big[[Y, Z], X\big] + \big[[Z, X], Y\big] = 0,$$

4. *satisfies*

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X.$$

Definition 2.1.15. A real vector space endowed with a bracket satisfying the first three properties is called a Lie algebra.

Recall the definition of left translation in a Lie group G by $L_g(h) = gh$ for all $h \in G$. This is a transitive action and $d(L_g) : \mathfrak{X}(G) \to \mathfrak{X}(G)$ is a diffeomorphism. A vector field $X \in \mathfrak{X}(G)$ is said to be left-invariant if for all $g, h \in G$

$$d(L_g)_h X(h) = X(gh).$$

[3]This circumstance gives rise to the notion of F-related vector fields.

Let us see what this means in local coordinates. Fix a chart neighborhood (U, φ) of $h \in G$ and one for gh, (V, ψ), so that following holds for $p \in U$ and $q \in V$

$$X(p) = \sum_{j=1}^{n} X(\mathrm{Pr}_j(\varphi))(p) \cdot \partial_j^{\varphi,p} \qquad \text{and} \qquad X(q) = \sum_{j=1}^{n} X(\mathrm{Pr}_j(\psi))(q) \cdot \partial_j^{\psi,q},$$

and, by (2.1)

$$d(L_g)_p(\partial_j^{\varphi,p}) = \sum_{k=1}^{n} \partial_j^{\varphi,p}(\mathrm{Pr}_k(\psi \circ L_g)) \cdot \partial_k^{\psi,gp}.$$

Now, left invariance means in particular, with some suggestive notation

$$X(\mathrm{Pr}_j(\psi))(gh) = \sum_{s=1}^{n} X(\mathrm{Pr}_s(\varphi))(h) \cdot \partial_s^{\varphi,h}(\mathrm{Pr}_j(\psi \circ L_g)) =: \langle X_h^{\varphi}, \nabla_h^{\varphi} L_g^{\psi,j} \rangle,$$

and we see that if in the right–hand–side we choose $h = e$ and let g run trough all elements of G, that X is completely determined at e and smooth, since L_g is. With this short hand notation and a chart (U, φ) containing e, we deduce that left invariant vector fields form a vector space $\mathfrak{g}$, since for two such fields X, Y and $a, b \in \mathbb{R}$ we have

$$a\langle X_e^{\varphi}, \nabla_e^{\varphi} L_g^{\psi,j} \rangle + b\langle Y_e^{\varphi}, \nabla_e^{\varphi} L_g^{\psi,j} \rangle = \langle aX_e^{\varphi} + bY_e^{\varphi}, \nabla_e^{\varphi} L_g^{\psi,j} \rangle.$$

Proposition 2.1.20. *For $X, Y \in \mathfrak{g}$, it follows that $[X, Y] \in \mathfrak{g}$. In particular, $\mathfrak{g}$ is a Lie algebra.*

We will denote this Lie algebra by $\mathrm{Lie}(G)$, and refer to it as the Lie algebra of G.

Theorem 2.1.21. *A vector space isomorphism is given by the evaluation map*

$$\begin{aligned} \epsilon : \mathrm{Lie}(G) &\to T_e(G) \\ X &\mapsto X(e). \end{aligned}$$

We readily see that every Lie group is parallelizable by a left-invariant frame. Further, by regarding $T_{\mathbb{1}}\mathrm{O}(n)$, we realize that the Lie algebra of $\mathrm{O}(n)$ is given by the vector space of skew symmetric matrices, and this is thus also the Lie algebra of $\mathrm{SO}(n)$.

Theorem 2.1.22 (Ado's Theorem)**.** *Every real finite dimensional Lie algebra admits a faithful finite dimensional representation.*

It thus follows that every real finite dimensional Lie algebra is isomorphic to a subalgebra of the matrix algebra with commutator bracket.

2.1.6. Tangent covectors on manifolds

For each $p \in M$, we denote the dual space of T_pM by T_p^*M, the cotangent space at p. Elements of T_p^*M are called (tangent) covectors at p. Given a chart neighborhood (U, φ) of p, we define the dual basis to satisfy

$$d^k_{\varphi,p}\big(\partial^{\varphi,p}_j\big) = \delta^k_j.$$

Hence given any element Λ_p of the cotangent space, we have in local coordinates $A_j := \Lambda_p\big(\partial^{\varphi,p}_j\big)$

$$\Lambda_p = \sum_{j=1}^{n} A_j \cdot d^j_{\varphi,p}.$$

Recall that for two overlapping chart domains (U, φ) and (V, ψ), the Jacobian matrices $J(\psi \circ \varphi^{-1})$ and $J(\varphi \circ \psi^{-1})$ are inverse to each other where defined, thus with (2.2) we obtain

$$d^j_{\varphi,p} = \sum_{k=1}^{m} \partial^{\psi,p}_k \mathrm{Pr}_j(\varphi) \cdot d^k_{\psi,p} =: \Big\langle J(\varphi \circ \psi^{-1})^t\big|_{\psi(p)} \cdot e_j, \vec{d}_{\psi,p} \Big\rangle =: \Big\langle \nabla^{\psi}_p \mathrm{Pr}_j(\varphi), \vec{d}_{\psi,p} \Big\rangle.$$

Definition 2.1.16. The cotangent bundle of M, denoted by T^*M, is the disjoint union of all cotangent spaces

$$T^*M = \bigcup T^*_p M.$$

There is a natural projection map $\pi : T^*M \to M$, sending $d^j_{\varphi,p}$ to p.

Proposition 2.1.23. *If M is a smooth manifold and n-dimensional, then T^*M is in a natural way a smooth $2n$-manifold.*

A chart neighborhood $(\mathcal{U}, \Phi)$ of T^*M at (p, Λ_p) is given via charts (U, φ) of M around p with $\mathbf{A} = (A_1, \ldots, A_n)^t$:

$$\mathcal{U} \simeq U \times \mathbb{R}^n \qquad \text{and} \qquad \Phi(p, \Lambda_p) = (\varphi(p), \mathbf{A})^t.$$

Thus local sections of this vector bundle are called covector fields, and a local coframe are n-many covector fields that are linearly independent and span T^*_pM. Clearly to each frame exists a coframe and vice versa. The real vector space of smooth covectors on M will be denoted by $\mathfrak{X}^*(M)$.

Geometrically, we can visualize a linear functional Λ_p as an affine $(n-1)$-dimensional subspace $I^p_\Lambda = \mathrm{Ker}(\Lambda_p) + v_p$; where v_p is the vector of minimal Euclidean length with $\Lambda_p(v_p) = 1$. Thus all affine subspaces parallel to I^p_Λ are level sets of Λ_p with value depending on the distance to the origin in a natural way: the plane I^p_Λ is the level set of the value 1. It follows that I^p_Λ contains all the information to recover Λ_p.

Definition 2.1.17. We define the differential of functions to be a covector field

$$\begin{aligned} d : \mathcal{C}^\infty(M) &\to \mathfrak{X}^*(M) \\ f &\mapsto df. \end{aligned}$$

For $\mathrm{w}_p \in T_pM$, df acts by the assignment $df_p(\mathrm{w}_p) = \mathrm{w}_p f$.

In local coordinates

$$df = \sum_{j=1}^{n} \partial_j^{\varphi} f \cdot d_{\varphi}^{j}.$$

The next statement is Proposition 2.1.8 for $\mathbb{R} = N$, where we identify $\mathbb{R} \simeq T\mathbb{R} \simeq T^*\mathbb{R}$.

Proposition 2.1.24. *Let $f, g \in \mathcal{C}^\infty(M)$, then*

- *d is $\mathbb{R}$-linear over $\mathcal{C}^\infty(M)$;*
- *$d(fg) = fdg + gdf$;*
- *$d\frac{f}{g} = \frac{gdf - fdg}{g^2}$ whenever $g \neq 0$;*
- *$d(g \circ f) = g' \circ f \cdot df$ whenever $g \circ f$ makes sense;*
- *if f is constant, then $df = 0$.*

It follows that $df = 0$ if and only if f is constant on connected components.

2.1.7. Riemannian structure

Definition 2.1.18. A Riemannian metric g on M is a smooth, symmetric covariant 2-tensor field that is positive definite at each point, i.e. for each p, the 2-tensor $g_p : T_pM^2 \to \mathbb{R}$ is an inner product. In local coordinates and in terms of tensor products, we can write

$$g_p = \sum_{j,s=1}^{n} g_{js}(p) d_{\varphi,p}^{j} \otimes d_{\varphi,p}^{s} = \sum_{j,s=1}^{n} g_{js}(p) d_{\varphi,p}^{j} \cdot d_{\varphi,p}^{s},$$

where $\{g_{js}\}$ is a strictly positive definite matrix with smooth coefficients for each p. The elements of the inverse matrix will be denoted by $\{g^{js}\}$.

Riemannian metrics can be assembled via partitions of unity and hence are not rare, quite the contrary, many different metrics would yield different geometric features.

Once a metric is set, we can define the norm or length of a tangent vector $v_p \in T_pM$:

$$|v_p|_g = \sqrt{\langle v_p, v_p \rangle_g} = \sqrt{g_p(v_p, v_p)},$$

and the angle θ between non-zero vectors v_p and w_p is defined as usual

$$\cos(\theta) = \frac{\langle v_p, w_p \rangle_g}{|v_p|_g |w_p|_g}.$$

Thus locally we can always have an orthonormal frame field via Gram-Schmidt algorithm.

Definition 2.1.19. For any smooth path $\gamma : (a, b) \to M$ with $a < b$, we define the length of γ to be

$$\ell(\gamma) = \int_a^b |d\gamma(\partial_1^{\mathbb{1},t})|_g \, \mathrm{d}t.$$

Let $p, q \in M$, then we set $P(p,q) = \{\gamma : [0,1] \to M \text{ smooth } : \gamma(0) = p,\ \gamma(1) = q\}$ and define the geodesic distance between these points as

$$d_g(p,q) = \inf_{\gamma \in P(p,q)} \ell(\gamma). \tag{2.3}$$

Since the Riemannian metric g is at each point essentially a strictly positive definite matrix, the geodesic distance is comparable to the Euclidean one for points close together.

Lemma 2.1.25 ([Gri09] Lem. 3.24)**.** *For any point p in a Riemannian manifold (M, g), there is a chart (U, φ) containing p, so that for any $x, y \in U$ and some $C = C(U) \geq 1$, we have*

$$C^{-1}\|\varphi(x) - \varphi(y)\| \ \leq\ d_g(x,y)\ \leq\ C\|\varphi(x) - \varphi(y)\|.$$

The product of two Riemannian manifolds (M, g_1) and (N, g_2) carries a natural Riemannian structure $g_1 \times g_2$: Given points $p \in M$ and $q \in N$, we identify $T_{p,q}(M \times N) \simeq T_pM \times T_qN$, and for $\xi_p, \zeta_p \in T_pM$, $\eta_q, \theta_q \in T_qN$ we define

$$\left(\begin{pmatrix} \xi_p \\ \eta_q \end{pmatrix}, \begin{pmatrix} \zeta_p \\ \theta_q \end{pmatrix} \right)_{g_1 \times g_2} = \langle \xi_p, \zeta_p \rangle_{g_1(p)} + \langle \eta_q, \theta_q \rangle_{g_2(q)}.$$

A flat Riemannian manifold is one that has a local isometry with $\mathbb{R}^n$ and its Euclidean metric, i.e. g is locally the pull back of the Euclidean metric.

Theorem 2.1.26. *For a Riemannian manifold (M, g) the following are equivalent.*

1. *g is flat,*
2. *for each $p \in M$, there is a smooth chart such that $g = \sum_{j=1}^n d_\varphi^j \cdot d_\varphi^j = \sum_{j=1}^n (d_\varphi^j)^2$,*
3. *for each $p \in M$, there is a smooth chart such that the resulting coordinate frame $\partial_1^\varphi, \dots, \partial_n^\varphi$ is orthonormal.*

It is the notion of curvature that measures how much g deviates from being flat.

2.1.8. The tangent-cotangent isomorphism

Every Riemannian metric induces a bundle homomorphism $\hat{g} : TM \to T^*M$ as follows: let $p \in M$ and $v_p \in T_pM$, then $\hat{g}[v_p] : T_pM \to \mathbb{R}$ acts via

$$\hat{g}[v_p](w_p) = g_p(v_p, w_p).$$

By linearity, and since $\hat{g}[v_p]$ is injective, it is a bundle isomorphism. In a chart (U, φ) around p, we have for vector fields X, Y

$$\hat{g}[X](Y) = g_{j,s} X^j Y^s,$$

where $X^j = X(\mathrm{Pr}_j(\varphi))$ and similar for Y^s. It follows that $\hat{g}[X] = \sum_{j=1}^n g_{js} X^j \cdot d_\varphi^s$ and it is customary to denote

$$X_s = g_{js} X^j,$$

whence we lowered the index. Using musical notation, one rather writes $X^\flat$ for $\hat{g}[X]$, thus the tone has been lowered. The inverse map $\hat{g}^{-1} : T^*M \to TM$, acts on Λ via

$$\hat{g}^{-1}[\Lambda] = g^{js}\Lambda(\partial_s^\varphi) \cdot \partial_j^\varphi.$$

Using again musical notation, one writes $\Lambda^\sharp$ for $\hat{g}^{-1}[\Lambda]$, where the index is raised. These are the musical isomorphisms. Using this, we can introduce the gradient of a function on Riemannian manifolds via

$$\mathrm{grad} f = \nabla_g f = df^\sharp = \sum_{s,j=1}^{n} g^{js}\partial_s^\varphi f \cdot \partial_j^\varphi.$$

It follows that $\nabla_g(fh) = h\nabla_g f + f\nabla_g h$, and we also obtain

$$\langle \nabla_g f, X\rangle_g = \sum_{s,j,k=1}^{n} g^{js}\partial_s^\varphi f g_{jk}X^k = \sum_{s,k=1}^{n} \delta_s^k \partial_s^\varphi f X^k = df(X).$$

Further, these operators introduce an inner product on T_p^*M at each p which is symmetric, strictly positive definite and smooth: let $\Lambda, \Gamma \in T^*M$, set

$$\langle \Lambda, \Gamma\rangle_{g^*} = \langle \Lambda^\sharp, \Gamma^\sharp\rangle_g.$$

This puts the isomorphism in musical isomorphisms.

2.2. Laplace-Beltrami Operator

There are many reasons to study the Laplacian, be it for applications in engineering or physics, or for its appearance in archetypical partial differential equations. An ansatz for the heat equation for instance, is to assume the solution to be of the form $u(x,t) = \alpha(t)\varphi(x)$ for $x \in \Omega$ and $t > 0$, which gives

$$\frac{\Delta\varphi(x)}{\varphi(x)} = -\frac{\dot{\alpha}(t)}{\alpha(t)}.$$

This automatically leads to the Helmholtz equation

$$\Delta\varphi = \lambda\varphi,$$

which for non-empty boundary $\partial\Omega$, usually demands φ to be zero on $\partial\Omega$, i.e. Dirichlet boundary conditions. But understanding or even finding the eigenvalues and -functions is a very hard task, and the literature has many results regarding upper and lower bounds for both these quantities, see Section 2.2.5 for more. Different compact shapes with the same eigenvalues and multiplicity share many properties, like volume, dimension, curvature, genus, etc. Results like these have been inspired by the classical paper of Kac and his question if one can hear the shape of a drum...

For a general treatment of partial differential operators on manifolds we refer to [BC17]. The following sections are based on the very well written and extremely useful [Gri09], and its predecessor [Gri01], which can be found on the homepage of A. Grigor'yan.

2.2.1. Riemannian volume measure and Laplace-Beltrami operator

Given a Riemannian manifold (M,g), then the change of coordinates for the metric looks as follows

$$\sum_{i,j=1}^{n} g_{ij}^{\varphi} d_{\varphi}^{i} d_{\varphi}^{j} = \sum_{k,s=1}^{n} \Big(\sum_{i,j=1}^{n} g_{ij}^{\varphi} \partial_k^{\psi} \mathrm{Pr}_i(\varphi) \partial_s^{\psi} \mathrm{Pr}_j(\varphi) \Big) \cdot d_{\psi}^{k} d_{\psi}^{s},$$

hence

$$\langle \partial_k^{\psi}, \partial_s^{\psi} \rangle_g = g_{ks}^{\psi} = \sum_{i,j=1}^{n} g_{ij}^{\varphi} \partial_k^{\psi} \mathrm{Pr}_i(\varphi) \partial_s^{\psi} \mathrm{Pr}_j(\varphi).$$

In matrix notation with $J = J(\varphi \circ \psi^{-1})$ this becomes particularly succinct

$$g^{\psi} = J g^{\varphi} J^{t}. \tag{2.4}$$

Every Riemannian manifold has a natural measure μ_g on the Borel σ-algebra, induced by the topology of M, and completed such that a set $A \subset M$ is measurable if $\varphi(U \cap A)$ is Lebesgue measurable for every chart (U,φ). Concretely, we define on U:

$$\mathrm{d}\mu_g|_U(p) = \sqrt{\det(g^{\varphi})}(p)\ \mathrm{d}\lambda(\varphi(p)). \tag{$\star$}$$

The value $\mu_g(A)$, for an open set $A \subset U$, is defined as the integral of $(\star)$ and does not depend on a particular chart. To see this let $A \subset V$, for a chart (V, ψ), then by (2.4)

$$\begin{aligned}\mu_g(A) &= \int_{\varphi(A)} \sqrt{\det(g^\varphi)}(\varphi^{-1}(x))\ \mathrm{d}\lambda(x) = \int_{\psi(A)} \sqrt{\det(g^\varphi)}\big(\psi^{-1}(x)\big)|\det J(\varphi \circ \psi^{-1})|\ \mathrm{d}\lambda(x) \\ &= \int_{\psi(A)} \sqrt{\det(Jg^\varphi J^t)}\big(\psi^{-1}(x)\big)\ \mathrm{d}\lambda(x) = \int_{\psi(A)} \sqrt{\det(g^\psi)}\big(\psi^{-1}(x)\big)\ \mathrm{d}\lambda(x).\end{aligned}$$

The measure μ_g is σ-finite, inner and outer regular. Further, since continuous functions are measurable, we have a bijection between measurable functions on U and $\varphi(U)$. One can now introduce the notion of measurable functions, and $L^p(U, \mu_g)$-spaces via lift from $\varphi(U)$ to U, and via partition of unity, to the whole of M. An immediate corollary is that if for all compactly supported smooth functions h we have

$$\int_M f \cdot h\ \mathrm{d}\mu_g = 0,$$

then $f \equiv 0$ on M.

Remark 2.2.1. Note that in the definition of μ_g, it is the term $\sqrt{\det(g)}$ that transforms in a way to make the definition independent of the chart. Integration on an oriented manifold uses that n-forms transform correctly once sign changes during variable transformations are removed. Nevertheless, orientation is still needed to introduce integration on a boundary and Stokes Theorem. Measures on manifolds are introduced with more detail in [Die72] on page 163.

Lemma 2.2.2. *Given a compact smooth manifold M with two Riemannian metrics g_1, g_2. Then $L^p(M, \mu_{g_1}) = L^p(M, \mu_{g_2})$ for all $0 < p \leq \infty$.*

Proof. This is immediate from Exercise 3.5 on page 64 of [Gri09]: the ratio of $\det(g_1)$ and $\det(g_2)$ is a well defined, positive function $r = r(g_1, g_2) \in \mathcal{C}^\infty(M)$ and

$$\mathrm{d}\mu_{g_1}|_U(p) = \sqrt{\det(g_1^\varphi)(p)}\ \mathrm{d}\lambda(\varphi(p)) = \sqrt{\frac{\det(g_1^\varphi)}{\det(g_2^\varphi)}}\sqrt{\det(g_2^\varphi)(p)}\ \mathrm{d}\lambda(\varphi(p)) = r(p)\ \mathrm{d}\mu_{g_2}|_U(p).$$

■

Theorem 2.2.3 (Divergence Theorem)**.** *Given a smooth vector field v on M, then there is a unique function $-\mathrm{div}_g v \in \mathcal{C}^\infty(M)$, such that for all $h \in \mathcal{C}_0^\infty(M)$ we have*

$$\int_M -\mathrm{div}_g v \cdot h\ \mathrm{d}\mu_g = \int_M \langle v, \nabla_g h\rangle_g\ \mathrm{d}\mu_g.$$

On compact manifolds $-\mathrm{div}_g(v)$ is hence a zero mean function. The Laplace-Beltrami operator is now defined as

$$\Delta_g = -\mathrm{div}_g(\nabla_g).$$

Functions f with $\Delta_g f = 0$ are called harmonic.

Theorem 2.2.4 ([Aub98] Thm. 1.71)**.** *Harmonic functions on compact smooth Riemannian manifolds are constants.*

Note that the theorem asks for an oriented manifold, but it easily follows via the measure theoretic approach that this is not necessary (and indeed it comes right before T. Aubin introduces Measure Theory).

2.2.2. Partial differential intermezzo

We will for the moment regard $\Omega \subset \mathbb{R}^n$ and consider the differential operator

$$L = \sum_{i,j=1}^{n} \partial_i(a_{ij}\partial_j) + \sum_{j=1}^{n} b_j\partial_j + c,$$

where $a_{ij}, b_j, c \in \mathcal{C}^\infty(\overline{\Omega})$ and we assume the matrix a_{ij} to be symmetric and uniformly elliptic with ellipticity constant $\lambda > 0$, and that b_j, c are bounded by λ, i.e. for $\xi = (\xi^1, \ldots, \xi^n) \in \mathbb{R}^n$

$$\sum_{i,j=1}^{n} \xi^i a_{ij}(x)\xi^j \geq \lambda\|\xi\|^2 \qquad \text{for } x \in \Omega.$$

The appearing Sobolev space will be defined in the next section on manifolds, we just state some theorems from PDE Theory for later reference.

Definition 2.2.1. Let $u \in W^{1,2}_{loc}(\Omega)$ and $f \in L^2_{loc}$, then we say that the equation $Lu = f$ holds weakly in Ω, if for all $h \in \mathcal{C}^\infty_0(\Omega)$, we have

$$-\sum_{i,j=1} \int_\Omega a_{ij}\partial_j u \cdot \partial_i h \,\mathrm{d}x + \sum_{j=1}^{n} \int_\Omega b_j\partial_j u \cdot h \,\mathrm{d}x + \int_\Omega uch \,\mathrm{d}x = \int_\Omega fh \,\mathrm{d}x.$$

Theorem 2.2.5. *If L is as above, and u is such that $u \in W^{1,2}_{loc}(\Omega)$ and $Lu \in W^{k,2}_{loc}(\Omega)$, then it follows that $u \in W^{k+2,2}_{loc}(\Omega)$. Further, for any open and compactly contained $U \subset \Omega$ and some constant C depending on U, Ω, n, λ, we have*

$$C\|u\|_{W^{k+2,2}(U)} \leq \|u\|_{W^{1,2}(\Omega)} + \|Lu\|_{W^{k,2}(\Omega)}.$$

We will write $L^k u = f$ if for each $1 \leq s < k$, $L^s u \in W^{1,2}_{loc}(\Omega)$ and $L\big(L^{k-1}u\big) = f$ weakly.

Corollary 2.2.5.1. *Assume $u, Lu, \ldots, L^k u \in W^{1,2}_{loc}(\Omega)$, then*

$$u \in W^{2k+1,2}_{loc}(\Omega),$$

and for any open and compactly contained $U \subset \Omega$ and some $C = C(U, \Omega, n, \lambda, k, \rho)$

$$C\|u\|_{W^{2k+1,2}(U)} \leq \sum_{j=1}^{n} \|L^j u\|_{W^{1,2}(\Omega)}.$$

By the Sobolev Embedding Theorem, if $2k+1 > \frac{n}{2}+m$ for $m \in \mathbb{N}$, then $u \in W^{2k+1,2}_{loc}(\Omega)$ can be identified with a $u \in \mathcal{C}^m(\Omega)$.

2.2.3. Test functions and Sobolev spaces on manifolds

This part consists mainly of pages 97-108 of [Gri09], where many results are stated for weighted Riemannian manifolds. We do not need this generality, but want to point it out. Also, the following notation is used for uniform convergence of functions f_j to f with same domains, which is implicit in Lemma 2.4:

$$f_j \rightrightarrows f \;:\Leftrightarrow\; \forall \varepsilon > 0 \; \exists N > 0 : |f_j(x) - f(x)| \leq \varepsilon \text{ for all } j > N \text{ and } x \in \mathrm{dom}(f).$$

We will use the multi-index notation, where if $\alpha = (\alpha_1, \ldots, \alpha_n)^t \in \mathbb{N}_0^n$, then $|\alpha| = \alpha_1 + \ldots + \alpha_n$ and

$$\partial^\alpha f = \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n} f.$$

Given a smooth manifold M, not necessarily Riemannian, define $\mathcal{D}(M)$ to be the space $\mathcal{C}_0^\infty(M)$ with following mode of convergence: Let $f_j, f \in \mathcal{C}_0^\infty(M)$, we say that $f_j \xrightarrow{\mathcal{D}} f$ if

1. for any chart (U, φ) and multi-index α, we have $\partial^\alpha f_j \circ \varphi^{-1} \rightrightarrows \partial^\alpha f \circ \varphi^{-1}$ in $\varphi(U)$,
2. and there is a compact set $K \subset M$, such that all involved functions have support contained in K.

Elements of $\mathcal{D}(M)$ are called test functions, and one can show that they are dense in $L^p(M, \mu_g)$ for $1 \leq p < \infty$. Distributions on M are defined as elements of the dual space to $\mathcal{D}(M)$ with weak convergence and topology. Any $h \in L^1_{loc}(M, \mu_g)$ can be identified as distribution via

$$h(f) = (h, f)_{L^2} = \int_M hf \; \mathrm{d}\mu_g \quad \text{for} \quad f \in \mathcal{D}(M).$$

Taking derivatives of distributions is done via relegation to test functions, i.e. if $X \in TM$, then

$$(Xh, f)_{L^2} = -(h, Xf)_{L^2} \qquad \text{or} \qquad (\Delta h, f)_{L^2} = (h, \Delta_g f)_{L^2}.$$

If the stronger requirement holds, that for some $u, v \in L^2_{loc}(M, \mu_g)$ and all $f \in \mathcal{D}(M)$

$$(\Delta u, f)_{L^2} = (v, f)_{L^2},$$

then we say that $\Delta u = v$ holds weakly. We define similar notions for compactly supported, smooth vector fields v, w

$$(v, w)_{\vec{L}^2} = \int_M \langle v, w \rangle_g \; \mathrm{d}\mu_g,$$

so that the dual space of $\vec{\mathcal{D}}(M)$ are distributional vector fields. For a test function $f \in \mathcal{D}(M)$ and distributional vector field $\vec{h}$, we can define distributional divergence again via relegation

$$(-\mathrm{div}\vec{h}, f)_{L^2} = (\vec{h}, \nabla_g f)_{\vec{L}^2}.$$

The gradient of a distribution is defined analogously and shows that distributions and distributional vector fields are embedded in each other.

We denote with $\Gamma_0^\infty(M)$ the module of smooth, compactly supported tangent vector fields on M, and by $\vec{L}^2(M,\mu_g)$ we associate the vector space of tangent vectors v, with measurable coefficients in all charts, satisfying in addition $|v|_g \in L^2(M,\mu_g)$. Both are Hilbert spaces with the inner products given above. We similarly define spaces of functions having these properties only on compact subsets, and add the index *loc* to them.

Definition 2.2.2. Let $u \in L^2(M,\mu_g)$. A weak gradient for u is given by a vector field $v \in \vec{L}^2_{loc}(M,\mu_g)$, such that for all $\vec{h} \in \Gamma_0^\infty(M)$ we have

$$(-\mathrm{div}_g(\vec{h}), u)_{L^2} = (\vec{h}, v)_{\vec{L}^2}.$$

We will denote the weak gradient of u by ∇u, which still depends on the metric g.

A weak gradient need not exist for an arbitrary L^2-function, an example is given in [Eva10] on page 257 for M being the interval (0,2). Next we define Sobolev spaces.

Definition 2.2.3. Let

$$W^1(M,\mu_g) = \big\{u \in L^2(M,\mu_g) \ : \ \exists v \in \vec{L}^2(M,\mu_g) \text{ with } v = \nabla u\big\},$$

and an inner product that turns $W^1(M,\mu_g)$ into a Hilbert space is given by

$$(u,f)_{W^1} = (u,f)_{L^2} + (\nabla u, \nabla f)_{\vec{L}^2}.$$

In chart domains, this Sobolev space and the L^p spaces can be identified with the ones in Euclidean space. The norms will be comparable too.

Definition 2.2.4. Let $u \in W^1_{loc}(M,\mu_g)$ and $f \in L^2_{loc}(M,\mu_g)$. We say the equation $\Delta u = f$ is satisfied weakly, if for all $h \in \mathcal{D}(M)$ we have

$$(\nabla u, \nabla_g h)_{\vec{L}^2} = (f,h)_{L^2}. \tag{2.5}$$

The same result on regularity as in Corollary 2.2.5.1 holds in this case too.

Definition 2.2.5. For an open and bounded subset Ω of M, we have $\mathcal{C}_0^\infty(\Omega) \subset W^1(\Omega,\mu_g)$, and we define the set $W_0^1(\Omega,\mu_g)$ to be the closure of $\mathcal{C}_0^\infty(\Omega)$ in W^1-norm. A domain where Δ is a self-adjoint operator with respect to $(\ ,\ )_{L^2}$ is given by the Hilbert space

$$W_0^2(\Omega,\mu_g) = \big\{u \in W_0^1(\Omega,\mu_g) \ : \ \exists f \in L^2(\Omega,\mu_g) \text{ such that } (2.5)\big\} \subset L^2(\Omega,\mu_g).$$

Note that for $u, v \in W_0^2(\Omega,\mu_g)$, we have sequences $u_j, v_j \in \mathcal{C}_0^\infty(\Omega)$ converging to u, v respectively in W^1-norm, and

$$(\Delta u, v)_{L^2} = \lim_{j\to\infty} (\nabla u, \nabla_g v_j)_{\vec{L}^2} = (\nabla u, \nabla v)_{\vec{L}^2} = \lim_{j\to\infty} (\nabla_g u_j, \nabla v_j)_{\vec{L}^2} = (u, \Delta v)_{L^2}.$$

2.2.4. Fredholm theory on manifolds

We will now consider the general Dirichlet problem of finding $u \in W_0^1(\Omega, \mu_g)$ for open and bounded $\Omega \subset M$, such that

$$\Delta u + \alpha u = f,$$

where $\alpha \in \mathbb{R}$ and $f \in L^2(\Omega, \mu_g)$. A function $u \in W_0^1(\Omega, \mu_g)$ is called weak solution if for all $h \in \mathcal{C}_0^\infty(\Omega)$

$$(\nabla u, \nabla_g h)_{\vec{L}^2} + \alpha(u, h)_{L^2} = (f, h)_{L^2}. \tag{2.6}$$

Theorem 2.2.6. *For any fixed $\alpha > 0$ and $f \in L^2(\Omega, \mu_g)$, the above Dirichlet problem has a unique solution $u_f = u_f^\alpha \in W_0^1(\Omega, \mu_g)$. Thus, we can define a resolvent operator*

$$\begin{array}{rcl} R_\alpha : L^2(\Omega, \mu_g) & \rightarrow & L^2(\Omega, \mu_g) \\ f & \mapsto & u_f. \end{array}$$

It turns out that R_α is a bounded linear operator, that is positive definite and self-adjoint on the Hilbert space $L^2(\Omega, \mu_g)$.

Proof. We deal with a Hilbert space, so we have the Cauchy-Schwarz inequality at our disposal and for all $h \in W^1(\Omega, \mu_g)$

$$|(f, h)_{L^2}| \leq \|f\|_{L^2}\|h\|_{L^2} \leq \|f\|_{L^2}\|h\|_{W^1};$$

and since we can use the Cauchy-Schwarz inequality for each inner product $\langle\ ,\ \rangle_g$, we obtain

$$|(\nabla u, \nabla h)_{\vec{L}^2}| \leq (|\nabla u|_g, |\nabla h|_g)_{L^2} \leq \|u\|_{W^1}\|h\|_{W^1}.$$

This shows that the terms in (2.6) define bounded linear functionals on $W^1(\Omega, \mu_g)$, and if (2.6) holds for $h \in \mathcal{C}_0^\infty(\Omega)$, then it will hold by density in $W_0^1(\Omega, \mu_g)$. We define now an inner product on $W_0^1(\Omega, \mu_g)$ via

$$[u, h]_\alpha := (\nabla u, \nabla_g h)_{\vec{L}^2} + \alpha(u, h)_{L^2}.$$

Obviously we have

$$\min\{\alpha, 1\}\|u\|_{W^1}^2 \leq [u, u]_\alpha \leq \max\{\alpha, 1\}\|u\|_{W^1}^2,$$

thus making the induced norm equivalent to the one of W^1, so $(W_0^1(\Omega, \mu_g), [\ ,\ ]_\alpha)$ is a Hilbert space. Further, since $[\ ,\ ]_\alpha$ is equivalent to $\|\ \|_{W^1}$,

$$h \mapsto (f, h)_{L^2}$$

is a bounded linear functional on $W_0^1(\Omega, \mu_g)$ and by the Riesz Representation Theorem, there is a unique $u_f \in W_0^1(\Omega, \mu_g)$, such that

$$[u_f, h]_\alpha = (f, h)_{L^2};$$

but this is (2.6). Further, if we plug in u_f for h, we obtain

$$\alpha(u_f, u_f)_{L^2} \leq [u_f, u_f]_\alpha = (f, u_f)_{L^2} \leq \|f\|_{L^2} \|u_f\|_{L^2},$$

which yields

$$\alpha \|u_f\|_{L^2} \leq \|f\|_{L^2}.$$

It hence follows that

$$\|R_\alpha\| = \sup_{f \in L^2(\Omega) \setminus \{0\}} \frac{\|R_\alpha f\|_{L^2}}{\|f\|_{L^2}} \leq \frac{1}{\alpha} < \infty.$$

For $f, h \in L^2(\Omega)$ we set $u_f = R_\alpha f$ and $u_h = R_\alpha h$ with $u_f, u_h \in W_0^1(\Omega)$ so that

$$(h, u_f)_{L^2} = (\nabla u_f, \nabla u_h)_{\vec{L}^2} + \alpha(u_f, u_h)_{L^2} = (f, u_h)_{L^2}.$$

If we set $h = f$ in the above formula, we deduce that $(f, u_f) \geq 0$, positive definiteness follows. ■

Finally we apply this theorem to the Helmholtz Equation/Eigenvalue Problem

$$\Delta u = \lambda u.$$

The next statement is a summary of Theorem 4.6 and 10.13 of [Gri09] and Theorem 2.11 of [Gri01].

Theorem 2.2.7 (Eigenfunction Basis)**.** *The operator Δ is positive definite and self-adjoint with domain containing $W_0^1(\Omega, \mu_g)$, for all relatively compact $\Omega \subset M$. The spectrum is discrete and an unbounded sequence in $[0, \infty)$. There is an orthonormal basis of smooth eigenfunctions $\{\phi_j\}$ in $L^2(\Omega, \mu_g)$.*

Proof. We only sketch the proof. Regard following modification

$$\Delta u + u = (1 + \lambda)u,$$

so the problem fits the previous theorem, and a solution u_λ would satisfy $u_\lambda = R((1+\lambda)u_\lambda)$, where $R = R_1$, or equivalently $\frac{1}{1+\lambda}u_\lambda = R(u_\lambda)$. Since R is a compact operator on a Hilbert space, there is an orthonormal basis of $L^2(\Omega, \mu_g)$ consisting of eigenfunctions ϕ_j for R, and a bounded sequence of eigenvalues a_j which accumulate around 0. All the eigenvalues are positive, since R is positive definite. Thus ϕ_j will be an orthonormal eigenfunction for Δ too, and $\lambda_j = \frac{1}{a_j} - 1$ will be its eigenvalue, which again has to be positive, since Δ is a positive definite operator. By Corollary 2.2.5.1, all the ϕ_j are smooth. ■

2.2.5. Miscellaneous results on eigenvalues of Δ_g

The source for this section is [Can13] and [Ura93], both containing many results that have nothing to do with the work of this author, but neither could he resist to take them in. It is for instance a reasonable question how the eigenvalues of Δ_g change with the metric g, and indeed, they depend continuously on g, see page 94 of [Can13].

Also one might wonder what sequences can appear as eigenvalues of a Laplacian?

Theorem 2.2.8 ([Ver87])**.** *If M is a smooth, compact n-dimensional manifold without boundary and $n > 2$, then to each finite sequence $0 < \lambda_1 \leq \lambda_2 \leq \ldots \lambda_k$, there is a metric g on M, such that Δ_g has this sequence as initial eigenvalues.*

Note that for $n = 2$ there are bounds on the multiplicity of eigenvalues. On the other hand, the asymptotic behavior of the number of eigenvalues below a threshold μ

$$N(\mu) := \big|\{\lambda_k \ : \ \lambda_k < \mu\}\big|$$

can not be altered, as following result shows (for a proof see page 99 of [Can13]).

Theorem 2.2.9 (Weyl's Asymptotic Formula)**.** *Let (M, g) be a smooth, compact Riemannian n-manifold and $\{\lambda_j\}$ be the eigenvalues of Δ_g. Then, with ω_n the volume of the n-dimensional unit ball*

$$\lim_{\mu\to\infty} \frac{N(\mu)}{\mu^{n/2}} = \frac{\omega_n \, Vol(M)}{(2\pi)^n} \qquad \textit{and} \qquad \lim_{k\to\infty} \frac{\lambda_k}{k^{2/n}} = \frac{\sqrt{2\pi}}{\big(\omega_n \, Vol(M)\big)^{2/n}}.$$

Upper bounds on eigenvalues for immersed manifolds have been obtained in [CDS10], lower bounds for closed manifolds with Ricci curvature bounded from below in [HKP16]. Eigenvalues of the Laplace-Beltrami operator determine the volume, dimension and total scalar curvature of a compact Riemannian manifold without boundary, see [Min53]. Yet there are different shapes with exactly the same eigenvalues, see [RWP06].

2.3. Green Functions

We introduce Green functions, i.e. fundamental solutions to self-adjoint linear partial differential operators. Applications of Green functions in science and engineering, and how to obtain them in practice for some special domains in $\mathbb{R}^n$ can be found in chapter 4 of [Col+11]. The eigenfunction expansion is proven in Section 2.3.2.

2.3.1. Definition and properties

We follow chapter 4 of [Aub98].

Definition 2.3.1. Given a compact smooth Riemannian manifold (M, g) of dimension n, with volume $V = \int_M 1\mu_g$. A kernel $\mathcal{G}(p, q) : M^2 \to \mathbb{R} \cup \{\infty\}$ is called Green function for the Laplace-Beltrami operator if in the sense of distributions and for any fixed $p \in M$ we have

$$\Delta_g \mathcal{G}(p, \cdot) = \delta_p(\cdot) - \frac{1}{V}.$$

The notion of Green function on manifolds with boundary is a bit different. Existence and properties are contained in the following theorem, which we state for $n > 2$ only.

Theorem 2.3.1 ([Aub98] Thm. 4.13)**.** *Let M be a compact smooth Riemannian manifold of dimension n. Then there exists a Green function $\mathcal{G}$ for Δ_g with the following properties:*

1. *For all $f \in \mathcal{C}^2(M)$ we have*

$$f(p) = \frac{1}{V} \int_M f \ \mathrm{d}\mu_g + \int_M \mathcal{G}(p, q) \Delta_g f(q) \ \mathrm{d}\mu_g(q). \tag{2.7}$$

2. *$\mathcal{G}(p, q)$ is smooth on the product manifold M^2 minus the diagonal.*

3. *If $r = d_g(p, q)$ is the distance in the induced metric, then for $n > 2$ and some finite constant C we have*

$$|\mathcal{G}(p, q)| < \frac{C}{r^{n-2}}.$$

4. *Green functions are bounded from below.*

5. *The integral of $\mathcal{G}$ over M^2 with respect to $\mathrm{d}\mu_g \mathrm{d}\mu_g$ exists.*

6. *We have the symmetry relation $\mathcal{G}(p, q) = \mathcal{G}(q, p)$ for all $p, q \in M$.*

Green functions on non-compact manifolds need not exist, see [Gri83] for non-existence results, which are closely related to capacities of compact sets.

2.3.2. Eigenfunction expansion

It is unfortunate, but the author of this thesis could not find a reference for the eigenfunction expansion on manifolds. So it will be derived in this section for $n = 3$. To do this, we will use a result of [Coc72], which deals with L^2 kernels on intervals. Alternatively we refer to [Smi38], which is much shorter and more succinct than [Coc72] – but we need the results for compact manifolds, and indeed, on page 5 of Cochran's book we find a remark, that says it is possible to lift the theory of Fredholm integral equations from an interval to surfaces.

Note that these references rely on properties of integrals, Measure Theory and Hilbert space techniques, which clearly all are present on $L^2(M, \mu_g)$ for a compact Riemannian manifold (M, g). To show for instance Fubini's Theorem or Dominated Convergence, we can use a partition of unity and work in the product of two charts, where we use the result for Lebesgue measures on $\mathbb{R}^{2n}$.

Theorem 2.3.2. *Let (M, g) be a 3-dimensional, compact smooth Riemannian manifold and $\mathcal{G}$ be a Green function for Δ_g with zero mean, i.e. $\int_M \mathcal{G}(p,q)\ \mathrm{d}\mu_g(q) = 0$. Let $\{\lambda_j\}$ and $\{\phi_j\}$ for $j \in \mathbb{N}$ be the set of eigenvalues and eigenfunctions of Δ_g with $\phi_0 = 1$ and $\lambda_0 = 0$. Then*

$$\mathcal{G}(p,q) = \sum_{j=1}^{\infty} \frac{\phi_j(p)\phi_j(q)}{\lambda_j},$$

and the series converges in $L^2(M^2) = L^2(M^2, \mu_g \times \mu_g)$.

Proof. We will use Theorem 3.3-1 of [Coc72], or Theorem 5 of [Smi38] in our setting, and construct an integral kernel H that is symmetric, in $L^2(M^2)$ and closed, i.e. there is no $h \in L^2(M)$ with

$$I_H(h) = \int_M H(p,q)h(q)\ \mathrm{d}\mu_g(q) = 0,$$

unless $h = 0$. In this case H has an eigenfunction expansion as we desire for the Green function, which converges in L^2. We set $H(p,q) = \mathcal{G}(p,q) + V^{-1}$ where $V = \int_M 1\ \mathrm{d}\mu_g$. Then H is symmetric because $\mathcal{G}$ is, and since the eigenfunctions ϕ_j are smooth and have zero mean

$$\begin{aligned} I_H(\phi_j) &= \frac{1}{\lambda_j V}\int_M \phi_j(q)\ \mathrm{d}\mu_g(q) + \frac{\lambda_j}{\lambda_j}\int_M \mathcal{G}(p,q)\phi_j(q)\ \mathrm{d}\mu_g(q) \\ &= \frac{1}{\lambda_j V}\int_M \phi_j(q)\ \mathrm{d}\mu_g(q) + \frac{1}{\lambda_j}\int_M \mathcal{G}(p,q)\Delta_g\phi_j(q)\ \mathrm{d}\mu_g(q) = \frac{1}{\lambda_j}\phi_j(p), \end{aligned}$$

by Theorem 2.3.1. This also shows that H is closed, as $I_H(\phi_0) = V^{-1}$ and the $\{\phi_j\}_{j\in\mathbb{N}_0}$ form a complete orthonormal system. If H were in L^2, we would deduce that $\{\lambda_j^{-1}\}$ and $\{\phi_j\}$ for $j \in \mathbb{N}$ are eigenvalues and eigenfunctions for H too, completed by $\lambda_0 = \frac{1}{V}$, and thus

$$H(p,q) = \mathcal{G}(p,q) + \frac{1}{V} = \frac{1}{V} + \sum_{j=1}^{\infty}\frac{\phi_j(p)\phi_j(q)}{\lambda_j}.$$

We are hence left to show that $\mathcal{G} \in L^2(M^2)$. By Theorem 2.3.1, $\mathcal{G}^2$ will be integrable if this is true along a neighborhood of the diagonal of M^2. We can cover M with finitely many charts $\mathcal{V}_s = (V_s, \psi_s)$ as in Lemma 2.1.25. There is thus an $\varepsilon > 0$, depending on the charts $\mathcal{V}_s$, such that for each $p \in M$ we can find an index α, with $B_\varepsilon(\psi_\alpha(p)) \subset \psi_\alpha(V_\alpha)$. Let us define $\Gamma_s = \{p \in V_s \ : \ B_\varepsilon(\psi_s(p)) \subset \psi_s(V_s)\}$, then (Γ_s, ψ_s) will be charts that cover M. We will then use this ε-buffer when we integrate with respect to polar coordinates:

$$\begin{aligned}\int_{\Gamma_s^2} \mathcal{G}(p,q)^2 \,\mathrm{d}\mu_g(p)\mathrm{d}\mu_g(q) &= \int_{\psi_s(\Gamma_s)^2} \mathcal{G}(\psi^{-1}(x), \psi^{-1}(y))^2 \sqrt{\det(g_s^\psi)(x)} \sqrt{\det(g_s^\psi)(y)} \,\mathrm{d}x\mathrm{d}y \\ &\leq \int_{\psi_s(\Gamma_s)\times\psi_s(V_s)} \frac{C_s}{\|x-y\|^2} \mathrm{d}x\mathrm{d}y \leq C_s' \int_{\psi_s(\Gamma_s)} \int_0^\varepsilon \frac{r^2}{r^2} \,\mathrm{d}r\mathrm{d}x < \infty,\end{aligned}$$

where we used 2.) + 3.) of Theorem 2.3.1 which is reflected in the constants C_s, C_s'; we used that $0 < c \leq \det(g)$ is continuous on M, and that the change to polar coordinates brings a factor of r^{n-1}, canceling the $(r^{n-2})^2$ for $n = 3$. ■

The last part of the proof above can be slightly modified to obtain following corollary.

Corollary 2.3.2.1. *Let (M, g) be a compact, smooth Riemannian n-manifold and $\mathcal{G}$ as in Theorem 2.3.2. Then*

$$\mathcal{G} \in L^p(M^2, \mu_g) \quad \textit{for} \quad 1 \leq p < \frac{n}{n-2}.$$

A different approach to the eigenfunction expansion is to equate $\mathcal{G}$ to the resolvent operator from Fredholm Theory, which is used to deduce the complete orthonormal basis for Δ_g. This is done in Lemma 2.3 of [AZ98] and should yield the result for arbitrary n.

We also note that by equation (4), on page 88 of [Coc72] and our approach in this section, it follows that the maximal number N of linearly independent eigenfunctions of the Laplacian for the eigenvalue λ is bounded by

$$N \leq \lambda^2 \|\mathcal{G}\|^2.$$

In combination with Theorem 2.2.8, this shows that $\|\mathcal{G}\|^2$ for a zero mean Green function can become arbitrary large.

Part II.

The Articles

3. Approximation to uniform distribution in SO(3)

This chapter is based on the joint work with Carlos Beltrán, [BF20], and has been reproduced here with kind permission of E. Saff.

3.1. Introduction and Results

In this paper we study properties of a finite collection of randomly generated points in SO(3), the rotation group of 3-dimensional Euclidean space, sampled by a certain *determinantal point process* (dpp). It turns out that these points tend to be well distributed, a property that is important for discretization, integration and approximation. Our goal is not to compute actual collections of evenly distributed rotation matrices, but rather to provide a comparison tool that allows to decide the effectiveness of any given method.

If one is given an algorithm to generate finite (but arbitrarily large) collections of matrices, common methods to measure how well distributed these are include either calculating some discrete energy of them or looking at the speed of convergence of the counting measure towards the uniform measure. Most work in this direction has been done on spheres of various dimensions, see the monograph [BHS19] for a very complete survey of the state of the art of this question; the particular question of finding collections of spherical points with small energy was posed by Shub and Smale in [SS93] and is nowadays known as Smale's 7th problem [Sma98].

In order to extend part of the work done on spheres to the context of rotation matrices, we will obtain bounds on various energies for points generated through a certain dpp (technically speaking, a dpp is a counting measure where one identifies the measure with its set of atoms). In few words, such a process is obtained by taking a Hilbert space $\mathcal{H}(X)$ (usually $\mathcal{H}(X) = L^2(X, \mu)$) of an underlying measure space (X, μ) and an N-dimensional subspace $\mathrm{H} \subset \mathcal{H}(X)$, with projection kernel $\mathcal{K}$ onto H – then, under mild conditions on X, one is guaranteed almost surely the existence of such a process with N distinct points in X associated to $\mathcal{K}$.

The theory of those processes has been summarized in [Hou+09]; there one also finds a pseudo-code which samples points from any given dpp. A main feature of the underlying points is that they tend to repel each other, and hence have become the theoretical basis of construction of well-distributed points on various symmetric spaces, see for instance [AZ15; BE18; BMOC16; MOC18].

Since one can sometimes compute the expected value of the energy of points coming from these processes with high precision, they have been used as a tool to understand the asymptotic properties of the discrete energy in that context; and in particular, for even dimensional spheres with exception of the usual 2-sphere, the best known bounds for some energies have been proved using this approach.

We will employ the same method for SO(3), considering first the (discrete) *Riesz s-energy* for $A = \{\alpha_1, \ldots, \alpha_N\} \subset \mathrm{SO}(3)$:

$$E_R^s(A) := \sum_{j \neq k} \frac{1}{\|\alpha_j - \alpha_k\|_F^s},$$

with α_j being thought of as rotation matrices, $\|\cdot\|_F$ being the Frobenius or L^2-norm, and $s \in (0, \infty)$. In contrast to this, the *continuous Riesz s-energy* is given by replacing the double sum by the double integral over SO(3). We further set

$$\mathcal{E}_R^s(N) = \inf_{|A|=N} E_R^s(A).$$

The investigation of these sums is very popular and results usually describe the behavior of the first leading terms. This seems particularly interesting in case s equals the dimension, where we have following result.

Theorem 3.1.1. *If $N = \binom{2L+3}{3}$ for $L \in \mathbb{N}$, then the Riesz 3-energy satisfies*

$$12\sqrt{2}\,\pi\ \mathcal{E}_R^3(N) \ \leq\ N^2 \log(N) + \left(3\gamma + \log(8^2 \cdot 6) - \tfrac{21}{4}\right)N^2 + O(N^{5/3}\log(N)),$$

where γ is the Euler–Mascheroni constant.

The right-hand side is the expected value of the Riesz 3-energy with underlying points generated by a certain dpp. Now, given any particular method of generating finite point sets in SO(3), one can numerically compute their 3-energy and compare it to the value above to decide if the points are evenly distributed. This comparison would clearly rise in significance at the presence of lower bounds on the 3-energy. From [BHS19, Th. 9.5.4] we have

$$\lim_{N\to\infty} \frac{\mathcal{E}_R^3(N)}{N^2 \log N} = \frac{\beta_3}{Vol(\mathrm{SO}(3))} = \frac{\frac{4\pi}{3}}{16\sqrt{2}\pi^2} = \frac{1}{12\sqrt{2}\,\pi}.$$

Here β_3 is the volume of the unit ball in $\mathbb{R}^3$ and $Vol(\mathrm{SO}(3))$ is the volume, i.e. the Hausdorff measure, of SO(3) as a subset of $\mathbb{R}^{3\times 3} \equiv \mathbb{R}^9$, see [Hua63] for a computation of that volume. We can thus see that random points from our dpp give the correct order of the asymptotic. The first order asymptotics for other s–energies are also understood (see the Poppy Seed Bagel Theorem [BHS19, Th. 8.5.2] for $s > 3$ and the Fundamental Theorem [BHS19, Th. 4.2.2] for $s < 3$). However, we have not found estimates on the next order term for the minimal 3–energy on SO(3), which leads to the following open question.

Open Problem 3.1.2. *Find bounds on the second term asymptotics for $\mathcal{E}_R^3(N)$ or more generally for $\mathcal{E}_R^s(N)$.*

We now turn our attention to the Green energy, where we obtain bounds with the continuous Green energy as coefficient of the factor N^2 (zero in this case), and narrow the domain of the leading coefficient of the second term.

To recap, a Green function $\mathcal{G}_{\mathrm{L}}$ for a linear differential operator L is an integral kernel to produce solutions for inhomogeneous differential equations and is unique modulo Ker(L). In our case, we deal with the Laplace-Beltrami operator Δ_g, and note that $\mathrm{Ker}(\Delta_g)$ is the set of harmonic functions – which are just constants on a compact Riemannian manifold (M, g). We will construct $\mathcal{G} = \mathcal{G}_{\Delta_g}$ in such a way that it integrates to zero and speak of *the* Green function.

The (discrete) *Green energy* for $A = \{\alpha_1, \dots, \alpha_N\} \subset \mathrm{SO}(3)$ will be given by

$$E_{\mathcal{G}}(A) := \sum_{i \neq j} \mathcal{G}(\alpha_i, \alpha_j),$$

and we let

$$\mathcal{E}_{\mathcal{G}}(N) = \inf_{|A|=N} E_{\mathcal{G}}(A).$$

It is noteworthy that $\mathcal{G}(\alpha, \beta) d(\alpha, \beta) \approx 1$ for α close to β in geodesic distance $d(\cdot, \cdot)$, and a set of points with small Green energy is hence expected to be well-distributed, which is indeed the main result in [BCC19]: We know that if $\{\alpha_1, \dots, \alpha_N\}$ attains the minimal possible energy, then the associated discrete measure approaches the uniform distribution in SO(3) as $N \to \infty$. A set of points with small Green energy is also expected to be well-separated, see [Cri19].

Now, $\mathcal{G}(\cdot, \beta)$ is for any $\beta \in \mathrm{SO}(3)$ a zero mean function by definition, and if $\{\alpha_1, \dots, \alpha_N\}$ were simply chosen uniformly and independently in SO(3), then the expected value of the Green energy would equal 0, so in particular we have $\mathcal{E}_{\mathcal{G}}(N) \leq 0$. In this note we prove the following much stronger result.

Theorem 3.1.3. *If* $N = \binom{2L+3}{3}$ *for* $L \in \mathbb{N}$*, then*

$$-3\sqrt[3]{\pi} N^{4/3} + O(N) \leq \ \mathcal{E}_{\mathcal{G}}(N) \ \leq -4 \left(\tfrac{3}{4}\right)^{4/3} N^{4/3} + O(N).$$

The right-hand side is the expected value of the Green energy with underlying points generated by a dpp, and that is where we have the restriction for N, as the process is related to subspaces H that we can project onto. The lower bound is valid for all N.

As mentioned above, another classical measure of the distribution properties of $\{\alpha_1, \dots, \alpha_N\}$ is the speed of convergence to uniform measure, which can be understood by choosing some range sets $\{A_j\}_{j \in I}$ measurable w.r.t. Haar measure μ and investigating the behavior of

$$\sup_{j \in I} \Big| \#\{k \ : \ \alpha_k \in A_j\} - N\mu(A_j) \Big|$$

as N grows large. We will tackle this problem probabilistically, where we turn the count of points in A_j into a random variable.

In analogy to spherical caps on spheres, the range sets for SO(3) will be the balls $B(\alpha, 2\varepsilon) := \{\beta \in \mathrm{SO}(3)\ :\ \omega(\alpha^{-1}\beta) < 2\varepsilon\}$ for $\varepsilon \in \left(0, \frac{\pi}{2}\right)$ and $\omega(\cdot)$ being the rotation angle distance introduced in the following sections. For given random points $\{\alpha_1, \ldots, \alpha_N\}$ and fixed $\alpha \in \mathrm{SO}(3)$, we define random variables via characteristic functions

$$X^k_{\alpha,\varepsilon} = \chi_{B(\alpha,2\varepsilon)}(\alpha_k) \qquad \text{and} \qquad \eta_{\alpha,\varepsilon} = \sum_{k=1}^{N} X^k_{\alpha,\varepsilon}.$$

Now, for a collection of random uniform points chosen independently in SO(3), denoting by $\mathbb{1}$ the identity matrix in SO(3), we have

$$\mathbb{E}[\eta_{\alpha,\varepsilon}] = N\mu(B(\alpha, 2\varepsilon)) = N\mu(B(\mathbb{1}, 2\varepsilon)),$$

and the variance can also be computed from the independence of the points:

$$\mathrm{Var}[\eta_{\alpha,\varepsilon}] = \mathbb{E}[\eta^2_{\alpha,\varepsilon}] - \mathbb{E}[\eta_{\alpha,\varepsilon}]^2 = N\left(\mu(B(\mathbb{1}, 2\varepsilon)) - \mu(B(\mathbb{1}, 2\varepsilon))^2\right).$$

We are able to bound the variance of this quantity for our dpp, proving that it is much smaller than in the previous case.

Theorem 3.1.4. *Let $N = \binom{2L+3}{3}$ for $L \in \mathbb{N}$, and $\varepsilon \in \left(0, \frac{\pi}{2}\right)$ be fixed. Then the points generated by the dpp given in Lemma 3.2.3 satisfy*

$$\mathbb{E}[\eta_{\alpha,\varepsilon}] = N\mu(B(\alpha, 2\varepsilon)) = N\mu(B(\mathbb{1}, 2\varepsilon)),$$

and moreover

$$\mathrm{Var}(\eta_{\alpha,\varepsilon}) = O\Big(\frac{\varepsilon^2}{\cos(\varepsilon)}\Big) N^{2/3}\log(N).$$

From Theorem 3.1.4 and for any fixed ε, we then have by Chebyshev's inequality

$$\sup_{\alpha \in \mathrm{SO}(3)} \mathbb{P}\Big(\big|\eta_{\alpha,\varepsilon} - N\mu(B(\mathbb{1}, 2\varepsilon))\big| \geq T\Big) \leq \mathrm{Var}(\eta_{\alpha,\varepsilon})T^{-2};$$

for example, letting $T = N^{1/3}\log(N)$ and with some little arithmetic we obtain

$$\sup_{\alpha \in \mathrm{SO}(3)} \mathbb{P}\Big(\big|\tfrac{1}{N}\eta_{\alpha,\varepsilon} - \mu(B(\mathbb{1}, 2\varepsilon))\big| \geq \tfrac{\log(N)}{N^{2/3}}\Big) = O\big(\tfrac{1}{\log(N)}\big).$$

In other words, for large N the counting and Haar measures are very similar with high probability.

3.2. Introductory Concepts

In this section we collect some definitions and previous results that we will use and that intend to make this manuscript reasonably self-contained. Definitions of Chebyshev polynomials and alike are postponed to section 3.2.4.

3.2.1. Structure, distances and integration in $\mathrm{SO}(3)$

The special orthogonal group SO(3) is the compact Lie group of 3 by 3 orthogonal matrices over $\mathbb{R}$ that represent rotations in $\mathbb{R}^3$, i.e. with determinant equal to one. It is a 3 dimensional manifold and since it is naturally included in $\mathbb{R}^9$ it is customary to let it inherit its Riemannian submanifold structure.

Following [HS11], using Euler angles $(\varphi_1, \theta, \varphi_2) \in [0, 2\pi) \times [0, \pi] \times [0, 2\pi)$, every element $R \in \mathrm{SO}(3)$ can be decomposed as $R = s_z(\varphi_1) s_x(\theta) s_z(\varphi_2)$ where

$$s_z(\varphi_1) := \begin{pmatrix} \cos(\varphi_1) & -\sin(\varphi_1) & 0 \\ \sin(\varphi_1) & \cos(\varphi_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad s_x(\theta) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

are rotations around the z-axis and x-axis respectively. The normalized Haar measure (i.e. the unique left and right invariant probability measure in SO(3)) is given by $\mathrm{d}\mu(R) = \frac{1}{8\pi^2} \sin(\theta) \mathrm{d}\varphi_1 \mathrm{d}\theta \mathrm{d}\varphi_2$, and it corresponds to the inherited Riemannian submanifold structure of SO(3) up to the normalizing constant.

The Riemannian distance associated to the structure of SO(3) is certainly a natural and useful concept, but for us it will be more convenient to use the so called *rotation angle distance* defined as follows: for $\alpha, \beta \in \mathrm{SO}(3)$,

$$\omega(\alpha^{-1}\beta) = \arccos\left(\frac{\mathrm{Trace}(\alpha^{-1}\beta) - 1}{2}\right) \in [0, \pi].$$

Its convenience stems from the following fact, see for example [HS11, page 173]: Given a function $f \in L^1(\mathrm{SO}(3))$ such that we can find $\tilde{f} \in L^1([0, \pi])$ with $f(x) = \tilde{f}(\omega(x))$, then

$$\int_{\mathrm{SO}(3)} f(x) \, \mathrm{d}\mu(x) = \frac{2}{\pi} \int_0^\pi \tilde{f}(t) \sin^2\left(\tfrac{t}{2}\right) \, \mathrm{d}t. \tag{3.1}$$

By the Monotone Convergence Theorem, (3.1) is also valid if $f, \tilde{f}$ are just assumed to be non-negative and measurable.

3.2.2. Laplace-Beltrami operator and Green function in $\mathrm{SO}(3)$

The Laplace-Beltrami operator Δ_g is defined on any Riemannian manifold (M, g) in terms of the Levi-Civita connection. Following [Car92], if $\gamma_1(t), \ldots, \gamma_n(t)$ is a set of geodesics in an n-dimensional manifold such that $\gamma_j(0) = p \in \mathrm{M}$ for all $1 \leq j \leq n$, and such that $\{\dot{\gamma}_j(0)\}$ form an orthonormal basis of the tangent space $T_p\mathrm{M}$ (geodesic normal coordinates), then the action of Δ_g on C^2-functions f at p is given by

$$\Delta_g f(p) = -\sum_{j=1}^{n} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big|_{t=0} f(\gamma_j(t)).$$

Note the convention given by the minus sign in front of the sum, which sometimes leads to confusion given the Laplacian in $\mathbb{R}^n$. The convention we use here is widely accepted, see for example [Jos11]. A Green function $\mathcal{G} = \mathcal{G}_{\Delta_g}$ is a distributional solution to

$$\Delta_g \mathcal{G}(\cdot, y) = \delta(\cdot, y) - \frac{1}{\mu_{dV}(\mathrm{M})},$$

where $\mu_{dV}(\mathrm{M})$ is the Riemannian volume form in M. This way defined it is unique modulo $\mathrm{Ker}(\Delta_g)$ and it is common practice to add a constant in such a way that for all $y \in \mathrm{M}$ the function $\mathcal{G}(\cdot, y)$ has zero mean, see [Aub98]. We use this convention and simply refer to $\mathcal{G}$ as *the* Green function.

It further follows from classical Fredholm theory that

$$\mathcal{G}_{\Delta_g}(x, y) = \sum_{j=1}^{\infty} \frac{\phi_j(x)\bar{\phi}_j(y)}{\lambda_j}, \tag{3.2}$$

where $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$ is the sequence of eigenvalues for Δ_g and $\{\phi_j\}$, $j \geq 1$ is a complete orthonormal set of associated eigenfunctions. This is hence true locally on any smooth manifold.
In the case $\mathrm{M} = \mathrm{SO}(3)$, we obtain a Green function which is independent of any particular chart, thus valid globally. The eigenvalues and eigenfunctions of Δ_g are known from the classical theory of continuous groups and have been intensively studied in the physics literature, see [HS11; Jos73], [Wig59, §15]:

Lemma 3.2.1. *The eigenvalues of Δ_g in $\mathrm{SO}(3)$ are $\lambda_\ell = \ell(\ell+1)$ for $\ell \geq 0$. Moreover, if H_ℓ is the eigenspace associated to λ_ℓ, then the dimension of H_ℓ is $(2\ell+1)^2$ and an orthonormal basis of H_ℓ is given by $\sqrt{2\ell+1}\mathcal{D}^\ell_{m,n}$ where $-\ell \leq m, n \leq \ell$ and $\mathcal{D}^\ell_{m,n}$ are Wigner's $\mathcal{D}$-functions.*

The actual form of the Wigner $\mathcal{D}$-functions will not be important for us, since we will only use the fact that they constitute an orthogonal basis and the following summation formula:

$$\sum_{m=-\ell}^{\ell} \sum_{n=-\ell}^{\ell} \mathcal{D}^\ell_{m,n}(\alpha)\overline{\mathcal{D}^\ell_{m,n}(\beta)} = \mathcal{U}_{2\ell}\Big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\Big); \tag{3.3}$$

where $\mathcal{U}_{2\ell}(x)$ is the Chebyshev polynomial of second kind and degree 2ℓ, which will be briefly introduced in section 3.2.4. For more on formula (3.3) see [Jos73, Eq. 4.65] or [Vol10, pp. 40-41] for a nice summary. The following simple form for the Green function is derived in section 3.2.4, and to the best of our knowledge, this is the first time it has been formulated.

Lemma 3.2.2. *The Green function for the Laplace-Beltrami operator on $\mathrm{SO}(3)$ can be written in terms of the metric ω, i.e. for $\alpha, \beta \in \mathrm{SO}(3)$ with $\alpha \neq \beta$:*

$$\mathcal{G}(\alpha, \beta) = \big(\pi - \omega(\alpha^{-1}\beta)\big)\cot\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big) - 1.$$

3.2.3. Determinantal point processes

We point the reader to the excellent monograph [Hou+09] for an introduction to point processes, and we briefly summarize part of this material below. As in [BMOC16] and [BE18], we will use only a fraction of the theory.

A *simple point process* on a locally compact Polish space Λ with reference measure μ is a random, integer-valued positive Radon measure η, that almost surely assigns at most measure 1 to singletons – we shall think of it as a counting measure

$$\eta = \sum_{j=1} \delta_{x_j},$$

with $x_j \neq x_s$ for $j \neq s$. One usually identifies η with a discrete subset of Λ.

The *joint intensities* of η w.r.t. μ, if they exist, are functions $\rho_k : \Lambda^k \to [0, \infty)$ for $k > 0$, such that for pairwise disjoint sets $\{D_s\}_{s=1}^k \subset \Lambda$, the expected value of the product of number of points falling into D_s is given by

$$\mathbb{E}\left[\prod_{s=1}^k \eta(D_s)\right] = \int_{D_1 \times \ldots \times D_k} \rho_k(y_1, \ldots, y_k) \, \mathrm{d}\mu(y_1) \ldots \, \mathrm{d}\mu(y_k),$$

and $\rho_k(y_1, \ldots, y_k) = 0$ in case $y_j = y_s$ for some $j \neq s$.

A simple point process is *determinantal* with kernel $\mathcal{K}$ iff for every $k \in \mathbb{N}$ and all y_j's

$$\rho_k(y_1, \ldots, y_k) = \det\Big(\mathcal{K}(y_j, y_s)\Big)_{1 \leq j,s \leq k}.$$

Let (M, g) be a compact Riemannian manifold with measure $d\mu = \mu_{dV}$. Let $\mathrm{H} \subseteq L^2(\mathrm{M})$ be any N-dimensional subspace in the set of square-integrable functions. It follows from the Macchi-Soshnikov theorem [Hou+09, Thm. 4.5.5] that a simple point process with N points exists in M associated to H. An important property of that dpp is given by [Hou+09, Form. (1.2.2)]: For any measurable function $f : \mathrm{M} \times \mathrm{M} \to [0, \infty]$,

$$\mathbb{E}\left[\sum_{i \neq j} f(x_i, x_j)\right] = \iint_{\mathrm{M}} f(x, y)\Big(\mathcal{K}_{\mathrm{H}}(x, x)\mathcal{K}_{\mathrm{H}}(y, y) - |\mathcal{K}_{\mathrm{H}}(x, y)|^2\Big) \, \mathrm{d}\mu(x, y); \qquad (3.4)$$

where we write $\mathrm{d}\mu(x, y)$ as an abbreviation for $\mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$ and

$\mathbb{E}\big[g(x_1, \ldots, x_N)\big]$ means expected value of some function defined from $\mathrm{M} \times \cdots \times \mathrm{M}$ (N copies of M) to $[0, \infty]$, when $x_1, \ldots, x_N$ are chosen from the point process associated to H;

$\mathcal{K}_{\mathrm{H}}(x, y)$ is the (orthogonal) projection kernel on H, namely for any $f \in L^2(\mathrm{M})$ the orthogonal projection of f onto H can be computed via:

$$\Pi_{\mathrm{H}}(f)(x) = \int_{y \in \mathrm{M}} f(y)\mathcal{K}_{\mathrm{H}}(x, y) \, \mathrm{d}\mu(y) \ \in L^2(\mathrm{H}).$$

Note that if $\varphi_1, \dots, \varphi_N$ is an orthonormal basis of H, then we can write

$$\mathcal{K}_{\mathrm{H}}(x, y) = \sum_{j=1}^{N} \varphi_j(x)\overline{\varphi_j(y)}, \tag{3.5}$$

and clearly

$$\int_{\mathrm{SO}(3)} \mathcal{K}_{\mathrm{H}}(x, x) \, \mathrm{d}\mu(x) = N.$$

Coming back to the case of interest and following ideas in [BMOC16], we choose as subspace H the span of the first eigenspaces of Δ_g. Recall the definition of classical Gegenbauer polynomials $\mathcal{C}_n^{(\lambda)}(t)$, a sequence of degree n polynomials orthogonal w.r.t the weight $(1-t^2)^{\lambda-1/2}$ in $[-1, 1]$, normalized in such a way that

$$\mathcal{C}_n^{(\lambda)}(1) = \binom{2\lambda + n - 1}{2\lambda - 1}.$$

An equivalent definition of these polynomials is given by the formal power series

$$(1 - 2t\alpha + \alpha^2)^{-\lambda} = \sum_{n=0}^{\infty} \mathcal{C}_n^{(\lambda)}(t)\alpha^n.$$

Lemma 3.2.3. *Let $L \geq 0$ and $\mathrm{H}_L \subseteq L^2(\mathrm{SO}(3))$ be the span of the union of eigenspaces for eigenvalues $\lambda_0, \dots, \lambda_L$ of Δ_g. Then, we define*

$$N := \dim(\mathrm{H}_L) = \binom{2L+3}{3} = \mathcal{C}_{2L}^{(2)}(1) = \frac{4}{3}L^3 + O(L^2).$$

Moreover, the projection kernel is:

$$\mathcal{K}_L(\alpha, \beta) := \mathcal{K}_{\mathrm{H}_L}(\alpha, \beta) = \mathcal{C}_{2L}^{(2)}\Big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\Big).$$

We then consider the dpp associated to H_L.

3.2.4. Proofs of the basic lemmas

The degree $n+1$ Chebyshev polynomials of first and second kind satisfy the recurrence relation

$$P_{n+1}(x) = 2xP_n(x) - P_{n-1}(x), \tag{3.6}$$

with $\mathcal{T}_0 \equiv 1$, $\mathcal{T}_1(x) = x$ and $\mathcal{U}_{-1} \equiv 0$, $\mathcal{U}_0(x) \equiv 1$ in their respective notation. With this said, using (3.2), (3.3) and (3.5), we obtain

$$\mathcal{K}_L(\alpha, \beta) = \sum_{\ell=0}^{L} (2\ell + 1)\, \mathcal{U}_{2\ell}\Big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\Big); \tag{3.7}$$

$$\mathcal{G}(\alpha,\beta) = \sum_{\ell=1}^{\infty} \frac{2\ell+1}{\ell(\ell+1)} \mathcal{U}_{2\ell}\Big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\Big). \tag{3.8}$$

Further we list some equations for later reference and the reader's convenience.

$$\begin{array}{rll} 2\mathcal{T}_{2\ell+1}(x) & = \mathcal{U}_{2\ell+1}(x) - \mathcal{U}_{2\ell-1}(x) & \text{[AS72, Eq. 22.5.8]}, \\ \mathcal{T}_n(1) & = 1 & \text{[Gra+00, Eq. 8.944.1]}, \\ \frac{\mathrm{d}}{\mathrm{d}x}\mathcal{T}_{2\ell+1}(x) & = (2\ell+1)\,\mathcal{U}_{2\ell}(x) & \text{[Gra+00, Eq. 8.949.1]}, \\ \frac{\mathrm{d}}{\mathrm{d}x}\mathcal{U}_{2L+1}(x) & = 2\mathcal{C}^{(2)}_{2L}(x) & \text{[Gra+00, Eq. 8.949.4]}, \\ \mathcal{C}^{(\lambda)}_n(1) & = \binom{2\lambda+n-1}{2\lambda-1} & \text{[Gra+00, Eq. 8.937.4]}. \end{array} \tag{3.9}$$

Proof of Lemma 3.2.3. Let $y := \cos\big(\frac{\omega(\alpha^{-1}\beta)}{2}\big)$, then by (3.7) and (4.3)

$$\mathcal{K}_L(\alpha,\beta) = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{\ell=0}^{L}\mathcal{T}_{2\ell+1}(x)\Big|_y = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{2}\mathcal{U}_{2L+1}(x)\Big|_y = \mathcal{C}^{(2)}_{2L}(y).$$

The formula for the dimension of H_L can be proved as follows. The eigenspace associated to $\lambda_\ell = \ell(\ell+1)$ has dimension $(2\ell+1)^2$ since this is the number of elements of its basis $D^\ell_{m,n}$. Thus $\dim(\mathrm{H}_L)$ is given by $\sum_{\ell=0}^{L}(2\ell+1)^2 = \binom{2L+3}{3}$. ■

Proof of Lemma 3.2.2. In (3.8) we apply the equality

$$\mathcal{U}_{2\ell}(\cos(t)) = \frac{\sin\big((2\ell+1)t\big)}{\sin(t)} \qquad \text{[Gra+00, Eq. 8.940.1]},$$

and argue, under the assumption $w := \omega(\alpha^{-1}\beta) \in (0,\pi]$, as follows

$$\begin{aligned} \mathcal{G}(\alpha,\beta) &= \sum_{\ell=1}^{\infty}\frac{2\ell+1}{\ell(\ell+1)}\frac{\sin\big((2\ell+1)\frac{w}{2}\big)}{\sin\big(\frac{w}{2}\big)} \\ &= \frac{1}{\sin\big(\frac{w}{2}\big)}\sum_{\ell=1}^{\infty}\left(\frac{\sin\big((2\ell+1)\frac{w}{2}\big)}{\ell+1} + \frac{\sin\big((2\ell+1)\frac{w}{2}\big)}{\ell}\right) \\ &= \frac{1}{i}\Big(-\log\big(1-e^{iw}\big) + \log\big(1-e^{-iw}\big)\Big)\cot\big(\tfrac{w}{2}\big) - 1; \end{aligned}$$

where we used the well known fact, that the power series for $\log(1-x)$ at 1 converges at the boundary of its disc of convergence (except for $x = 1$) and equals the logarithm at these values:

$$\begin{aligned} \sum_{\ell=1}^{\infty}\frac{\sin\big((2\ell+1)\frac{w}{2}\big)}{\ell+1} &= \frac{1}{2i}\sum_{\ell=1}^{\infty}\frac{e^{i\frac{w}{2}(2\ell+1)} - e^{-i\frac{w}{2}(2\ell+1)}}{\ell+1} \\ &= \frac{e^{-i\frac{w}{2}}}{2i}\sum_{\ell=1}^{\infty}\frac{e^{iw(\ell+1)}}{\ell+1} - \frac{e^{i\frac{w}{2}}}{2i}\sum_{\ell=1}^{\infty}\frac{e^{-iw(\ell+1)}}{\ell+1} \\ &= \frac{-e^{-i\frac{w}{2}}}{2i}\big(\log\big(1-e^{iw}\big) + e^{iw}\big) + \frac{e^{i\frac{w}{2}}}{2i}\big(\log\big(1-e^{-iw}\big) + e^{-iw}\big) \\ &= \frac{-e^{-i\frac{w}{2}}}{2i}\log\big(1-e^{iw}\big) + \frac{e^{i\frac{w}{2}}}{2i}\log\big(1-e^{-iw}\big) - \sin(\tfrac{w}{2}), \end{aligned}$$

and similarly

$$\sum_{\ell=1}^{\infty} \frac{\sin\big((2\ell+1)\frac{w}{2}\big)}{\ell} = \frac{-e^{i\frac{w}{2}}}{2i} \log\big(1-e^{iw}\big) + \frac{e^{-i\frac{w}{2}}}{2i} \log\big(1-e^{-iw}\big).$$

Further, by $1 - e^{-iw} = 2ie^{-i\frac{w}{2}} \sin(\frac{w}{2})$, we conclude

$$\begin{aligned}
\log\big(1-e^{-iw}\big) - \log\big(1-e^{iw}\big) &= \log\big(2ie^{-i\frac{w}{2}} \sin(\tfrac{w}{2})\big) - \log\big(-2ie^{i\frac{w}{2}} \sin(\tfrac{w}{2})\big) \\
&= \log\big(2e^{i\frac{-w+\pi}{2}} \sin(\tfrac{w}{2})\big) - \log\big(2e^{i\frac{w-\pi}{2}} \sin(\tfrac{w}{2})\big) \\
&= (-w+\pi)\tfrac{i}{2} - (w-\pi)\tfrac{i}{2} = i(\pi - w),
\end{aligned}$$

where we used a property of the principal branch of the complex logarithm: $\log(re^{i\varphi}) = \log(r) + i\varphi$. ■

3.3. Riesz s-Energy: Proof of Theorem 3.1.1

Recall that if A is a real matrix, we have $\|A\|_F^2 := \mathrm{Trace}(A^tA)$. We set throughout $N = N(L) = \mathcal{C}_{2L}^{(2)}(1)$ for $L \in \mathbb{N}$.

Lemma 3.3.1. *For $\alpha, \beta \in \mathrm{SO}(3)$, we have $\|\alpha - \beta\|_F = \sqrt{8}\sin\big(\frac{\omega(\alpha^{-1}\beta)}{2}\big)$.*

Proof. We abbreviate $w = \omega(\alpha^{-1}\beta)$, and use the half-angle formula for sine:

$$\begin{aligned}
\|\alpha-\beta\|_F^2 &= \mathrm{Trace}\big[(\alpha-\beta)^t(\alpha-\beta)\big] = 6 - 2\mathrm{Trace}(\alpha^{-1}\beta) \\
&= 8\frac{2-\big(\mathrm{Trace}(\alpha^{-1}\beta)-1\big)}{4} = 8\frac{1-\cos(w)}{2} = 8\,\sin^2\big(\tfrac{w}{2}\big).
\end{aligned}$$ ■

Recall the definition of Euler's Beta function $\mathcal{B}(a,b) := \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t$ for $a, b > 0$. We are now ready to state our first proposition.

Proposition 3.3.2. *For $s \in (0,3)$ and $N = N(L) = \binom{2L+3}{3}$, we have*

$$\mathcal{E}_R^s(N) \;\le\; \tfrac{2}{8^{s/2}\pi}\mathcal{B}\big(\tfrac{3-s}{2}, \tfrac{1}{2}\big)N^2 + O(N^{1+s/3}).$$

If $s \in \{1,2\}$, we have more information on the term $O(N^{1+s/3})$: It is respectively

$$-\tfrac{\sqrt{2}}{\pi}\big(\tfrac{3}{4}\big)^{4/3}N^{4/3} + O(N) \quad \textit{and} \quad -\tfrac{4}{15}\big(\tfrac{3}{4}\big)^{5/3}N^{5/3} + O(N^{4/3}).$$

Proof. We use (3.4), Lemma 3.2.3, Lemma 3.3.1, invariance of Haar measure, and (3.1):

$$\begin{aligned}
&\iint_{\mathrm{SO}(3)} \frac{\mathcal{K}_L(\alpha,\alpha)^2 - \mathcal{K}_L(\alpha,\beta)^2}{\|\alpha-\beta\|_F^s}\,\mathrm{d}\mu(\alpha,\beta) \\
&= \iint_{\mathrm{SO}(3)} \frac{\big[\mathcal{C}_{2L}^{(2)}(1)\big]^2 - \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\frac{\omega(\alpha^{-1}\beta)}{2}\big)\big)\big]^2}{8^{\frac{s}{2}}\,\sin^s\big(\frac{\omega(\alpha^{-1}\beta)}{2}\big)}\,\mathrm{d}\mu(\alpha,\beta) \\
&= \frac{2}{8^{\frac{s}{2}}\pi}\int_0^{\pi}\Big(N^2 - \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{t}{2}\big)\big)\big]^2\Big)\sin^{2-s}\big(\tfrac{t}{2}\big)\,\mathrm{d}t \\
&= \frac{4}{8^{\frac{s}{2}}\pi}N^2\int_0^{\pi/2}\sin^{2-s}(t)\,\mathrm{d}t - \frac{4}{8^{\frac{s}{2}}\pi}\int_0^1\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2(1-t^2)^{\frac{1-s}{2}}\,\mathrm{d}t.
\end{aligned}$$

The next line is, apart of the factor $\frac{4}{8^{s/2}\pi}$, the continuous Riesz s-energy:

$$\int_0^{\pi/2} \sin^{2-s}(t)\ \mathrm{d}t = \int_0^1 \frac{t^{1-s}t}{\sqrt{1-t^2}}\ \mathrm{d}t = \frac{1}{2}\int_0^1 t^{\frac{1-s}{2}}(1-t)^{-1/2}\ \mathrm{d}t = \frac{1}{2}\mathcal{B}\big(\tfrac{3-s}{2},\tfrac{1}{2}\big).$$

On the other hand, for $0<s<3$ we have

$$\begin{aligned}\int_0^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2(1-t^2)^{\frac{1-s}{2}}\ \mathrm{d}t &= \int_0^{\pi/2} \big[C_{2L}^{(2)}(\cos(t))\big]^2 \sin^{2-s}(t)\ \mathrm{d}t\\ &\le \int_0^{1/L} \big[C_{2L}^{(2)}(\cos(t))\big]^2 t^{2-s}\ \mathrm{d}t + \int_{1/L}^{\pi/2} \big[C_{2L}^{(2)}(\cos(t))\big]^2 t^{2-s}\ \mathrm{d}t\\ &\le \big[C_{2L}^{(2)}(1)\big]^2 \frac{t^{3-s}}{3-s}\Big|_0^{1/L} - \frac{CL^2}{1+s}\frac{1}{t^{1+s}}\Big|_{1/L}^{\pi/2} = O(L^{3+s});\end{aligned}$$

where we have used that $|C_{2L}^{(2)}(t)| \le |C_{2L}^{(2)}(1)|$ for all $t\in[-1,1]$ and [Sze39, Eq. 7.33.6], i.e. for every $c>0$ there is $C\ge 0$ such that

$$|C_{2L}^{(2)}(\cos(\theta))| \le \frac{CL}{\theta^2}, \qquad \frac{c}{L}\le\theta\le\frac{\pi}{2}.$$

The case $s=1$ is Lemma 3.6.2; the case $s=2$ follows from Lemma 3.6.4:

$$\int_0^1 \frac{\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2}{\sqrt{1-t^2}}\ \mathrm{d}t = \int_0^{\pi/2} \big[\mathcal{C}_{2L}^{(2)}(\cos(t))\big]^2\ \mathrm{d}t = \frac{\pi}{2}\sum_{u=0}^{2L} c_{u,u} = \frac{8\pi}{15}L^5 + O(L^4),$$

where $c_{u,u} = c_{u,u}^{(2)}(2L)$ with notation as in Lemma 3.6.4. ■

In the next proof we use (3.1) and the digamma function ψ, see Section 3.6.

Proof of Theorem 3.1.1. We proceed as in the previous proof and use Lemma 3.6.4, in particular, we use the notation of that lemma for $c_{j,k} = c_{j,k}^{(2)}(2L)$:

$$\begin{aligned}\int_0^{\pi/2} \frac{\big[\mathcal{C}_{2L}^{(2)}(1)\big]^2 - \big[\mathcal{C}_{2L}^{(2)}\big(\cos(t)\big)\big]^2}{\sin(t)}\ \mathrm{d}t &= 2\sum_{r=1}^{2L}\int_0^{\pi/2} \frac{1-\cos(2rt)}{\sin(t)}\ \mathrm{d}t \sum_{u=0}^{2L-r} c_{r+u,u}\\ &= 4\sum_{r=1}^{2L}\int_0^{\pi/2} \big[\mathcal{U}_{r-1}\big(\cos(t)\big)\big]^2 \sin(t)\ \mathrm{d}t \sum_{u=0}^{2L-r} c_{r+u,u}\\ &= 4\sum_{r=1}^{2L}\int_0^1 \big[\mathcal{U}_{r-1}(t)\big]^2\ \mathrm{d}t \sum_{u=0}^{2L-r} c_{r+u,u} = (\star).\end{aligned}$$

We use (3.17) and obtain

$$(\star) = 2(\gamma+\log(4))\sum_{r=1}^{2L}\sum_{u=0}^{2L-r} c_{r+u,u} + 2\sum_{r=1}^{2L}\psi\big(r+\tfrac{1}{2}\big)\sum_{u=0}^{2L-r} c_{r+u,u} =: S_1 + S_2.$$

By $c_{r+u,u} = c^{(2)}_{r+u,u}(2L) = (r+u+1)(2L-r-u+1)(u+1)(2L-u+1)$, we have

$$\sum_{u=0}^{2L-r} c_{r+u,u} = \frac{16}{15}L^5 + \frac{2}{3}L^2r^3 - \frac{4}{3}L^3r^2 - \frac{r^5}{30} + O_{a+b<5}(L^a r^b),$$

and hence by well known formulas for the sum of powers of integers:

$$\begin{aligned} S_1 &= 2\big(\gamma + \log(4)\big)\left(\frac{16}{15}L^5 2L + \frac{2}{3}L^2 4L^4 - \frac{4}{3}L^3\frac{8}{3}L^3 - \frac{1}{30}\frac{32}{3}L^6\right) + O(L^5) \\ &= \frac{16}{9}\big(\gamma + \log(4)\big)L^6 + O(L^5). \end{aligned}$$

Invoking Lemma 3.3.3 yields

$$\begin{aligned} \frac{1}{2}S_2 &= \frac{16}{15}L^5\big(2L\ \psi(2L) - 2L\big) + \frac{2}{3}L^2\Big(\frac{(2L)^4}{4}\psi(2L) - \frac{(2L)^4}{4^2}\Big) \\ &\quad - \frac{4}{3}L^3\Big(\frac{(2L)^3}{3}\psi(2L) - \frac{(2L)^3}{3^2}\Big) - \frac{1}{30}\Big(\frac{(2L)^6}{6}\psi(2L) - \frac{(2L)^6}{6^2}\Big) \\ &\quad + O(L^5\log(L)) \\ &= \frac{8}{9}L^6\psi(2L) - \frac{14}{9}L^6 + O(L^5\log(L)). \end{aligned}$$

Since $N^2 = \mathcal{C}^{(2)}_{2L}(1)^2 = \frac{16}{9}L^6\big(1 + O(L^{-1})\big)$, and $\big(\frac{3}{4}N\big)^{1/3} = L\big(1+O(L^{-1})\big)^{1/6}$ we see

$$\frac{1}{3}\log\big(\tfrac{3}{4}N\big) = \log(L) + O(L^{-1});$$

and using harmonic numbers $H_n := \sum_{k=1}^n \frac{1}{k} = \log(n) + \gamma + O(n^{-1})$ which satisfy $\psi(2L) = H_{2L-1} - \gamma$, see [AS72, Eq. 6.3.2]:

$$\begin{aligned} (\star) &= \frac{16}{9}L^6\Big(\psi(2L) + \gamma + \log(4)\Big) - \frac{7}{4}\frac{16}{9}L^6 + O(L^5\log(L)) \\ &= N^2\Big(\log\big(2(3N/4)^{1/3} - 1\big) + \gamma + \log(4)\Big) - \frac{7}{4}N^2 + O(N^{5/3}\log(N)) \\ &= \frac{1}{3}N^2\log(N) + \frac{1}{3}\Big(3\gamma + \log\big(8^3\tfrac{3}{4}\big) - \frac{21}{4}\Big)N^2 + O(N^{5/3}\log(N)); \end{aligned}$$

proving the claim when multiplied by $\frac{4}{8^{3/2}\pi}$. ∎

Lemma 3.3.3. *Let $\psi(t)$ be the digamma function and $m \geq 0$, then*

$$\sum_{k=1}^n k^m\psi\big(k+\tfrac{1}{2}\big) = \frac{n^{m+1}}{m+1}\psi(n) - \frac{n^{m+1}}{(m+1)^2} + O(n^m\log(n)).$$

Proof. Since $\psi(t) = \log(t) + O(\frac{1}{t})$ for $t > 2$, we have

$$\sum_{k=1}^n k^m\psi\big(k+\tfrac{1}{2}\big) = \int_1^n t^m\log(t)\ \mathrm{d}t + O(n^m\log(n));$$

as the sum can be bounded from above and below by the same integral, apart from integration boundaries, where we obtain the error term. We finish by integrating: $\Big(\frac{t^{m+1}}{m+1}\log(t) - \frac{t^{m+1}}{(m+1)^2}\Big)\Big|_1^n$. ∎

3.4. Green Energy: Proof of Theorem 3.1.3

We prove the lower and upper bound separately in the following two sections.

3.4.1. Estimate of the Green Energy: Lower Bound

We follow an exposition due to N. Elkies, found in [Lan88, Lem. 5.2 pp. 149-154]. The results in [Lan88] are stated in detail for Riemann surfaces, i.e. one–dimensional complex manifolds, although it is mentioned that the argument can be extended to more general manifolds. Here we work out the details for SO(3).

The idea is to find a function with nice properties smaller than $\mathcal{G}$, and to bound its energy from below. For $\alpha, \beta \in \mathrm{SO}(3)$ and $t > 0$, we define:

$$\mathcal{G}_t(\alpha,\beta) = \sum_{\ell=1}^{\infty} e^{-\ell(\ell+1)t} \frac{2\ell+1}{\ell(\ell+1)} \mathcal{U}_{2\ell}\Big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\Big).$$

Quantitative estimates depend on asymptotics for this function.

Lemma 3.4.1 (N. Elkies)**.** *For all $t > 0$ and $\alpha \neq \beta$ we have*

$$\mathcal{G}(\alpha,\beta) \geq \mathcal{G}_t(\alpha,\beta) - t.$$

Proof. Using uniform convergence, we differentiate term by term and define

$$h_t(\alpha,\beta) := -\partial_t \mathcal{G}_t(\alpha,\beta) = \sum_{\ell=1}^{\infty} e^{-\ell(\ell+1)t}(2\ell+1) \sum_{m=-\ell}^{\ell}\sum_{n=-\ell}^{\ell} \mathcal{D}^{\ell}_{m,n}(\alpha)\overline{\mathcal{D}^{\ell}_{m,n}(\beta)}.$$

Given a smooth test function ϕ, with uniformly converging representation as $\sum_{\ell=0}^{\infty}\phi_\ell$, where $\phi_\ell = \sum_{m,n}\varphi^{\ell}_{m,n}\mathcal{D}^{\ell}_{m,n}\sqrt{2\ell+1}$, we set

$$u(\alpha,t) := \int_{\mathrm{SO}(3)} h_t(\alpha,\beta)\phi(\beta)\ \mathrm{d}\mu(\beta) = \sum_{\ell=1}^{\infty} e^{-\ell(\ell+1)t}\phi_\ell(\alpha),$$

where we interchanged integration and summation by uniform convergence and used that $\{\mathcal{D}^{\ell}_{m,n}\sqrt{2\ell+1}\}$ is an orthonormal basis. Now we have uniformly

$$\lim_{t\to 0} u(\alpha,t) = \phi(\alpha) - \int_{\mathrm{SO}(3)} \phi(\beta)\ \mathrm{d}\mu(\beta) = \phi(\alpha) - \phi_0.$$

For $t > 0$ fixed, we can interchange differentiation and integration yielding

$$\Delta_g u(\alpha,t) + \partial_t u(\alpha,t) = 0.$$

By the strong maximum principle (Theorem 3.7.2), we have for every $t > 0$:

$$\min_{\alpha\in\mathrm{SO}(3)} u(\alpha,t) \geq \min_{\alpha\in\mathrm{SO}(3)} u(\alpha,0).$$

The same PDE and estimates hold for

$$v(\alpha,t) = u(\alpha,t) + \phi_0.$$

If $\phi \geq 0$, then so is $v(\alpha,t)$ for all $t>0$ by the maximum principle as $v(\alpha,0) = \phi(\alpha)$. Hence

$$u(\alpha,t) = v(\alpha,t) - \phi_0 \geq -\phi_0 \quad \text{for } \phi \geq 0.$$

We further set

$$\mathbb{I}(\alpha,t) := \int_{\mathrm{SO}(3)} \mathcal{G}_t(\alpha,\beta)\phi(\beta)\, \mathrm{d}\mu(\beta) = \sum_{\ell=1}^{\infty} e^{-\ell(\ell+1)t} \frac{\phi_\ell(\alpha)}{\ell(\ell+1)},$$

where we interchanged sum and integral again. Differentiating term-wise for $t>0$ yields

$$\partial_t \mathbb{I}(\alpha,t) = -\sum_{\ell=1}^{\infty} e^{-\ell(\ell+1)t}\phi_\ell(\alpha) = -u(\alpha,t) \leq \phi_0 \quad \text{for } \phi \geq 0.$$

Finally, for fixed α let $t > \epsilon > 0$, then by the fundamental theorem of calculus:

$$\mathbb{I}(\alpha,t) - \mathbb{I}(\alpha,\epsilon) = \int_\epsilon^t -u(\alpha,s)\, \mathrm{d}s \leq \phi_0(t-\epsilon)$$

and thus, for all non-negative test functions ϕ

$$\int_{\mathrm{SO}(3)} \Big(\mathcal{G}_t(\alpha,\beta) - \mathcal{G}_\epsilon(\alpha,\beta) - (t-\epsilon)\Big)\phi(\beta)\, \mathrm{d}\mu(\beta) \leq 0.$$

Since the integrand is continuous, this proves that for $t > \epsilon$

$$\mathcal{G}_t(\alpha,\beta) - t \leq \mathcal{G}_\epsilon(\alpha,\beta) + \epsilon,$$

and for any fixed α, β with $\alpha \neq \beta$ taking the limit as $\epsilon \to 0$ proves the result. ■

Now by Lemma 3.4.1, we have for some $t > 0$ which will be determined later, and any collection of distinct points $\{\alpha_1, \ldots, \alpha_N\} \subset \mathrm{SO}(3)$:

$$\begin{aligned}
&\sum_{s\neq k}^{N} \mathcal{G}(\alpha_s,\alpha_k) + N(N-1)2t \geq \sum_{s\neq k}^{N} \mathcal{G}_{2t}(\alpha_s,\alpha_k) \\
&= \sum_{\ell=1}^{\infty} \frac{2\ell+1}{\ell(\ell+1)} \sum_{m=-\ell}^{\ell}\sum_{n=-\ell}^{\ell}\sum_{s\neq k}^{N} e^{-\ell(\ell+1)2t}\mathcal{D}^\ell_{m,n}(\alpha_s)\overline{\mathcal{D}^\ell_{m,n}(\alpha_k)} \quad = \\
&\sum_{\ell=1}^{\infty} \frac{2\ell+1}{\ell(\ell+1)} \sum_{m=-\ell}^{\ell}\sum_{n=-\ell}^{\ell} \left(\left| \sum_{k=1}^{N} e^{-\ell(\ell+1)t}\mathcal{D}^\ell_{m,n}(\alpha_k)\right|^2 - \sum_{k=1}^{N} e^{-\ell(\ell+1)2t}\left|\mathcal{D}^\ell_{m,n}(\alpha_k)\right|^2 \right) \\
&\geq -\sum_{\ell=1}^{\infty} \frac{2\ell+1}{\ell(\ell+1)} \sum_{m=-\ell}^{\ell}\sum_{n=-\ell}^{\ell}\sum_{k=1}^{N} e^{-\ell(\ell+1)2t}\left|\mathcal{D}^\ell_{m,n}(\alpha_k)\right|^2 = -N\mathcal{G}_{2t}(\alpha,\alpha).
\end{aligned}$$

Thus our remaining task is to find an asymptotic for $\mathcal{G}_t(\alpha,\alpha)$ in t. First we note that

$$\frac{e^{-\ell(\ell+1)t}}{\ell(\ell+1)} = 4\frac{e^{-\ell(\ell+1)t}}{(2\ell+1)^2}\Big(1+\frac{1}{4\ell(\ell+1)}\Big) = 4\frac{e^{-\ell(\ell+1)t}}{(2\ell+1)^2} + O(\ell^{-4}).$$

For $0 < t \ll 1$ we then obtain

$$\begin{aligned}
\mathcal{G}_t(\alpha,\alpha) &= \sum_{\ell=1}^{\infty} e^{-\ell(\ell+1)t}\frac{(2\ell+1)^2}{\ell(\ell+1)} = \sum_{\ell=1}^{\infty}\left(4e^{-\ell(\ell+1)t} + \frac{e^{-\ell(\ell+1)t}}{\ell(\ell+1)}\right)\\
&= 4e^{t/4}\int_0^\infty e^{-(2x+1)^2 t/4}\ \mathrm{d}x + O(1)\\
&= 2e^{t/4}\int_1^\infty e^{-x^2 t/4}\ \mathrm{d}x + O(1)\\
&= \frac{4e^{t/4}}{\sqrt{t}}\int_{\sqrt{t}/2}^\infty e^{-x^2}\ \mathrm{d}x + O(1)\\
&= \frac{4e^{t/4}}{\sqrt{t}}\int_0^\infty e^{-x^2}\ \mathrm{d}x + O(1) = 2\sqrt{\frac{\pi}{t}} + O(1).
\end{aligned} \tag{3.10}$$

If we choose $2t = \frac{\sqrt[3]{\pi}}{N^{2/3}}$, then by (3.10)

$$\mathcal{G}_{2t}(\alpha,\alpha) = 2\sqrt[3]{\pi}N^{\frac{1}{3}} + O(1),$$

and hence

$$\sum_{s\neq k}^{N}\mathcal{G}(\alpha_s,\alpha_k) \geq -3\sqrt[3]{\pi}N^{\frac{4}{3}} + O(N),$$

proving the lower bound in Theorem 3.1.3.

3.4.2. Estimate of the Green Energy: Upper Bound

According to (3.4), we have to estimate the integral

$$I = \iint_{\mathrm{SO}(3)} \mathcal{G}(\alpha,\beta)\left(\mathcal{K}_L(\alpha,\alpha)^2 - \mathcal{K}_L(\alpha,\beta)^2\right)\ \mathrm{d}\mu(\alpha,\beta),$$

which by Lemmas 3.2.2 and 3.2.3 and by invariance of Haar measure equals

$$\int_{\mathrm{SO}(3)}\Big(\big(\pi-\omega(\alpha)\big)\cot\big(\tfrac{\omega(\alpha)}{2}\big) - 1\Big)\left(\mathcal{C}_{2L}^{(2)}(1)^2 - \Big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{\omega(\alpha)}{2}\big)\big)\Big]^2\right)\ \mathrm{d}\mu(\alpha).$$

The integrand is in $L^1(\mathrm{SO}(3))$ since the singularity of the cotangent is removed by the zero of the difference of Gegenbauer polynomials, thus being a continuous function on a compact set. We can then apply (3.1) getting:

$$I = \frac{2}{\pi}\int_0^\pi \Big(\big(\pi - t\big)\cot\big(\tfrac{t}{2}\big) - 1\Big)\left(\mathcal{C}_{2L}^{(2)}(1)^2 - \Big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{t}{2}\big)\big)\Big]^2\right)\sin^2\big(\tfrac{t}{2}\big)\ \mathrm{d}t.$$

Since
$$\int_0^\pi \Big((\pi-t)\cot\left(\tfrac{t}{2}\right)-1\Big)\sin^2\left(\tfrac{t}{2}\right)\,\mathrm{d}t=0,$$
we indeed have
$$-I=\frac{2}{\pi}\int_0^\pi \Big((\pi-t)\cot\left(\tfrac{t}{2}\right)-1\Big)\Big[\mathcal{C}_{2L}^{(2)}\big(\cos\left(\tfrac{t}{2}\right)\big)\Big]^2\sin^2\left(\tfrac{t}{2}\right)\,\mathrm{d}t. \tag{3.11}$$
We simplify by noticing that
$$\begin{aligned}\int_0^\pi \Big[\mathcal{C}_{2L}^{(2)}\big(\cos\left(\tfrac{t}{2}\right)\big)\Big]^2\sin^2\left(\tfrac{t}{2}\right)\,\mathrm{d}t &= 2\int_0^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\sqrt{1-t^2}\,\mathrm{d}t\\ &=\int_{-1}^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\sqrt{1-t^2}\,\mathrm{d}t\\ &=\int_{-1}^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\sqrt{1-t^2}(1+t)\,\mathrm{d}t,\end{aligned}$$
where we used that odd functions integrate to zero over symmetric intervals. But
$$\int_{-1}^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\sqrt{1-t}(1+t)^{3/2}\,\mathrm{d}t=\frac{\pi}{2}\binom{2L+3}{2L}, \tag{3.12}$$
by the following equality, valid for $\nu>\frac{1}{2}$ and found in [Gra+00, Eq. 7.314]:
$$\int_{-1}^1 (1-x)^{\nu-\frac{3}{2}}(1+x)^{\nu-\frac{1}{2}}\big|\mathcal{C}_n^{(\nu)}(x)\big|^2\,\mathrm{d}x=\frac{\pi^{1/2}\Gamma(\nu-\frac{1}{2})\Gamma(2\nu+n)}{n!\Gamma(\nu)\Gamma(2\nu)}. \tag{3.13}$$
We have then proved that
$$\begin{aligned}-I&=\frac{2}{\pi}\int_0^\pi (\pi-t)\cot\left(\tfrac{t}{2}\right)\Big[\mathcal{C}_{2L}^{(2)}\big(\cos\left(\tfrac{t}{2}\right)\big)\Big]^2\sin^2\left(\tfrac{t}{2}\right)\,\mathrm{d}t+O(L^3)\\ &=\frac{4}{\pi}\int_0^1 \big(\pi-2\arccos(t)\big)t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t+O(L^3)\\ &=4\int_0^1 t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t-\frac{4}{\pi}\int_0^1 2\arccos(t)t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t+O(L^3).\end{aligned}$$
Next we use Lemma 3.6.1 and Lemma 3.6.2 in
$$\int_0^1 t^2\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t<\int_0^1 t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t<\int_0^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t,$$
and obtain
$$\int_0^1 t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\,\mathrm{d}t=L^4+O(L^3).$$
Finally we use
$$0\le 2\arccos(t)\le\pi\sqrt{1-t},\qquad \text{for } t\in[0,1]$$

so that, by (3.12),

$$\int_0^1 2\arccos(t)t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\ \mathrm{d}t < \int_0^1 \pi\sqrt{1-t}\,t\big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\ \mathrm{d}t$$
$$< \pi\int_{-1}^1 \big[\mathcal{C}_{2L}^{(2)}(t)\big]^2\sqrt{1-t}(1+t)^{3/2}\ \mathrm{d}t = O(L^3).$$

Hence

$$I = -4L^4 + O(L^3),$$

and the upper bound in Theorem 3.1.3 follows from $N = \frac{4}{3}L^3 + O(L^2)$.

3.5. Variance: Proof of Theorem 3.1.4

Let $A = B(\mathbb{1}, 2\varepsilon) \subseteq \mathrm{SO}(3)$ be as in the introduction, namely

$$A = \{\beta \in \mathrm{SO}(3) : \omega(\beta) < 2\varepsilon\} = \big\{\beta \in \mathrm{SO}(3)\ : \|\beta - \mathbb{1}\|_F < \sqrt{8}\sin(\varepsilon)\big\},$$

where the equality follows from Lemma 3.3.1. Note that by rotation invariance it suffices to study the variance of the random variable

$$\eta_A = \sum_{k=1}^N \chi_A(\alpha_k),$$

where $\{\alpha_1, \ldots, \alpha_N\}$ are generated by our dpp. The expected value of η_A satisfies $\mathbb{E}[\eta_A] = \mu(A)N$, and the variance of η_A is by definition (using $\chi_A(\alpha_k)^2 = \chi_A(\alpha_k)$):

$$\mathrm{Var}(\eta_A) = \mathbb{E}[\eta_A^2] - \mathbb{E}[\eta_A]^2 = \mathbb{E}\Bigg[\sum_{i\neq j}\chi_A(\alpha_i)\chi_A(\alpha_j)\Bigg] + \mu(A)N - \mu(A)^2N^2.$$

The expected value of the right-hand side equals by (3.4), (with $f(x, y) = \chi_A(x)\chi_A(y)$)

$$\iint_{\alpha,\beta\in A} \big[\mathcal{C}_{2L}^{(2)}(1)\big]^2 - \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\big)\big]^2\ \mathrm{d}\mu(\beta,\alpha) =$$
$$\mu(A)^2N^2 - \iint_{\alpha,\beta\in A} \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\big)\big]^2\ \mathrm{d}\mu(\beta,\alpha).$$

In other words, we have

$$\mathrm{Var}(\eta_A) = \mu(A)N - \iint_{\alpha,\beta\in A} \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\big)\big]^2\ \mathrm{d}\mu(\beta,\alpha),$$

and therefore, using invariance of the Haar measure, (3.1) and (3.12)

$$\begin{aligned}
\mathrm{Var}(\eta_A) - &\int_A\int_{A^c} \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{\omega(\alpha^{-1}\beta)}{2}\big)\big)\big]^2\ \mathrm{d}\mu(\beta)\ \mathrm{d}\mu(\alpha)\\
&= \mu(A)N - \int_{\mathrm{SO}(3)}\chi_A(\alpha)\int_{\mathrm{SO}(3)} \big[\mathcal{C}_{2L}^{(2)}\big(\cos\big(\tfrac{\omega(\beta)}{2}\big)\big)\big]^2\ \mathrm{d}\mu(\beta)\ \mathrm{d}\mu(\alpha)\\
&= \mu(A)N - \int_{\mathrm{SO}(3)}\chi_A(\alpha)N\ \mathrm{d}\mu(\alpha) = 0.
\end{aligned}$$

All in one we have proved the variance version of [RV07, Eq. 28]:

$$\mathrm{Var}(\eta_A) = \int_A \int_{A^c} \left[\mathcal{C}_{2L}^{(2)}\left(\cos\left(\tfrac{\omega(\alpha^{-1}\beta)}{2}\right)\right)\right]^2 \,\mathrm{d}\mu(\beta)\,\mathrm{d}\mu(\alpha).$$

Now, note that

$$A^c = \left\{\beta \in \mathrm{SO}(3)\ : \|\beta - \mathbb{1}\|_F \geq \sqrt{8}\sin(\varepsilon)\right\},$$

and by the triangle inequality $\|\beta - \mathbb{1}\|_F \leq \|\beta - \alpha\|_F + \|\mathbb{1} - \alpha\|_F$ for $\alpha \in A$, we see that

$$A^c \subset S_\alpha := \left\{\beta \in \mathrm{SO}(3)\ : \omega(\alpha^{-1}\beta) \geq f\big(\omega(\alpha)\big)\right\},$$

where $f\big(\omega(\alpha)\big) := 2\arcsin\left(\sin(\varepsilon) - \sin\left(\tfrac{\omega(\alpha)}{2}\right)\right)$. With the characteristic function χ_α of S_α, $\chi_\alpha(\beta) = \chi_{[f(\omega(\alpha)),\pi]}(\omega(\alpha^{-1}\beta))$, we integrate over SO(3) and use (3.1):

$$\begin{aligned}
\int \chi_\alpha(\beta)\left[\mathcal{C}_{2L}^{(2)}\left(\cos\left(\tfrac{\omega(\alpha^{-1}\beta)}{2}\right)\right)\right]^2 \mathrm{d}\mu(\beta) &= \int \chi_\alpha(\alpha\beta)\left[\mathcal{C}_{2L}^{(2)}\left(\cos\left(\tfrac{\omega(\beta)}{2}\right)\right)\right]^2 \mathrm{d}\mu(\beta)\\
&= \frac{2}{\pi}\int_{f(\omega(\alpha))}^{\pi} \left[\mathcal{C}_{2L}^{(2)}\left(\cos\left(\tfrac{t}{2}\right)\right)\right]^2 \sin^2\left(\tfrac{t}{2}\right)\,\mathrm{d}t\\
&= \frac{4}{\pi}\int_{\frac{f(\omega(\alpha))}{2}}^{\pi/2} \left[\mathcal{C}_{2L}^{(2)}\left(\cos(t)\right)\right]^2 \sin^2(t)\,\mathrm{d}t\\
&= \frac{4}{\pi}\int_0^{\cos(\frac{f(\omega(\alpha))}{2})} \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2 \sqrt{1-t^2}\,\mathrm{d}t.
\end{aligned}$$

Applying (3.1) one more time yields with $\chi_A(\beta) = \chi_{[0,2\varepsilon)}(\omega(\beta))$

$$\begin{aligned}
\mathrm{Var}(\eta_A) &\leq \int_{\mathrm{SO}(3)} \chi_A(\alpha) \int_{\mathrm{SO}(3)} \chi_\alpha(\beta)\mathcal{C}_{2L}^{(2)}\left(\cos\left(\tfrac{\omega(\alpha^{-1}\beta)}{2}\right)\right)^2 \mathrm{d}\mu(\beta)\,\mathrm{d}\mu(\alpha)\\
&= \frac{4}{\pi}\int_{\mathrm{SO}(3)} \chi_A(\alpha) \int_0^{\cos(\frac{f(\omega(\alpha))}{2})} \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2 \sqrt{1-t^2}\,\mathrm{d}t\,\mathrm{d}\mu(\alpha)\\
&= \frac{16}{\pi^2}\int_0^{\varepsilon} \sin^2(x) \int_0^{\sqrt{1-(\sin(\varepsilon)-\sin(x))^2}} \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2 \sqrt{1-t^2}\,\mathrm{d}t\,\mathrm{d}x\\
&= \frac{16}{\pi^2}\int_0^{\varepsilon} \sin^2(x) \int_{\cos(\varepsilon)}^{\sqrt{1-(\sin(\varepsilon)-\sin(x))^2}} \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2 \sqrt{1-t^2}\,\mathrm{d}t\,\mathrm{d}x\\
&\qquad + \frac{16}{\pi^2}\int_0^{\varepsilon} \sin^2(x) \int_0^{\cos(\varepsilon)} \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2 \sqrt{1-t^2}\,\mathrm{d}t\,\mathrm{d}x =: I_1 + I_2.
\end{aligned}$$

Next we change the order of integration, thus for $t \in [\cos(\varepsilon), 1]$, we integrate over $\{t\} \times [z(t), \varepsilon]$, where $z(t) := \arcsin\left(\sin(\varepsilon) - \sqrt{1-t^2}\right)$. We do this since $x \in [z(t), \varepsilon]$ implies $\sqrt{1-(\sin(\varepsilon)-\sin(x))^2} \in [t, 1]$. Thus

$$I_1 = \frac{16}{\pi^2}\int_{\cos(\varepsilon)}^1 \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2 \sqrt{1-t^2} \int_{z(t)}^{\varepsilon} \sin^2(x)\,\mathrm{d}x\,\mathrm{d}t.$$

Further, by a standard estimate and the mean value theorem, we get

$$\begin{aligned}\int_{z(t)}^{\varepsilon} \sin^2(x)\ \mathrm{d}x &\le \sin^2(\varepsilon)\Big(\arcsin\big(\sin(\varepsilon)\big) - \arcsin\big(\sin(\varepsilon) - \sqrt{1-t^2}\big)\Big) \\ &\le \sin^2(\varepsilon)\frac{\sqrt{1-t^2}}{\cos(\varepsilon)},\end{aligned}$$

and hence by Lemma 3.6.1

$$I_1 \le \frac{16\sin^2(\varepsilon)}{\pi^2\cos(\varepsilon)}\int_0^1 \left[\mathcal{C}_{2L}^{(2)}(t)\right]^2(1-t^2)\ \mathrm{d}t = \frac{\sin^2(\varepsilon)}{\cos(\varepsilon)}O(L^2\log(L)).$$

We now estimate I_2. Using $\sin(\varepsilon) = \sqrt{1-\cos^2(\varepsilon)} \le \sqrt{1-t^2}$ for $t \in [0, \cos(\varepsilon)]$, Lemma 3.6.1, and $\frac{\sin(x)}{\sin(\varepsilon)} \le 1$ yield

$$I_2 \le \frac{16}{\pi^2}\int_0^{\varepsilon}\sin^2(x)\int_0^{\cos(\varepsilon)}\left[\mathcal{C}_{2L}^{(2)}(t)\right]^2\sqrt{1-t^2}\frac{\sqrt{1-t^2}}{\sin(\varepsilon)}\ \mathrm{d}t\ \mathrm{d}x = \varepsilon^2 O(L^2\log(L)).$$

Theorem 3.1.4 is now proved.

3.6. The L^2–Norm of Gegenbauer Polynomials

First we recall the digamma function $\psi(x) := \frac{\mathrm{d}}{\mathrm{d}x}\log\left(\Gamma(x)\right)$ and its property:

$$\psi(n+\tfrac{1}{2}) = \sum_{k=1}^{n}\frac{2}{2k-1} - \gamma - \log(4), \text{ for } n \in \mathbb{N}, \tag{3.14}$$

see [AS72, Eq. 6.3.4], where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

Lemma 3.6.1. *The Gegenbauer polynomials $\mathcal{C}_{n-2}^{(2)}(x)$ satisfy*

$$\int_0^1 (x^2-1)\left[\mathcal{C}_{n-2}^{(2)}(x)\right]^2\ \mathrm{d}x = -\frac{2n^2-1}{16}\Big(\psi(n+\tfrac{1}{2}) + \gamma + \log(4)\Big) + \frac{n^2}{8}.$$

Lemma 3.6.2. *The Gegenbauer polynomials $\mathcal{C}_{n-2}^{(2)}(x)$ satisfy*

$$\int_0^1 \left[\mathcal{C}_{n-2}^{(2)}(x)\right]^2\ \mathrm{d}x = \frac{n^4}{16} + \frac{4n^2-1}{64}\Big(\psi(n+\tfrac{1}{2}) + \gamma + \log(4)\Big) - \frac{5}{32}n^2.$$

For the proofs, we need a result from [Det93], showing the following recursive formula for squares of Gegenbauer polynomials:

$$\left(\frac{n}{2\lambda}\right)^2\left[\mathcal{C}_n^{(\lambda)}(x)\right]^2 = \sum_{k=0}^{n-1}\frac{\lambda+k}{\lambda}\left[\mathcal{C}_k^{(\lambda)}(x)\right]^2 - (1-x^2)\left[\mathcal{C}_{n-1}^{(\lambda+1)}(x)\right]^2,$$

which, for $\lambda = 1$, i.e. Chebyshev polynomials of 2nd kind [Det93, Corollary 6.2], is

$$\frac{(n+1)^2}{4}\left[\mathcal{U}_{n+1}(x)\right]^2 - \sum_{k=0}^{n}(k+1)\left[\mathcal{U}_k(x)\right]^2 = (x^2-1)\left[\mathcal{C}_n^{(2)}(x)\right]^2. \tag{3.15}$$

Proof of Lemma 3.6.1. We will use a well known identity for $m \le n$:

$$\mathcal{U}_m(x)\mathcal{U}_n(x) = \sum_{k=0}^{m} \mathcal{U}_{n-m+2k}(x), \tag{3.16}$$

which follows by induction on m, starting and re-applying the recurrence (3.6). Using (3.16) with $m = n$ in (4.2) and integrating yields

$$\begin{aligned}
&\int_0^1 (x^2-1)\big[\mathcal{C}_n^{(2)}(x)\big]^2 \,\mathrm{d}x \\
&\quad = \frac{(n+1)^2}{4} \sum_{k=0}^{n+1} \int_0^1 \mathcal{U}_{2k}(x) \,\mathrm{d}x - \sum_{k=0}^{n} (k+1) \sum_{s=0}^{k} \int_0^1 \mathcal{U}_{2s}(x) \,\mathrm{d}x \\
&\quad = \frac{(n+1)^2}{4} \sum_{k=0}^{n+1} \frac{\mathcal{T}_{2k+1}(1) - \mathcal{T}_{2k+1}(0)}{2k+1} - \sum_{k=0}^{n} (k+1) \sum_{s=0}^{k} \frac{\mathcal{T}_{2s+1}(1) - \mathcal{T}_{2s+1}(0)}{2s+1} \\
&\quad = \frac{(n+1)^2}{4} \sum_{k=0}^{n+1} \frac{1}{2k+1} - \sum_{k=0}^{n}\sum_{s=0}^{k} \frac{k+1}{2s+1},
\end{aligned}$$

where we used (4.3) and that $\mathcal{T}_{2n+1}(x)$ is odd. By (3.14), we state for later use:

$$\int_0^1 [\mathcal{U}_n(x)]^2 \,\mathrm{d}x = \sum_{k=0}^{n} \frac{1}{2k+1} = \frac{1}{2}\Big(\psi(n+\tfrac{3}{2}) + \gamma + \log(4)\Big), \text{ for } n \in \mathbb{N}_0. \tag{3.17}$$

We continue

$$\begin{aligned}
\int_0^1 (x^2-1)\big[\mathcal{C}_n^{(2)}(x)\big]^2 \,\mathrm{d}x &= \frac{(n+1)^2}{8}\Big(\psi(n+\tfrac{5}{2}) + \gamma + \log(4)\Big) \\
&\quad - \sum_{k=0}^{n} \frac{k+1}{2}\psi(k+\tfrac{3}{2}) - (\gamma+\log(4))\frac{(n+2)(n+1)}{4} \\
&= \frac{(n+1)^2}{8}\psi(n+\tfrac{5}{2}) - \sum_{k=1}^{n+1}\frac{k}{2}\psi(k+\tfrac{1}{2}) - \frac{(n+3)(n+1)}{8}(\gamma+\log(4)).
\end{aligned}$$

Also, we find by induction:

$$\sum_{k=1}^{n} \frac{k}{2}\psi(k+\tfrac{1}{2}) = \frac{1}{16}\left[(2n+1)^2\psi(n+\tfrac{3}{2}) - 2(n+1)^2 + \gamma + \log(4)\right],$$

where we used the recurrence $\psi(z+1) = \psi(z) + \frac{1}{z}$, see [AS72, Eq. 6.3.5]. Thus

$$\begin{aligned}
\int_0^1 (x^2-1)\big[\mathcal{C}_{n-2}^{(2)}(x)\big]^2 \,\mathrm{d}x &= \frac{2(n-1)^2 - (2n-1)^2}{16}\psi(n+\tfrac{1}{2}) + \frac{n^2}{8} \\
&\quad - \frac{2(n+1)(n-1)+1}{16}(\gamma+\log(4)) \\
&= -\frac{2n^2-1}{16}\Big(\psi(n+\tfrac{1}{2}) + \gamma + \log(4)\Big) + \frac{n^2}{8},
\end{aligned}$$

finishing the proof. ■

The proof of Lemma 3.6.2 first needs some preparation.

Lemma 3.6.3. *Let $c_{j,k}$ for $j,k \in \{0,\dots,n\}$ be real numbers such that*

1. $c_{j,k} = c_{j+r,k+r}$ *for* $j+k = n-r$ *with* $r \in \{1,\dots,n\}$,
2. $c_{j,k} = c_{n-j,k}$ *for* $j \geq k$,
3. $c_{j,k} = c_{k,j}$.

Then for any function $f : \mathbb{N}_0 \to \mathbb{R}$, we have[1]

$$\sum_{j,k=0}^{n} c_{j,k} f(|j-k|) = \sum_{j,k=0}^{n} c_{j,k} f(|n-j-k|) = 2 \sum_{r=0}^{n}{}' f(r) \sum_{u=0}^{n-r} c_{r+u,u}. \tag{3.18}$$

Proof. We first fix some $r \in \{1,\dots,n\}$ and regard the second sum. Observe that for all tuples such that $j_i+k_i = n-r$ and $\hat{j}_i+\hat{k}_i = n+r$, we also have $|n-j_i-k_i| = |n-\hat{j}_i-\hat{k}_i| = r$. These tuples are listed in the following table:

i	1	2	$\dots$	$n-r+1$
j_i	0	1		$n-r$
k_i	$n-r$	$n-r-1$	$\dots$	0
$\hat{j}_i$	r	$r+1$		n
$\hat{k}_i$	n	$n-1$		r

So for all r, $(j_i,k_i) \mapsto (j_i+r, k_i+r) = (\hat{j}_i, \hat{k}_i)$ is a bijection with $c_{j_i,k_i} = c_{\hat{j}_i,\hat{k}_i}$ and

$$\sum_{j,k=0}^{n} c_{j,k} f(|n-j-k|) = 2 \sum_{\substack{j,k=0 \\ j+k<n}}^{n} c_{j,k} f(n-j-k) + f(0) \sum_{u=0}^{n} c_{n-u,u}.$$

The first sum of (3.18) can be restricted to $j > k$ when doubled, apart of the sum $f(0)\sum_{u=0}^{n} c_{u,u}$. Again, we list all tuples with $j_i - k_i = r = n - \hat{j}_i - \hat{k}_i$:

i	1	2	$\dots$	$n-r+1$
j_i	r	$r+1$		n
k_i	0	1	$\dots$	$n-r$
$\hat{j}_i$	$n-r$	$n-r-1$		0
$\hat{k}_i$	0	1		$n-r$

Similarly, $(j_i,k_i) \mapsto (n-j_i, k_i) = (\hat{j}_i, \hat{k}_i)$ is a bijection with $c_{j_i,k_i} = c_{\hat{j}_i,\hat{k}_i}$, and

$$\sum_{j>k=0}^{n} c_{j,k} f(j-k) = \sum_{\substack{j,k=0 \\ j+k<n}}^{n} c_{j,k} f(n-j-k).$$

Rewriting the first sum above via $j = r+u$ and $k = u$ for some $u \in \{0,\dots,n-r\}$ and using that $c_{n-u,u} = c_{u,u}$ finishes the argument. ■

[1] The apostrophe on the sum-symbol sigma means taking half the first term.

Requirement 2. in Lemma 3.6.3 is valid for all j, k. To see this, let $j < k$, then

$$c_{j,k} \overset{2.+3.}{=} c_{j,n-k} \overset{1.}{=} c_{j+(k-j),n-k+(k-j)} = c_{k,n-j} \overset{3.}{=} c_{n-j,k}.$$

Lemma 3.6.4. *Let $n, \lambda \in \mathbb{N}$ be fixed, $j, k \in \{0, \dots, n\}$ and define*

$$c_{j,k}^{(\lambda)} = c_{j,k}^{(\lambda)}(n) = \frac{1}{[\Gamma(\lambda)]^4} \frac{\Gamma(\lambda+j)\Gamma(\lambda+n-j)}{j!(n-j)!} \frac{\Gamma(\lambda+k)\Gamma(\lambda+n-k)}{k!(n-k)!},$$

then

$$\big[\mathcal{C}_n^{(\lambda)}\big(\cos(t)\big)\big]^2 = \sum_{u=0}^{n} c_{u,u}^{(\lambda)} + 2\sum_{r=1}^{n} \cos(2rt) \sum_{u=0}^{n-r} c_{r+u,u}^{(\lambda)}.$$

In particular,

$$\int_0^{\pi/2} [C_n^{(\lambda)}(\cos\varphi)]^2 \, d\varphi = \frac{\pi}{2} \sum_{u=0}^{n} c_{u,u}^{(\lambda)}.$$

Proof. We will use Lemma 3.6.3 with [Gra+00, Eq. 8.934]:

$$\mathcal{C}_n^{(\lambda)}(\cos(\varphi)) = \sum_{\substack{k,\ell=0 \\ k+\ell=n}}^{n} \frac{\Gamma(\lambda+k)\Gamma(\lambda+\ell)}{k!\ell![\Gamma(\lambda)]^2} \cos((k-\ell)\varphi), \tag{3.19}$$

in conjunction with the angle-sum and half-angle formula for cosine and sine:

$$\begin{aligned}
&\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 - \big[\mathcal{C}_n^{(\lambda)}\big(\cos(t)\big)\big]^2 = \sum_{j,k=0}^{n} c_{j,k}^{(\lambda)} \Big(1 - \cos\big((n-2j)t\big)\cos\big((n-2k)t\big)\Big) \\
&= \sum_{j,k=0}^{n} c_{j,k}^{(\lambda)} \frac{1}{2}\Big(1 - \cos\big((j-k)2t\big) + 1 - \cos\big((n-j-k)2t\big)\Big) \\
&= \sum_{j,k=0}^{n} c_{j,k}^{(\lambda)} \Big(\sin^2\big((j-k)t\big) + \sin^2\big((n-j-k)t\big)\Big) = 4\sum_{r=1}^{n} \sin^2(rt) \sum_{u=0}^{n-r} c_{r+u,u}^{(\lambda)}.
\end{aligned}$$

Hence, with $\big[\mathcal{C}_n^{(\lambda)}\big(\cos(t)\big)\big]^2 = \big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 - \Big(\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 - \big[\mathcal{C}_n^{(\lambda)}\big(\cos(t)\big)\big]^2\Big)$

$$\begin{aligned}
\big[\mathcal{C}_n^{(\lambda)}\big(\cos(t)\big)\big]^2 &= 2{\sum_{r=0}^{n}}' \sum_{u=0}^{n-r} c_{r+u,u}^{(\lambda)} - 4\sum_{r=1}^{n} \sin^2(rt) \sum_{u=0}^{n-r} c_{r+u,u}^{(\lambda)} \\
&= \sum_{u=0}^{n} c_{u,u}^{(\lambda)} + 2\sum_{r=1}^{n} \big(1 - 2\,\sin^2(rt)\big) \sum_{u=0}^{n-r} c_{r+u,u}^{(\lambda)},
\end{aligned}$$

and we finish using $1 - 2\,\sin^2(rt) = \cos(2rt)$. ■

Proof of Lemma 3.6.2. With the notation of Lemma 3.6.4, where $c_{j,k} = c^{(2)}_{j,k}(n-2)$:

$$\sum_{u=0}^{n-2-r} c_{r+u,u} = \sum_{u=1}^{n-1-r} (r+u)(n-u)u(n-r-u)$$
$$= \tfrac{4r^2-1}{120}\Big(r\big(5n^2-\tfrac{1}{4}\big) - r^3 - 5n(n^2-1)\Big) - \frac{2r}{64}(4n^2-1) + \binom{2n+2}{5}\frac{1}{8}.$$

Further, we see by induction $\sum_{r=1}^{n-2} \frac{1}{4r^2-1} = \frac{n-2}{2n-3}$, and thus by Lemma 3.6.4

$$\begin{aligned}\int_0^1 \big[\mathcal{C}^{(2)}_{n-2}(x)\big]^2 \,\mathrm{d}x &= \sum_{u=0}^{n-2} c_{u,u} + 2\sum_{r=1}^{n-2}\int_0^{\frac{\pi}{2}} \cos(2rt)\sin(t)\,\mathrm{d}t \sum_{u=0}^{n-2-r} c_{r+u,u} \\ &= \sum_{u=0}^{n-2} c_{u,u} - 2\sum_{r=1}^{n-2}\frac{1}{4r^2-1}\sum_{u=0}^{n-2-r} c_{r+u,u} \\ &= \frac{n^5-n}{30} - \frac{1}{60}\sum_{r=1}^{n-2}\Big(r\big(5n^2-\tfrac{1}{4}\big) - r^3 - 5n(n^2-1)\Big) \\ &\qquad + 2\sum_{r=1}^{n-2}\frac{1}{4r^2-1}\left(\frac{2r}{64}(4n^2-1) - \binom{2n+2}{5}\frac{1}{8}\right) \\ &= \frac{2n^4-5n^2}{32} + \sum_{r=0}^{n-1}\frac{1}{4r^2-1}\frac{2r-1}{32}(4n^2-1),\end{aligned}$$

as $\big(\frac{4n^2-1}{32} - \binom{2n+2}{5}\frac{1}{4}\big)\frac{n-2}{2n-3}$ has a simple form. Equation (3.17) finishes the proof. ■

3.7. The Strong Maximum Principle on Manifolds

We state the classical strong maximum principle Theorem 3.7.1 for open, bounded, and connected subsets $U \subset \mathbb{R}^n$, and regard second order parabolic partial differential operators $\mathrm{L} + \frac{\partial}{\partial t}$ acting on functions $C^2_1(U \times (0,T])$, i.e. twice differentiable with respect to spatial variables and once w.r.t. time. $T > 0$. A special case of this is extended in Theorem 3.7.2. We set for smooth coefficients:

$$\mathrm{L}u(x,t) = -\sum\nolimits_{i,j}^{n} a_{ij}(x,t)\tfrac{\partial}{\partial x_i}\tfrac{\partial}{\partial x_j}u(x,t) + \sum\nolimits_{j}^{n} b_j(x,t)\tfrac{\partial}{\partial x_j}u(x,t), \tag{3.20}$$

and without loss of generality, $a_{ij}(x,t) = a_{ji}(x,t)$.

Definition 3.7.1. $\mathrm{L} + \frac{\partial}{\partial t}$ is said to be uniformly parabolic if there is a $C > 0$, s.t.

$$\sum_{i,j} a_{ij}(x,t)\xi_i\xi_j \geq C\|\xi\|_2^2, \quad \text{where } \xi \in \mathbb{R}^n,\ (x,t) \in U \times (0,T]. \tag{3.21}$$

Theorem 3.7.1 (Thm. 11, page 396 of [Eva10]). *Let $u \in C_1^2(U \times (0,T]) \cap C(\bar{U} \times [0,T])$ be such that*

$$\mathrm{L}u + \frac{\partial}{\partial t} u = 0,$$

for $U \subset \mathbb{R}^n$ as above, $\mathrm{L} + \frac{\partial}{\partial t}$ uniformly parabolic, and L as in (3.20). If the maximum or minimum of u is attained at a point $(x_0, t_0) \in U \times (0,T]$, then u equals this value everywhere in $U \times [0, t_0]$.

Given a manifold M with or without boundary, we set $\mathrm{M}^\circ = \mathrm{M} \setminus \partial \mathrm{M}$, and for $x \in \mathrm{M}$, define M_x as the connected component of M containing x. Now, the next theorem should be known, but we haven't found a reference.

Theorem 3.7.2. *Let (M, g) be an n-dimensional (smooth) compact Riemannian manifold with or without boundary, not necessarily connected. Suppose $u \in C_1^2(\mathrm{M}^\circ \times (0,T]) \cap C(\mathrm{M} \times [0,T])$ satisfies for $(x,t) \in \mathrm{M}^\circ \times (0,T]$:*

$$\Delta_g u(x,t) + \frac{\partial}{\partial t} u(x,t) = 0.$$

If the maximum or minimum of u is attained at a point $(x_0, t_0) \in \mathrm{M}^\circ \times (0,T]$, then u equals this value everywhere in $\mathrm{M}_{x_0} \times [0, t_0]$. In particular, the maximum and minimum of u are attained in $\big(\partial \mathrm{M} \times [0,T]\big) \cup \big(\mathrm{M}^\circ \times \{0\}\big)$.

Proof. For every $\alpha \in \mathrm{M}^\circ$, there is an open neighborhood $U_\alpha \subset \mathrm{M}$ and a chart $\boldsymbol{x}_\alpha : U_\alpha \to B_\alpha \subset \mathbb{R}^n$, such that $\boldsymbol{x}_\alpha(U_\alpha)$ is an open ball B_α, and the local representation of Δ_g in U_α is of type (3.20), and satisfies (3.21) for $C = 1/2$. This follows from the fact that the Laplace-Beltrami operator at a point β in the interior can be written as the usual Laplacian at β, and by continuity of the coefficients, there is an open set of β where the inequality (3.21) is true for $C = 1/2$.

Assume there were a $t_0 > 0$ such that the maximum/minimum of u would be attained at (α, t_0). Writing Δ_g w.r.t. the chart $\boldsymbol{x}_\alpha$ as Δ_α, and regarding the equation

$$\Delta_\alpha u(\boldsymbol{x}_\alpha^{-1}(x), t) + \frac{\partial}{\partial t} u(\boldsymbol{x}_\alpha^{-1}(x), t) = 0,$$

in $B_\alpha \times (0,T]$, a neighborhood of $(\boldsymbol{x}_\alpha(\alpha), t_0)$, we deduce by Theorem 3.7.1 that $u(x,t) \equiv u(\alpha, t_0)$ for all $(x,t) \in B_\alpha \times [0, t_0]$.

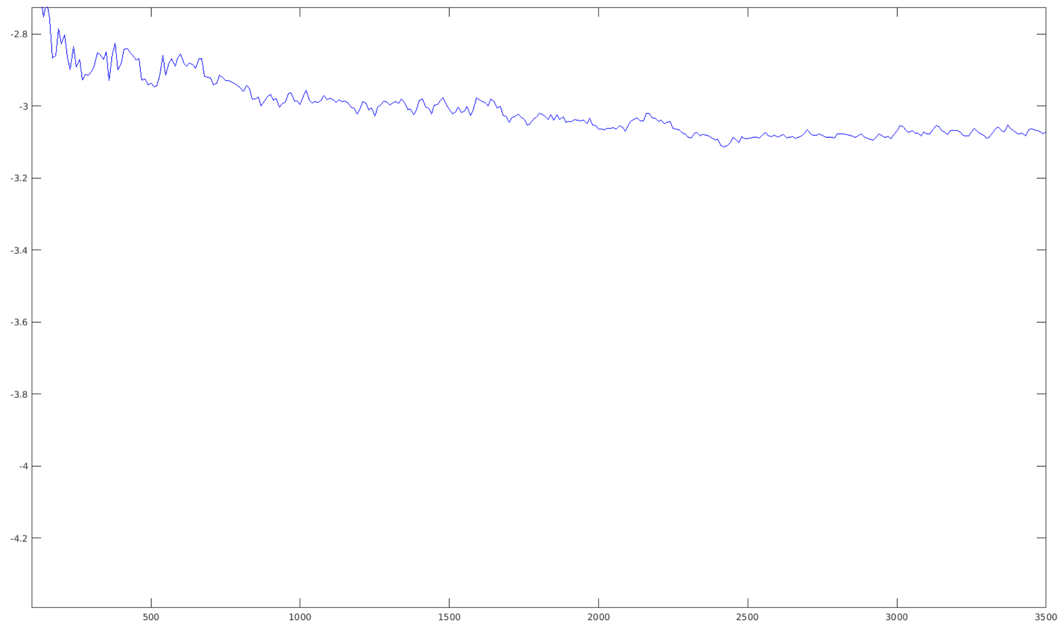
Further, M_α is covered by finitely many intersecting charts as above, and Theorem 3.7.1 would yield that u is constant and equals $u(\alpha, t_0)$ in all of $\mathrm{M}_\alpha \times [0, t_0]$. The maximum/minimum is in particular attained at the boundary as claimed. ■

3.8. Sampling on $\mathrm{SO}(3)$

This algorithm can still be found on the arXiv–version of [BF20] in Appendix C, and this section is based on that page. The theoretical upper bounds for the Green energy on SO(3) cannot be best possible, as it is an expected value – and hence there must be fluctuations above and in particular below that value.

We will introduce an algorithm to sample points in SO(3), that is simple to implement and numerically outperforms points sampled by a dpp. We are not giving any proofs regarding this algorithm, but rather show that it exists and how our bounds could be used as a comparison tool.

In 1987 a probabilistic algorithm was introduced by P. Diaconis and M. Shahshahani for compact groups in [DS87] and seemingly a special case of that was re-discovered by J. Arvo for SO(3) in [Arv12]. We will use a variant of this, replacing random points by a Halton sequence in the unit cube, which we baptize HArDiSh algorithm, and it does very well according to numerics.



The graphic shows the evolution of the Green energy divided by $N^{3/4}$ for HArDiSh – generated points, here $N = k * 10$ for $k \in \{10, \dots, 350\}$. The boundaries for the y-axis are chosen to be our theoretical bounds.

Following closely to [Arv12], we sample N points as follows: For x_1, x_2, x_3 to be determined later, let $M = -HR$ where $H = \mathbb{1} - 2vv^t$,

$$v = \frac{1}{\sqrt{N}} \begin{pmatrix} \cos(2\pi x_2)\sqrt{x_3} \\ \sin(2\pi x_2)\sqrt{x_3} \\ \sqrt{N - x_3} \end{pmatrix}, \text{ and } R = \begin{pmatrix} \cos(2\pi x_1) & \sin(2\pi x_1) & 0 \\ -\sin(2\pi x_1) & \cos(2\pi x_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.22}$$

In [Arv12], the x_j were chosen uniformly at random, and as Arvo already mentions, generating x_j by stratified or jittered sampling should yields less clumping for the matrices M. Our humble modification is to sample x_j via Halton sequences, i.e. let vdC(p, j)

denote the j-th element of the van der Corput seqence in base p, set

$$\mathcal{H} = \begin{pmatrix} 1/3 & 2/3 & 1/9 & 4/9 & 7/9 & \dots & \mathrm{vdC}(3,N) \\ 1/2 & 1/4 & 3/4 & 1/8 & 5/8 & \dots & \mathrm{vdC}(2,N) \\ 1 & 2 & 3 & 4 & 5 & \dots & N \end{pmatrix};$$

then we obtain matrices $\{M_k\}_{k=1}^N$ via (3.22) by setting $x_j(k) = \mathcal{H}(j,k)$. We do not know if the algorithm will continue to perform well for high numbers N.

3.9. Acknowledgment

Elkies' general argument for the lower bound of Green Energy (see Section 3.4.1) has been pointed out to the authors by E. Saff, and his help is thankfully acknowledged. We also thank the two anonymous referees for helpful comments.

4. On the L^2-norm of Gegenbauer polynomials

This chapter is the preprint version of [Fer21] with slight changes. We also take the opportunity to quickly recall the beta function: Let $a, b > \frac{1}{2}$, then

$$\mathcal{B}(a,b) := \int_0^1 t^{a-1}(1-t)^{b-1}\ \mathrm{d}t = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

In what follows we will give the leading terms of the asymptotic expansion of integrals as in [BF20, Cor. 4.1.3.1]. Note that computer algebra systems like Mathematica or Matlab were not able to numerically support these results at the time of this writing; but then again if one divides the integrals by the maximum of the integrand squared, a quickly growing sequence of numbers will emerge – a sign that the mentioned systems in fact cannot integrate Gegenbauer polynomials squared correctly.

4.1. Notation and Results

Gegenbauer polynomials $\mathcal{C}_n^{(\lambda)}$, where $\lambda \in I_G := (-\frac{1}{2}, 0) \cup (0, \infty)$ is called the index and $n \in \mathbb{N}_0$ is the degree, are the coefficients of following power series expansion in α:

$$(1 - 2x\alpha + \alpha^2)^{-\lambda} = \sum_{n=0}^{\infty} \mathcal{C}_n^{(\lambda)}(x)\alpha^n.$$

The case $\lambda = 0$ is not considered here. $\{\mathcal{C}_n^{(\lambda)}\}_{n\in\mathbb{N}_0}$ are orthogonal with respect to the measure $(1-x^2)^{\lambda-1/2}\ \mathrm{d}x$ over $[-1,1]$, and by [Gra+00, Eq. 8.930]:

$$\forall \lambda \in I_G : \qquad \mathcal{C}_0^{(\lambda)}(x) = 1, \qquad \mathcal{C}_1^{(\lambda)}(x) = 2\lambda x. \tag{4.1}$$

For continuous $f : [0,1] \to \mathbb{R}$, the following notation will be used:

$$\|f\|_2^2 := \int_0^1 [f(x)]^2\ \mathrm{d}x.$$

We derive an asymptotic formula for $\|\mathcal{C}_n^{(\lambda)}\|_2^2$ when $\lambda > 0$ in Corollary 4.1.3.1. Indeed, one of the key ingredients in [BF20] was the asymptotic nature of $\|\mathcal{C}_n^{(2)}\|_2^2$ in n, and the following lemma was proved in [BF20, Lemmas 6.1 and 6.2]:

Lemma 4.1.1. *Let ψ denote the digamma function and γ the Euler-Mascheroni constant. Then the Gegenbauer polynomials satisfy for $n \geq 2$:*

$$\begin{aligned} \big\| \sqrt{1-x^2}\ \mathcal{C}_{n-2}^{(2)} \big\|_2^2 &= \tfrac{1}{16}(2n^2-1)\big(\psi(n+\tfrac{1}{2})+\gamma+\log(4)\big) - \tfrac{1}{8}n^2, \\ \big\| \mathcal{C}_{n-2}^{(2)} \big\|_2^2 &= \tfrac{1}{16}n^4 + \tfrac{1}{64}(4n^2-1)\big(\psi(n+\tfrac{1}{2})+\gamma+\log(4)\big) - \tfrac{5}{32}n^2. \end{aligned}$$

The following result of Corollary 5.2 from [Det93] will prove to be indispensable.

Theorem 4.1.2 (Dette [Det93])**.** *The Gegenbauer polynomials satisfy for $\lambda \in I_G$*

$$\Big(\frac{n}{2\lambda}\Big)^2 \big[\mathcal{C}_n^{(\lambda)}(x)\big]^2 + (1-x^2)\big[\mathcal{C}_{n-1}^{(\lambda+1)}(x)\big]^2 = \sum_{k=0}^{n-1} \frac{\lambda+k}{\lambda} \big[\mathcal{C}_k^{(\lambda)}(x)\big]^2. \tag{4.2}$$

Our main theorem is as follows and we will use it to derive the asymptotic behavior of $\|\mathcal{C}_n^{(\lambda)}\|_2^2$.

Theorem 4.1.3 (Main Result)**.** *The Gegenbauer polynomials satisfy for $\lambda \in I_G$ and $n > 1$:*

$$\|\mathcal{C}_{n-2}^{(\lambda+1)}\|_2^2 = \frac{n^2-2\lambda n}{2^4\lambda^3}\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 + \frac{n(2n+1)}{2^3\lambda^2}\|\mathcal{C}_n^{(\lambda)}\|_2^2 - \sum_{k=0}^{n-1}\frac{\lambda+k}{2^2\lambda^2}\|\mathcal{C}_k^{(\lambda)}\|_2^2.$$

Corollary 4.1.3.1. *Let $\mathcal{B}(x,y)$ denote the beta function. The following asymptotic formulas in n hold for $\lambda \in (0,1)$ and $\delta(\lambda) := \max\{4\lambda-1, 2\lambda\}$:*

$$\begin{aligned} \|\mathcal{C}_n^{(\lambda)}\|_2^2 &< \ \mathcal{B}\big(1-\lambda, \tfrac{1}{2}\big)\frac{2^{1-2\lambda}}{\Gamma(\lambda)^2}\frac{1}{n^{2-2\lambda}}, \\ \|\mathcal{C}_n^{(\lambda+1)}\|_2^2 &= \frac{n^{4\lambda}}{4\lambda\Gamma(2\lambda+1)^2} + O\big(n^{\delta(\lambda)}\big), \\ \big\|\sqrt{1-x^2}\ \mathcal{C}_{n-1}^{(\lambda+1)}\big\|_2^2 &< \frac{\mathcal{B}\big(1-\lambda,\tfrac{1}{2}\big)}{\Gamma(\lambda+1)^2}\frac{n^{2\lambda}}{2^{1+2\lambda}} + O\big(n^{\delta(\lambda)-1}\big). \end{aligned}$$

The following asymptotic formulas hold for $\lambda > 1$:

$$\begin{aligned} \|\mathcal{C}_{n-2}^{(\lambda+1)}\|_2^2 &= \frac{n^{4\lambda}}{4\lambda\Gamma(2\lambda+1)^2} + \frac{\lambda-1}{\Gamma(2\lambda+1)^2}n^{4\lambda-1} + O(n^{4\lambda-2}), \\ \big\|\sqrt{1-x^2}\ \mathcal{C}_{n-1}^{(\lambda+1)}\big\|_2^2 &= \frac{2\lambda-1}{4(\lambda-1)\Gamma(2\lambda+1)^2}n^{4\lambda-2} + O(n^{\delta(\lambda-1)+2}). \end{aligned}$$

The identity $2\cdot\|\mathcal{C}_n^{(1)}\|_2^2 = \psi(n+\frac{3}{2})+\gamma+\log(4)$ is given by [BF20, Eq. 14].

4.2. Ingredients for the Proof of the Theorem

In this section we collect known results concerning Gegenbauer polynomials for later reference and the reader's convenience, and we derive some technical lemmas in Section 4.3 to prove Theorem 4.1.3. To avoid repetition, we will assume $\lambda \in I_G$ for the rest of the text if not stated otherwise. Note first that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\mathcal{C}^{(\lambda)}_{n+1}(x) &= 2\lambda\, \mathcal{C}^{(\lambda+1)}_{n}(x) && \text{[Gra+00, Eq. 8.935]}, \\ \mathcal{C}^{(\lambda)}_{n}(1) &= \tfrac{\Gamma(n+2\lambda)}{\Gamma(2\lambda)n!} = \frac{\prod_{j=1}^{n}(2\lambda+n-j)}{n!} && \text{[Gra+00, Eq. 8.937]}; \end{aligned} \tag{4.3}$$

and $\mathcal{C}^{(\lambda)}_{n}(1)$ is the maximum on $[-1,1]$ for $\lambda > 0$ by [Sze39, Eq. 7.33.1]. Also, by (4.3):

$$(n+2)\, \mathcal{C}^{(\lambda)}_{n+2}(x) = 2\lambda\Big(x\, \mathcal{C}^{(\lambda+1)}_{n+1}(x) - \mathcal{C}^{(\lambda+1)}_{n}(x)\Big) \qquad \text{[Gra+00, Eq. 8.933.2]}, \tag{4.4}$$

$$(n+\lambda)\, \mathcal{C}^{(\lambda)}_{n}(x) = \lambda\Big(\mathcal{C}^{(\lambda+1)}_{n}(x) - \mathcal{C}^{(\lambda+1)}_{n-2}(x)\Big) \qquad \text{[Gra+00, Eq. 8.939.6]}. \tag{4.5}$$

4.3. Identities for Gegenbauer polynomials

Lemma 4.3.1. *The Gegenbauer polynomials satisfy the following identities:*

$$\mathcal{C}^{(\lambda+1)}_{n}(x) + \mathcal{C}^{(\lambda+1)}_{n-2}(x) = 2x\, \mathcal{C}^{(\lambda+1)}_{n-1}(x) + \mathcal{C}^{(\lambda)}_{n}(x), \tag{$\star$}$$

$$\int_0^1 \big[\mathcal{C}^{(\lambda+1)}_{n}(x)\big]^2 - \big[\mathcal{C}^{(\lambda+1)}_{n-2}(x)\big]^2\, \mathrm{d}x = \frac{n+\lambda}{2\lambda^2}\Big(\big[\mathcal{C}^{(\lambda)}_{n}(1)\big]^2 + (2\lambda-1)\big\|\mathcal{C}^{(\lambda)}_{n}\big\|_2^2\Big).$$

Proof. First we use (4.5); then apply (4.4) to the right-hand side below proving ($\star$):

$$\mathcal{C}^{(\ell)}_{n}(x) + \mathcal{C}^{(\ell)}_{n-2}(x) = \frac{n+\lambda}{\lambda}\mathcal{C}^{(\lambda)}_{n}(x) + 2x\, \mathcal{C}^{(\ell)}_{n-1}(x) - \Big(2x\, \mathcal{C}^{(\ell)}_{n-1}(x) - 2\mathcal{C}^{(\ell)}_{n-2}(x)\Big),$$

where $\ell := \lambda + 1$. Next we obtain by the binomial theorem with (4.5), ($\star$) and (4.3)

$$\begin{aligned} \big[\mathcal{C}^{(\lambda+1)}_{n}(x)\big]^2 - \big[\mathcal{C}^{(\lambda+1)}_{n-2}(x)\big]^2 &= \frac{n+\lambda}{\lambda}\mathcal{C}^{(\lambda)}_{n}(x)\Big(2x\mathcal{C}^{(\lambda+1)}_{n-1}(x) + \mathcal{C}^{(\lambda)}_{n}(x)\Big) \\ &= \frac{n+\lambda}{\lambda}\Big(\frac{x}{2\lambda}\frac{\mathrm{d}}{\mathrm{d}x}\big[\mathcal{C}^{(\lambda)}_{n}(x)\big]^2 + \big[\mathcal{C}^{(\lambda)}_{n}(x)\big]^2\Big). \end{aligned}$$

Integration by parts then finishes the argument. ∎

Lemma 4.3.2. *The Gegenbauer polynomials satisfy the following identity:*

$$\begin{aligned} &\int_0^1 x^2\big[\mathcal{C}^{(\lambda+1)}_{n+1}(x)\big]^2 + \big[\mathcal{C}^{(\lambda+1)}_{n}(x)\big]^2\, \mathrm{d}x + \frac{1}{2\lambda}\int_0^1 (1-x^2)\big[\mathcal{C}^{(\lambda+1)}_{n+1}(x)\big]^2\, \mathrm{d}x \\ &\qquad = \frac{(n+2)^2}{8\lambda^3}\Big[\mathcal{C}^{(\lambda)}_{n+2}(1)\Big]^2 + \frac{2\lambda-1}{2\lambda}\frac{(n+2)^2}{4\lambda^2}\big\|\mathcal{C}^{(\lambda)}_{n+2}\big\|_2^2. \end{aligned}$$

Proof. Let $n = 2m$. By Lemma 4.3.1 and a telescoping sum argument:

$$\|\mathcal{C}_n^{(\lambda+1)}\|_2^2 - \|\mathcal{C}_0^{(\lambda+1)}\|_2^2 = \sum_{j=1}^{m} \frac{2j+\lambda}{2\lambda^2}\Big(\big[\mathcal{C}_{2j}^{(\lambda)}(1)\big]^2 + (2\lambda-1)\|\mathcal{C}_{2j}^{(\lambda)}\|_2^2\Big),$$

$$\|\mathcal{C}_{n+1}^{(\lambda+1)}\|_2^2 - \|\mathcal{C}_1^{(\lambda+1)}\|_2^2 = \sum_{j=1}^{m} \frac{2j+1+\lambda}{2\lambda^2}\Big(\big[\mathcal{C}_{2j+1}^{(\lambda)}(1)\big]^2 + (2\lambda-1)\|\mathcal{C}_{2j+1}^{(\lambda)}\|_2^2\Big).$$

Using (4.1) and summing up, and an application of Dette's result (4.2) yields:

$$\begin{aligned}
&\int_0^1 \big[\mathcal{C}_{n+1}^{(\lambda+1)}(x)\big]^2 + \big[\mathcal{C}_n^{(\lambda+1)}(x)\big]^2 \,\mathrm{d}x \\
&= \frac{4}{3}(\lambda+1)^2 + 1 + \frac{1}{2\lambda}\sum_{j=2}^{n+1}\frac{j+\lambda}{\lambda}\big[\mathcal{C}_j^{(\lambda)}(1)\big]^2 + \frac{2\lambda-1}{2\lambda}\sum_{j=2}^{n+1}\frac{j+\lambda}{\lambda}\|\mathcal{C}_j^{(\lambda)}\|_2^2 \\
&= \frac{(n+2)^2}{8\lambda^3}\big[\mathcal{C}_{n+2}^{(\lambda)}(1)\big]^2 + \frac{2\lambda-1}{2\lambda}\sum_{j=0}^{n+1}\frac{j+\lambda}{\lambda}\|\mathcal{C}_j^{(\lambda)}\|_2^2 \\
&= \frac{(n+2)^2}{8\lambda^3}\big[\mathcal{C}_{n+2}^{(\lambda)}(1)\big]^2 + \frac{2\lambda-1}{2\lambda}\Big(\frac{(n+2)^2}{4\lambda^2}\|\mathcal{C}_{n+2}^{(\lambda)}\|_2^2 + \big\|\sqrt{1-x^2}\,\mathcal{C}_{n+1}^{(\lambda+1)}\big\|_2^2\Big).
\end{aligned}$$

The case $n+1 = 2m$ is analogous. ■

Lemma 4.3.3. *The Gegenbauer polynomials satisfy the following identity:*

$$\int_0^1 x^2\big[\mathcal{C}_{n+1}^{(\lambda+1)}(x)\big]^2 - \big[\mathcal{C}_n^{(\lambda+1)}(x)\big]^2 \,\mathrm{d}x = \frac{n+2}{4\lambda^2}\Big(\big[\mathcal{C}_{n+2}^{(\lambda)}(1)\big]^2 - (n+3)\|\mathcal{C}_{n+2}^{(\lambda)}\|_2^2\Big).$$

Proof. Note first that by (4.4) and by quadratic completion

$$\begin{aligned}
2\frac{n+2}{2\lambda}\mathcal{C}_{n+2}^{(\lambda)}(x)\mathcal{C}_n^{(\lambda+1)}(x) &= 2x\mathcal{C}_{n+1}^{(\lambda+1)}(x)\mathcal{C}_n^{(\lambda+1)}(x) - 2\big[\mathcal{C}_n^{(\lambda+1)}(x)\big]^2 \\
&= x^2\big[\mathcal{C}_{n+1}^{(\lambda+1)}(x)\big]^2 - \big[\mathcal{C}_n^{(\lambda+1)}(x)\big]^2 - \Big(x\mathcal{C}_{n+1}^{(\lambda+1)}(x) - \mathcal{C}_n^{(\lambda+1)}(x)\Big)^2.
\end{aligned} \tag{4.6}$$

Hence by the binomial theorem and again by (4.4)

$$\begin{aligned}
&2\int_0^1 x^2\big[\mathcal{C}_{n+1}^{(\lambda+1)}(x)\big]^2 - \big[\mathcal{C}_n^{(\lambda+1)}(x)\big]^2 \,\mathrm{d}x \\
&= \frac{n+2}{\lambda}\int_0^1 \Big(x\mathcal{C}_{n+1}^{(\lambda+1)}(x) + \mathcal{C}_n^{(\lambda+1)}(x)\Big)\mathcal{C}_{n+2}^{(\lambda)}(x)\,\mathrm{d}x \\
&= \frac{n+2}{\lambda}\int_0^1 \frac{x}{4\lambda}\frac{\mathrm{d}}{\mathrm{d}x}\big[\mathcal{C}_{n+2}^{(\lambda)}(x)\big]^2 + \mathcal{C}_n^{(\lambda+1)}(x)\mathcal{C}_{n+2}^{(\lambda)}(x)\,\mathrm{d}x \\
&= \frac{n+2}{4\lambda^2}\Big(\big[\mathcal{C}_{n+2}^{(\lambda)}(1)\big]^2 - \int_0^1 \big[\mathcal{C}_{n+2}^{(\lambda)}(x)\big]^2\,\mathrm{d}x\Big) + 2\frac{n+2}{2\lambda}\int_0^1 \mathcal{C}_n^{(\lambda+1)}(x)\mathcal{C}_{n+2}^{(\lambda)}(x)\,\mathrm{d}x
\end{aligned}$$

which proves the result when we substitute (4.6) and use (4.4) one last time. ■

4.4. Proof of the Main Results

Proof of Theorem 4.1.3. Subtract the left hand sides of Lemma 4.3.2 and Lemma 4.3.3:

$$2\|\mathcal{C}_n^{(\lambda+1)}\|_2^2 + \frac{1}{2\lambda}\|\sqrt{1-x^2}\,\mathcal{C}_{n+1}^{(\lambda+1)}\|_2^2 = \Big(\frac{(n+2)^2}{8\lambda^3} - \frac{n+2}{4\lambda^2}\Big)\big[\mathcal{C}_{n+2}^{(\lambda)}(1)\big]^2 + \Big(\frac{(n+2)^2}{4\lambda^2} + \frac{(n+2)(n+3)}{4\lambda^2}\Big)\|\mathcal{C}_{n+2}^{(\lambda)}\|_2^2 - \frac{1}{2\lambda}\frac{(n+2)^2}{4\lambda^2}\|\mathcal{C}_{n+2}^{(\lambda)}\|_2^2;$$

an application of Dette's formula (4.2) then gives the desired expression. ■

Corollary 4.4.0.1. *The Gegenbauer polynomials satisfy the following identity:*

$$\|\sqrt{1-x^2}\,\mathcal{C}_{n-1}^{(\lambda+1)}\|_2^2 = \big[\mathcal{C}_n^{(\lambda)}(1)\big]^2\,\frac{n+2\lambda}{n+1}\frac{1-2\lambda}{2^3\lambda^2} + \frac{(n+1)(2n+3)}{2^3\lambda^2}\|\mathcal{C}_{n+1}^{(\lambda)}\|_2^2 - \|\mathcal{C}_n^{(\lambda)}\|_2^2\,\frac{n+2\lambda}{2^3\lambda^2}.$$

Proof. We use Lemma 4.3.3, add zero and obtain with Theorem 4.1.3

$$\begin{aligned}
&\frac{n}{4\lambda^2}\Big(\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 - (n+1)\|\mathcal{C}_n^{(\lambda)}\|_2^2\Big) + \int_0^1 (1-x^2)\big[\mathcal{C}_{n-1}^{(\lambda+1)}(x)\big]^2\,\mathrm{d}x\\
&= \int_0^1 \big[\mathcal{C}_{n-1}^{(\lambda+1)}(x)\big]^2 - \big[\mathcal{C}_{n-2}^{(\lambda+1)}(x)\big]^2\,\mathrm{d}x\\
&= \frac{(n+1)^2-2\lambda(n+1)}{2^4\lambda^3}\big[\mathcal{C}_{n+1}^{(\lambda)}(1)\big]^2 + \frac{(n+1)(2n+3)}{2^3\lambda^2}\|\mathcal{C}_{n+1}^{(\lambda)}\|_2^2 - \sum_{k=0}^{n}\frac{\lambda+k}{2^2\lambda^2}\|\mathcal{C}_k^{(\lambda)}\|_2^2\\
&\qquad - \frac{n^2-2\lambda n}{2^4\lambda^3}\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 - \frac{n(2n+1)}{2^3\lambda^2}\|\mathcal{C}_n^{(\lambda)}\|_2^2 + \sum_{k=0}^{n-1}\frac{\lambda+k}{2^2\lambda^2}\|\mathcal{C}_k^{(\lambda)}\|_2^2\\
&= \big[\mathcal{C}_n^{(\lambda)}(1)\big]^2\Big(\frac{(n+1)^2-2\lambda(n+1)}{2^4\lambda^3}\frac{(n+2\lambda)^2}{(n+1)^2} - \frac{n^2-2\lambda n}{2^4\lambda^3}\Big)\\
&\qquad + \frac{(n+1)(2n+3)}{2^3\lambda^2}\|\mathcal{C}_{n+1}^{(\lambda)}\|_2^2 - \|\mathcal{C}_n^{(\lambda)}\|_2^2\Big(\frac{n(2n+1)}{2^3\lambda^2} + \frac{\lambda+n}{2^2\lambda^2}\Big)\\
&= \big[\mathcal{C}_n^{(\lambda)}(1)\big]^2\,\frac{2n^2+3n+2\lambda-2\lambda n-4\lambda^2}{2^3\lambda^2(n+1)}\\
&\qquad + \frac{(n+1)(2n+3)}{2^3\lambda^2}\|\mathcal{C}_{n+1}^{(\lambda)}\|_2^2 - \|\mathcal{C}_n^{(\lambda)}\|_2^2\Big(\frac{2n^2+3n+2\lambda}{2^3\lambda^2}\Big).
\end{aligned}$$

We re-order to obtain the result. ■

For our asymptotic analysis we will need the following identity, which follows from the proof of Theorem 4.1.3 and Corollary 4.4.0.1:

Corollary 4.4.0.2. *The Gegenbauer polynomials satisfy the following identity:*

$$\begin{aligned}
\|\mathcal{C}_{n-2}^{(\lambda+1)}\|_2^2 &= \frac{n^2-2\lambda n}{2^4\lambda^3}\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 + \frac{2n^2(4\lambda-1)+n(4\lambda+1)+2\lambda}{2^5\lambda^3}\|\mathcal{C}_n^{(\lambda)}\|_2^2\\
&\quad - \big[\mathcal{C}_n^{(\lambda)}(1)\big]^2\,\frac{n+2\lambda}{n+1}\frac{1-2\lambda}{2^5\lambda^3} - \frac{(n+1)(2n+3)}{2^5\lambda^3}\|\mathcal{C}_{n+1}^{(\lambda)}\|_2^2.
\end{aligned} \tag{4.7}$$

Using the following asymptotic form, see [TE51]: For $|z| \to \infty$ and $\alpha, \beta \geq 0$:

$$\frac{\Gamma(z+\alpha)}{\Gamma(z+\beta)} = z^{\alpha-\beta}\Big(1 + \frac{(\alpha-\beta)(\alpha+\beta-1)}{2z} + O(|z|^{-2})\Big), \tag{4.8}$$

we obtain by (4.3) for $\lambda > 0$:

$$\Gamma(2\lambda)^2 \cdot \big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 = n^{4\lambda-2} + 2\lambda(2\lambda-1)n^{4\lambda-3} + O(n^{4\lambda-4}). \tag{4.9}$$

Proof of Corollary 4.1.3.1. We will write $\|\mathcal{C}_n^{(\lambda)}\|_2^2 = \Theta(n^{\Phi(\lambda)})$ if there are some constants $c_1, c_2 > 0$ such that $c_1 n^{\Phi(\lambda)} \leq \|\mathcal{C}_n^{(\lambda)}\|_2^2 \leq c_2 n^{\Phi(\lambda)}$ for all n big enough. First we use (4.7) to show by induction that $\Phi(\lambda)$ exists for $\lambda > 1$, and that $\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 = \Theta(n^{\Phi(\lambda)+2})$.

The case $\lambda = m \in \mathbb{N}_{>1}$: Lemma 4.1.1 gives the result for $\lambda = 2$, and if it holds for m, then with (4.7) and abuse of notation we have:

$$\big\|\mathcal{C}_{n-2}^{(m+1)}\big\|_2^2 = n^2\ \Theta\big(n^{\Phi(m)+2}\big) + n^2\ \Theta\big(n^{\Phi(m)}\big) + \Theta\big(n^{\Phi(m)+2}\big) + n^2\ \Theta\big(n^{\Phi(m)}\big).$$

This proves the claim as it shows that $\|\mathcal{C}_n^{(m+1)}\|_2^2 = \Theta(n^{\Phi(m)+4})$, but by (4.3):

$$\mathcal{C}_n^{(\lambda+1)}(1) = \frac{(2\lambda+n+1)(2\lambda+n)}{2\lambda(2\lambda+1)}\mathcal{C}_n^{(\lambda)}(1) = \Theta\big(n^2\mathcal{C}_n^{(\lambda)}(1)\big), \tag{4.10}$$

which, when squared and $\lambda = m$, is of order $\Phi(m) + 6$.

The case $\lambda \in (m, m+1)$, $m \in \mathbb{N}$: For $\lambda \in (0,1)$ and $\theta \in [0, \pi]$:

$$\sin(\theta)^\lambda\big|\mathcal{C}_n^{(\lambda)}(\cos(\theta))\big| < \frac{2^{1-\lambda}}{\Gamma(\lambda)}n^{\lambda-1} \qquad \text{see [Sze39, Eq. 7.33.5].}$$

We square this inequality, multiply by $\sin(\theta)^{1-2\lambda}$ and integrate:

$$\big\|\mathcal{C}_n^{(\lambda)}\big\|_2^2 < \frac{2^{2-2\lambda}}{\Gamma(\lambda)^2}n^{2\lambda-2}\int_0^{\pi/2}\sin(\theta)^{1-2\lambda}\ \mathrm{d}\theta = \mathcal{B}\big(1-\lambda, \tfrac{1}{2}\big)\frac{2^{1-2\lambda}}{\Gamma(\lambda)^2}n^{2\lambda-2};$$

where we used a change of variables $\theta = \arcsin(x)$ and $\mathcal{B}(x,y)$ is the beta function. This in combination with (4.9) and (4.7) gives for $\delta = \max\{4\lambda-1, 2\lambda\}$:

$$\big\|\mathcal{C}_n^{(\lambda+1)}\big\|_2^2 = \frac{n^{4\lambda}}{2^4\lambda^3\Gamma(2\lambda)^2} + O\big(n^\delta\big). \tag{4.11}$$

Thus for $\lambda \in (0,1)$: $\Phi(\lambda+1) = 4\lambda$, and $[\mathcal{C}_n^{(\lambda+1)}(1)]^2 = \Theta\big(n^{4\lambda+2}\big)$ by (4.9), which finishes the case for the interval $(1,2)$ and we use induction with (4.7) and (4.10).

Thus the two leading terms in the asymptotic form of $\|\mathcal{C}_n^{(\lambda+1)}\|_2^2$ are in the expansion of $\mathcal{C}_n^{(\lambda)}(1)$ when $\lambda > 1$; using once more (4.9) and (4.7) yields

$$\frac{n^2-2\lambda n}{2^4\lambda^3}\big[\mathcal{C}_n^{(\lambda)}(1)\big]^2 = \frac{n^{4\lambda}}{4\lambda\Gamma(2\lambda+1)^2} + \frac{2\lambda(2\lambda-2)}{4\lambda\Gamma(2\lambda+1)^2}n^{4\lambda-1} + O(n^{4\lambda-2}).$$

The asymptotic of the rest term of $\|\mathcal{C}_n^{(\lambda+1)}\|_2^2$ follows by (4.7), equation (4.11) and induction for non-integer $\lambda \in \mathbb{R}_{>2}$ or else Lemma 4.1.1 and induction when $\lambda \in \mathbb{N}_{>2}$. Now, these asymptotic formulas in combination with Corollary 4.4.0.1 and (4.9) will finish the argument: As an illustration, let $1 < \lambda \leq 2$, the cases $0 < \lambda < 1$ and $\lambda > 2$ are similar; let $\rho = \max\{4\lambda - 3, 2\lambda\}$, then

$$\begin{aligned}
\left\|\sqrt{1-x^2}\, \mathcal{C}_{n-1}^{(\lambda+1)}\right\|_2^2 &= \frac{1-2\lambda}{2^3\lambda^2}\left[\mathcal{C}_n^{(\lambda)}(1)\right]^2 + \frac{2n^2}{2^3\lambda^2}\|\mathcal{C}_{n+1}^{(\lambda)}\|_2^2 + O(n^\rho) \\
&= \frac{1-2\lambda}{2^3\lambda^2}\frac{n^{4\lambda-2}}{\Gamma(2\lambda)^2} + \frac{2n^2}{2^3\lambda^2}\frac{n^{4\lambda-4}}{4(\lambda-1)\Gamma(2\lambda-1)^2} + O(n^\rho) \\
&= \frac{n^{4\lambda-2}}{2^3\lambda^2\Gamma(2\lambda-1)^2}\left(\frac{1-2\lambda}{(2\lambda-1)^2} + \frac{1}{2(\lambda-1)}\right) + O(n^\rho) \\
&= \frac{n^{4\lambda-2}}{2^3\lambda^2\Gamma(2\lambda-1)^2}\frac{1}{(2\lambda-1)(\lambda-1)2} + O(n^\rho). \qquad \blacksquare
\end{aligned}$$

Remark 4.4.1. One can use Lemma 4.1.1, Corollary 4.4.0.1 and identity (4.7) to find formulas for $\|\sqrt{1-x^2}\, \mathcal{C}_n^{(m)}\|_2^2$ and $\|\mathcal{C}_n^{(m)}\|_2^2$ where $m \in \mathbb{N}_{>1}$.

4.5. Acknowledgement

The help of J. Brauchart, P. Grabner and J. Thuswaldner is gratefully appreciated; who proof read the manuscript, made useful remarks on presentation and suggested to generalize Corollary 4.1.3.1 from $\lambda \in \frac{1}{2}\mathbb{N}$ to $\lambda > 0$.

5. Potential Theory with Multivariate Kernels

This chapter is essentially [Bil+21b] and is a joint work with Dmitriy Bilyk, Alexey Glazyrin, Ryan Matzke, Josiah Park and Oleksandr Vlasiuk. It has been modified slightly and section 5.7 has been added by this author which gives interesting counterexamples to natural conjectures related to multivariate kernels.

Abstract

In the present paper we develop the theory of minimization for energies with multivariate kernels, i.e. energies, in which pairwise interactions are replaced by interactions between triples or, more generally, n-tuples of particles. Such objects, which arise naturally in various fields, present subtle differences and complications when compared to the classical two-input case. We introduce appropriate analogues of conditionally positive definite kernels, establish a series of relevant results in potential theory, explore rotationally invariant energies on the sphere, and present a variety of interesting examples, in particular, some optimization problems in probabilistic geometry which are related to multivariate versions of the Riesz energies.

5.1. Introduction and main results

Numerous questions, which arise in such different disciplines as discrete geometry, physics, signal processing, and many others, can be reformulated as problems of minimization of discrete or continuous pairwise interaction energies, i.e. expressions of the type

$$\frac{1}{N^2}\sum_{x,y\in\omega_N} K(x,y) \quad \text{or} \quad \int_\Omega\int_\Omega K(x,y)\,d\mu(x)\,d\mu(y), \tag{5.1}$$

where ω_N is a discrete set of N points, μ is a Borel probability measure on the domain Ω, and K is the potential function describing the pairwise interaction. Perhaps one of the most celebrated examples of such problems is the 1904 Thomson problem, asking for an equilibrium distribution of N electrons on the sphere, which is notoriously still open for most values of N [Th]. This and many other problems stimulated the study of such energies, which has now developed into a full-blown theory, see e.g. [Bjö56; Fug60; HS04], whose state-of-the-art is very well presented in a recent book [BHS19].

While classical energies (5.1) model pairwise interactions between particles, the present paper, in contrast, initiates the study of optimization problems for more complicated energies, defined by interactions of triples, quadruples, or even higher numbers of particles, i.e. energies of the type

$$E_K(\omega_N) = \frac{1}{N^n} \sum_{x_1,\dots,x_n \in \omega_N} K(x_1, ..., x_n), \tag{5.2}$$

$$I_K(\mu) = \int_\Omega \cdots \int_\Omega K(x_1, \dots, x_n)\, d\mu(x_1) \, \dots \, d\mu(x_n), \tag{5.3}$$

with $n \geq 3$. Energies of this type arise naturally in various fields:

1. In different branches of *physics* (nuclear, quantum, chemical, condensed matter, material science etc.), it has been suggested that, if the behavior of the system cannot be accurately modeled by two-body interactions, more precise information may be obtained from three-body or many-body interactions. Such forces are observed among nucleons in atomic nuclei (three-nucleon force) [Zel09], in carbon nanostructures [MS14], crystallization of atomistic configurations [FT15], cold polar molecules in optical lattices [BMZ07], interactions of solid and liquid forms of silicon [SW85], interactions between atoms [AT43], in "perfect glass" potentials [ZST16], and many other areas.

2. Energy integrals with multivariate kernels defined in (5.3) play the role of polynomials on the space $\mathbb{P}(\Omega)$ of probability measures on Ω – e.g., their linear span over all $n \in \mathbb{N}$ is dense in the space of continuous functions on $\mathbb{P}(\Omega)$, according to the Stone–Weierstrass theorem. Such functionals on the space of measures appear in optimal transport [San15] and mean field games [Lio06].

3. A classical example of a three-input energy, coming from geometric measure theory, is given by the total Menger curvature of a measure μ

$$c^2(\mu) = \int_\Omega \int_\Omega \int_\Omega c^2(x, y, z)\, d\mu(x)\, d\mu(y)\, d\mu(z), \tag{5.4}$$

where $c(x, y, z) = \frac{1}{R(x,y,z)}$ and $R(x, y, z)$ is the circumradius of the triangle xyz. This object plays an important role in the study of the L^2 boundedness of the Cauchy integral, analytic capacity, and uniform rectifiability [Dav99; MMV96].

4. Some questions in *probabilistic geometry* admit natural reformulations in terms of multi-input energies (5.2) or (5.3). For example, assume that three points are chosen in a domain Ω, e.g. $\Omega = \mathbb{S}^2$, independently at random, according to the probability distribution μ. Which probability distribution maximizes the expected area of the triangle generated by these random points or the volume of the parallelepiped spanned by the random vectors? These quantities can be written as energy integrals (5.2) with $n = 3$, and higher dimensional versions of such questions call for energies with more inputs, which may be viewed as natural extensions of the classical Riesz energy. Questions of this type are discussed in Section 5.6.4 and are explored in more detail in [Bil+21a].

5. Energies with more than two arguments akin to (5.2) appear in three-point bounds [CW12] and, more generally, k-point bounds [Laa+19; Mus14] in semidefinite programming [BV08] – a very fruitful method, which led to numerous breakthroughs in discrete geometry. A discussion of this method in the context of the multivariate energy optimization and, in particular, applications to the geometric problems described in Section 5.6.4 can be found in our follow-up work [Bil+21a].

6. Relations between the L^2 discrepancy and the two-input energies, in particular, the Stolarsky principle [Sto73], are well known [BDM18; Skr20]. In a similar spirit, other L^n norms of the discrepancy or "number variance" with integer values of n lead to n-particle interaction energies (5.2). Some similar ideas have been put forward in [Tor10].

Despite the abundance of applications, there seems to have been no systematic development of a general theory of multi-input energies, unlike the case of classical two-input energies which has been deeply and extensively explored. The present paper makes a first attempt to remedy this shortcoming and to study the general properties of point configurations and measures, minimizing the multi-input energies (5.2)-(5.3), and the relations between the structure of the multivariate kernel K and the energy minimizers. This theory presents many intrinsic obstacles and is far from a straightforward generalization of the two-input case. In particular, in the spherical case $\Omega = \mathbb{S}^{d-1}$ with rotationally-invariant two-input kernels $K(x, y) = F(\langle x, y \rangle)$, classical Schoenberg's theory [Sch41] proves that the uniform surface measure σ minimizes the energy integral in (5.1) if and only if the kernel K is conditionally positive definite. However, in the multi-input case, such a characterization is still elusive: while we obtain various natural sufficient conditions for the surface measure σ to minimize the energy (5.2) in Section 5.5, counterexamples presented in Section 5.6 show that none of them are necessary.

The outline of the paper is as follows. In Section 5.2 we introduce the notation and some of the main definitions, including the notion of n-positive definiteness. In Section 5.3 we explore some basic properties of multivariate energies. In particular, we analyze the connections between (conditional) positive definiteness of the kernel K, convexity of the energy functional $I_K(\mu)$, and arithmetic and geometric mean inequalities for the mixed energies. The meat of the paper, i.e. the results about minimizers of the n-input energies are concentrated in Sections 5.4–5.6.

Section 5.4 deals with analogues of classical potential theoretic results [Bjö56; BHS19; Fug60], which provide certain necessary (Theorem 5.4.1) and sufficient (Theorem 5.4.2) conditions for a measure μ to be a minimizer of the n-input energy integral in terms of the $(n-1)$-fold potential of the kernel K with respect to μ. Even though some of these results by themselves are clear-cut generalizations of standard statements for two-input energies, they yield several interesting consequences in the n-input case. In particular, Theorem 5.4.5 states that, under some additional assumptions (e.g., if K is n-positive definite), for any $1 \leq k \leq n-2$, if the measure μ minimizes the $(n-k)$-input energy I_U, where U is the k-fold integral of K with respect to μ, then μ also minimizes the n-input energy I_K. This statement allows one to simplify proving that a given measure is a minimizer of a multi-input energy by considering energies with a lower number of inputs. A partial converse to Theorem 5.4.5, for $k = n - 2$, is given in Theorem 5.4.6. In addition, in Lemma 5.4.3, we show that, for n-positive definite kernels, every local minimizer of I_K is necessarily a global minimizer.

In Section 5.5 we adapt the methods of Section 5.4 to energies with rotationally invariant kernels on the sphere $\mathbb{S}^{d-1}$, where symmetries allow for a more delicate analysis, and one has a natural candidate for a minimizer: the uniform surface measure σ. Theorem 5.5.1 states that energies with conditionally n-positive definite rotationally invariant kernels on the sphere are minimized by the surface measure σ (without any additional assumptions). As mentioned above, it turns out that, in contrast to the classical case $n = 2$, conditional n-positive definiteness is not necessary for σ to minimize the n-input energy, which is shown by examples presented in Propositions 5.6.9 and 5.6.10. Nevertheless, Theorem 5.5.1 allows one to prove that σ minimizes a variety of interesting energies, which did not seem to be accessible by different methods, see e.g. Corollary 5.5.1.1. In Theorem 5.5.2 we obtain very close necessary and sufficient conditions for σ to be a *local* minimizer of the n-input energy I_K in terms of the minimization properties of the two-input energy with the kernel given by the $(n-2)$-fold integral of K (or the conditional positive definiteness of this kernel). We also conjecture these are the correct conditions for σ to be a *global* minimizer of I_K.

Section 5.6 is dedicated to constructing various classes of n-positive definite kernels, proving that certain kernels of interest are (conditionally) n-positive definite, as well as exhibiting some naturally arising 3-input kernels on the sphere which are not conditionally 3-positive definite, yet the corresponding energies are minimized by the surface measure σ. These examples are presented in Propositions 5.6.5, 5.6.9, and 5.6.10. The first one is closely related to the semidefinite programming method as presented in [BV08], while the last two are geometric. The latter kernels are studied in Section 5.6.4 which addresses some problems from probabilistic discrete geometry. Their main objects may be viewed as multi-input analogues of the classical Riesz energies. In particular, we show that if three random vectors are chosen in the sphere $\mathbb{S}^{d-1}$ independently according to the probability distribution μ, then the expected volume squared of the tetrahedron generated by these vectors (Theorem 5.6.6) as well as the square of the area of the triangle defined by these points (Theorem 5.6.7) are maximized if the distribution is uniform, i.e. $\mu = \sigma$. A more detailed study of such geometric questions is conducted by the authors in [Bil+21a].

While many of the results presented in this paper hold (or can be extended) to a larger class of kernels (e.g., bounded lower semi-continuous, or even singular kernels), given that this is the first effort to establish a theory of multi-input energies, for the sake of brevity and clarity of the exposition, we shall only consider continuous kernels on compact metric spaces in this paper. We shall also restrict our attention to symmetric n-input kernels, i.e. functions invariant with respect to any permutation of variables. These assumptions are implicitly present in all of the results presented below, even if not stated explicitly.

5.2. Background and definitions

In what follows, we always assume that (Ω, ρ) is a compact metric space, $n \in \mathbb{N} \setminus \{1\}$, and the kernel $K : \Omega^n \to \mathbb{R}$ is continuous and symmetric, i.e. for any permutation $\pi \in S_n$ and $x_1, ..., x_n \in \Omega$, $K(x_1, ..., x_n) = K(x_{\pi(1)}, ..., x_{\pi(n)})$. We denote by $\mathcal{M}(\Omega)$ the set of finite signed Borel measures on Ω, and by $\mathbb{P}(\Omega)$ the set of Borel probability measures on Ω. Let $\omega_N = \{x_1, x_2, ..., x_N\}$ be an N-point configuration (multiset) in Ω for $N \geq n$. We define the discrete K-energy of ω_N to be

$$E_K(\omega_N) := \frac{1}{N^n} \sum_{j_1=1}^{N} \cdots \sum_{j_n=1}^{N} K(x_{j_1}, ..., x_{j_n}), \tag{5.5}$$

and the minimal discrete N-point K-energy of Ω as

$$\mathcal{E}_K(\Omega, N) := \inf_{\omega_N \subseteq \Omega} E_K(\omega_N). \tag{5.6}$$

Let $\mu_1, ..., \mu_n \in \mathcal{M}(\Omega)$, then we define their mutual energy as

$$I_K(\mu_1, ..., \mu_n) = \int_\Omega \cdots \int_\Omega K(x_1, ..., x_n) d\mu_1(x_1) \cdots d\mu_n(x_n), \tag{5.7}$$

and, for $j < n$, the j-th potential function as

$$U_K^{\mu_1,...,\mu_j}(x_{j+1}, ..., x_n) = \int_\Omega \cdots \int_\Omega K(x_1, ..., x_n) d\mu_1(x_1) \cdots d\mu_j(x_j). \tag{5.8}$$

Note that since we are working with continuous K, the energy is well defined for all finite signed Borel measures. We will abuse notation, by writing μ^k if k of the measures are the same, and define the **K-energy functional** on $\mathcal{M}(\Omega)$ by

$$I_K(\mu) = I_K(\mu^n) = I_K(\mu, ..., \mu). \tag{5.9}$$

The definitions of discrete (5.5) and continuous (5.9) energies are compatible in the sense that

$$E_K(\omega_N) = I_K(\mu_{\omega_N}), \qquad \text{where } \mu_{\omega_N} = \frac{1}{N} \sum_{j=1}^{N} \delta_{x_j} \tag{5.10}$$

and due to the weak-$*$ density of the linear span of Dirac masses in $\mathbb{P}(\Omega)$

$$\lim_{N \to \infty} \mathcal{E}_K(\Omega, N) = \inf_{\mu \in \mathbb{P}(\Omega)} I_K(\mu). \tag{5.11}$$

We now recall the classical notion of positive definiteness for two-input kernels, which plays an extremely important role in energy optimization problems and which we seek to generalize to n-input kernels. We state the definition in the form which is most relevant to our exposition.

Definition 5.2.1. A kernel $K : \Omega^2 \to \mathbb{R}$ is called positive definite if for every finite signed Borel measure $\nu \in \mathcal{M}(\Omega)$, the energy integral satisfies $I_K(\nu) \geq 0$.

If the inequality $I_K(\nu) \geq 0$ holds for all $\nu \in \mathcal{M}(\Omega)$ satisfying $\nu(\Omega) = 0$, we call the kernel conditionally positive definite.

A more standard way of stating the definition of positive definiteness of K is by requiring that for all $N \in \mathbb{N}$ and $x_1, ..., x_N \in \Omega$, the matrix $[K(x_i, x_j)]_{0 \leq i,j \leq N}$ is positive semi-definite, i.e.

$$I_K\left(\sum_{i=1}^{N} c_i \delta_{x_i}\right) = \sum_{i,j=1}^{N} K(x_i, x_j) c_i c_j \geq 0$$

for all $c_1, \ldots, c_N \in \mathbb{R}$. Since K is continuous, this is clearly equivalent to Definition 5.2.1 due to weak-$*$ density of discrete measures.

We extend this notion to n-input kernels by demanding that, if one fixes arbitrary values of all but two variables, the resulting two-input kernel is positive definite in the classical sense. For every $m < n$ and $z_1, z_2, ..., z_m \in \Omega$, we define

$$K_{z_1,z_2,...,z_m}(x_1, ..., x_{n-m}) := K(z_1, ..., z_m, x_1, ..., x_{n-m}). \tag{5.12}$$

Definition 5.2.2. We shall say that a continuous symmetric kernel $K : \Omega^n \to \mathbb{R}$ is (conditionally) n-positive definite if, for all $z_1, z_2, ..., z_{n-2} \in \Omega$, the two-input kernel $K_{z_1,...,z_{n-2}}$ is (conditionally) positive definite in the sense of Definition 5.2.1.

We would like to emphasize that this definition relies more on the pointwise two-variable structure, rather than the full set of variables. In particular, it doesn't have any connection to positive definite tensors [Qi05]. Thus, it may appear that the name *n-positive definite* might be somewhat misleading. However, from the point of view of energy optimization, which is the main theme of this paper, this nomenclature seems absolutely justified. Indeed, in various statements about minimal energy (e.g., Theorem 5.4.2, Corollary 5.4.2.1, Theorem 5.5.1), this condition naturally replaces positive definiteness of classical two-input kernels. In addition, non-symmetric multivariate kernels of similar flavor have been considered in the context of k-point bounds in semidefinite programming [Laa+19; Mus14]. The class of n-positive definite kernels is rather rich: throughout the text, in particular, in Section 5.6, we present numerous examples of functions with this property.

We immediately observe that this property is inherited by kernels with a lower number of inputs, which are obtained as potentials of K with respect to arbitrary probability measures.

Lemma 5.2.1. *Assume that K is (conditionally) n-positive definite. Then for every $\mu \in \mathbb{P}(\Omega)$, the potential $U_K^\mu(x_1, \ldots, x_{n-1})$ is (conditionally) $(n-1)$-positive definite.*

Proof. Let ν be a finite signed Borel measure on Ω (with $\nu(\Omega) = 0$ if K is conditionally n-positive definite). Then by Fubini–Tonelli

$$\begin{aligned} I_{\left(U_K^\mu\right)_{z_2,\ldots,z_{n-2}}}(\nu) &= \int_\Omega \int_\Omega \int_\Omega K(z_1, z_2, \ldots, z_{n-3}, z_{n-2}, x, y) d\mu(z_1)\, d\nu(x) d\nu(y) \\ &= \int_\Omega \int_\Omega \int_\Omega K_{z_1,\ldots,z_{n-2}}(x, y) d\nu(x) d\nu(y)\, d\mu(z_1) \geq 0, \end{aligned}$$

since $K_{z_1,\ldots,z_{n-2}}$ is (conditionally) positive definite for all $z_1, \ldots, z_{n-2} \in \Omega$. ■

As a corollary of Lemma 5.2.1, we observe that if $K : \Omega^n \to \mathbb{R}$ is (conditionally) n-positive definite, then for all $\mu_1, \ldots, \mu_k \in \mathbb{P}(\Omega)$, with $k \leq n-2$, $U_K^{\mu_1,\ldots,\mu_k}(x_{k+1}, \ldots, x_n)$ is (conditionally) $(n-k)$-positive definite.

Naturally, (conditionally) n-positive definite kernels enjoy the same basic properties as their classical two-variable counterparts.

Lemma 5.2.2. *If K and L are n-positive definite, then so are $K+L$ and KL. If $K_1, K_2, \ldots$ are n-positive definite and $\lim_{n\to\infty} K_n = K$ (uniformly), then K is n-positive definite. The statements about the sum and the limit (but not about the product) continue to hold if we replace n-positive definite with conditionally n-positive definite.*

The proof of this lemma is straightforward. The statement about the product KL follows from the classical Schur's product theorem, and positive definiteness in this statement cannot be replaced by conditional positive definiteness (since, for example, a negative constant is a conditionally n-positive definite function).

5.3. First principles

In this section we explore some basic properties related to (conditional) n-positive definiteness, such as inequalities for mixed energies and convexity of the energy functionals, as well as connections between these notions. All the kernels in this section are assumed to be continuous and symmetric.

5.3.1. Bounds on mutual energies

In the classical case, mixed energies can be bounded by averages of energies of each individual measure. We refer the reader to Chapter 4 of [BHS19] for details.

Lemma 5.3.1. *Suppose K is a conditionally positive definite kernel on Ω^2. Then for every pair of Borel probability measures μ_1 and μ_2 on Ω, the mutual energy $I_K(\mu_1, \mu_2)$ satisfies*

$$I_K(\mu_1, \mu_2) \leq \frac{1}{2}\big(I_K(\mu_1) + I_K(\mu_2)\big).$$

Furthermore, if K is positive definite, then

$$I_K(\mu_1, \mu_2) \leq \sqrt{I_K(\mu_1) I_K(\mu_2)}.$$

These inequalities can be extended to n-input energies with (conditionally) n-positive definite kernels.

Lemma 5.3.2. *Suppose K is a conditionally n-positive definite kernel on Ω^n. Then for every n-tuple of Borel probability measures $\mu_1, \ldots, \mu_n$ on Ω, the mutual energy $I_K(\mu_1, \ldots, \mu_n)$ satisfies*

$$I_K(\mu_1, \ldots, \mu_n) \leq \frac{1}{n} \sum_{j=1}^{n} I_K(\mu_j). \tag{5.13}$$

If, moreover, K is n-positive definite,

$$I_K(\mu_1, \dots, \mu_n) \leq \prod_{j=1}^{n} \sqrt[n]{I_K(\mu_j)}. \tag{5.14}$$

Proof. We only prove (5.14), as one could repeat the proof below verbatim, with the multiplicative notation replaced by the additive, to arrive at (5.13) (when K is n-positive definite, it would also follow from the arithmetic–geometric mean inequality).

By Lemma 5.3.1, our claim holds for $n = 2$. Now, suppose our claim holds for some $k \geq 2$, and let $\mu_1, ..., \mu_{k+1} \in \mathbb{P}(\Omega)$. Lemma 5.2.1 tells us that for $1 \leq j \leq k+1$, $U_K^{\mu_j}$ is k-positive definite, so by our inductive hypothesis

$$I_K(\mu_1, ..., \mu_{k+1}) = I_{U_K^{\mu_1}}(\mu_2, ..., \mu_{k+1}) \leq \prod_{j=1}^{k} \sqrt[k]{I_K(\mu_1, \mu_{j+1}^k)}. \tag{5.15}$$

Again using the inductive hypothesis, and the fact that K is symmetric, we have that for $1 \leq j \leq k$,

$$\begin{aligned}
I_K(\mu_1, \mu_{j+1}^k) &= I_K(\mu_{j+1}, \mu_1, \mu_{j+1}^{k-1}) \\
&\leq \sqrt[k]{I_K(\mu_{j+1}, \mu_1^k)} \sqrt[k]{I_K(\mu_{j+1})^{k-1}} \\
&= \sqrt[k]{I_K(\mu_1, \mu_{j+1}, \mu_1^{k-1})} \sqrt[k]{I_K(\mu_{j+1})^{k-1}} \\
&\leq \sqrt[k^2]{I_K(\mu_1)^{k-1}} \sqrt[k^2]{I_K(\mu_1, \mu_{j+1}^k)} \sqrt[k]{I_K(\mu_{j+1})^{k-1}},
\end{aligned}$$

where in the second and last lines we have used (5.15). Rearranging the terms, we have

$$\Big(I_K(\mu_1, \mu_{j+1}^k)\Big)^{\frac{k^2-1}{k^2}} \leq I_K(\mu_1)^{\frac{k-1}{k^2}} I_K(\mu_{j+1})^{\frac{k-1}{k}},$$

so that

$$\sqrt[k]{I_K(\mu_1, \mu_{j+1}^k)} \leq I_K(\mu_1)^{\frac{1}{k(k+1)}} I_K(\mu_{j+1})^{\frac{1}{k+1}}.$$

Plugging this back into (5.15), we have

$$I_K(\mu_1, ..., \mu_{k+1}) \leq \prod_{j=1}^{k} \sqrt[k]{I_K(\mu_1, \mu_{j+1}^k)} \leq \prod_{j=1}^{k+1} \sqrt[k+1]{I_K(\mu_j)}. \tag{5.16}$$

Our claim then follows via induction. ■

The upper bound (5.13) allows us to prove a corresponding lower bound for the mixed energy:

Corollary 5.3.2.1. *If K is n-positive definite on Ω^n, then for all $\mu_1, ..., \mu_n \in \mathbb{P}(\Omega)$,*

$$-\frac{1}{n} \sum_{j=1}^{n} I_K(\mu_j) \leq I_K(\mu_1, ..., \mu_n). \tag{5.17}$$

Proof. Suppose $n = 2$, and let $\mu_1, \mu_2 \in \mathbb{P}(\Omega)$. Setting $\mu = \frac{1}{2}(\mu_1 + \mu_2)$, we have

$$0 \leq 4I_K(\mu) = I_K(\mu_1) + I_K(\mu_2) + 2I_K(\mu_1, \mu_2),$$

since K is positive definite, and (5.17) follows.

Now suppose our claim holds for some $k \geq 2$, and let $\mu_1, ..., \mu_{k+1} \in \mathbb{P}(\Omega)$. Since by Lemma 5.2.1 the potential $U_K^{\mu_1}$ is k-positive definite, the inductive hypothesis implies that

$$-\frac{1}{k}\sum_{j=1}^{k} I_{U_K^{\mu_1}}(\mu_{j+1}) \leq I_{U_K^{\mu_1}}(\mu_2, ..., \mu_{k+1}) = I_K(\mu_1, ..., \mu_{k+1}).$$

For $1 \leq j \leq k$, Lemma 5.3.2 gives us that

$$I_{U_K^{\mu_1}}(\mu_{j+1}) = I_K(\mu_1, \mu_{j+1}^k) \leq \frac{1}{k+1}\Big(I_K(\mu_1) + kI_K(\mu_{j+1})\Big),$$

leading to

$$-\frac{1}{k+1}\sum_{j=1}^{k+1} I_K(\mu_j) \leq I_K(\mu_1, ..., \mu_{k+1}),$$

which finishes the proof of the claim. ■

Lemma 5.3.2 and Corollary 5.3.2.1 imply that if K is n-positive definite on Ω^n and $\mu_1, ..., \mu_n \in \mathbb{P}(\Omega)$, then

$$\big|I_K(\mu_1, ..., \mu_n)\big| \leq \frac{1}{n}\sum_{j=1}^{n} I_K(\mu_j). \tag{5.18}$$

Of course, since we can choose the probability measures μ_k to be Dirac masses, inequality (5.18) yields pointwise bounds on K. For instance, if K is n-positive definite, then for all $z_1, ..., z_n \in \Omega$,

$$\big|K(z_1, ..., z_n)\big| \leq \frac{1}{n}\sum_{j=1}^{n} K(z_j, ..., z_j),$$

and for conditionally n-positive definite kernels K, this inequality holds without the absolute value. Clearly then, K must achieve its maximum value on its diagonal, something that is already known for the two-input case.

Corollary 5.3.2.2. *Suppose K is a conditionally n-positive definite kernel. Then*

$$K(z_1, \ldots, z_n) \leq \max_{z\in\Omega}\{K(z, \ldots, z)\}. \tag{5.19}$$

5.3.2. Convexity

Convexity of the underlying energy functionals naturally plays an important role in energy minimization.

Definition 5.3.1. Suppose $K : \Omega^n \to \mathbb{R}$. We say that I_K is convex at $\mu \in \mathbb{P}(\Omega)$ if for every $\nu \in \mathbb{P}(\Omega)$ there exists some $t_\nu \in (0,1]$, such that for all $t \in [0, t_\nu)$

$$I_K((1-t)\mu + t\nu) \leq (1-t)I_K(\mu) + tI_K(\nu). \tag{5.20}$$

We say I_K is convex on $\mathbb{P}(\Omega)$ if it is convex at all $\mu \in \mathbb{P}(\Omega)$.

Conditional positive definiteness of the kernel K is closely related to convexity of the corresponding energy functional I_K. In fact, as we shall see in Proposition 5.3.5, when $n = 2$, the two notions are equivalent. For further discussions about the connections between various conditions related to positive definiteness in the classical two-input case, see [BMV21].

One-sided implication holds for all $n \geq 2$: as the next proposition shows, convexity of I_K can be deduced from relaxed arithmetic or geometric mean inequalities akin to (5.13) and (5.14). This implies, due to Lemma 5.3.2, that conditionally n-positive definite kernels K give rise to convex energies.

Proposition 5.3.3. *Let $K : \Omega^n \to \mathbb{R}$ be continuous and symmetric and fix $\mu \in \mathbb{P}(\Omega)$. Suppose that for all $\nu \in \mathbb{P}(\Omega)$ and $0 \leq k \leq n$,*

$$I_K(\mu^k, \nu^{n-k}) \leq \frac{k}{n} I_K(\mu) + \frac{n-k}{n} I_K(\nu). \tag{5.21}$$

Alternatively, assume that for all $\nu \in \mathbb{P}(\Omega)$ we have $I_K(\nu) \geq 0$ and for all $0 \leq k \leq n$

$$I_K(\mu^k, \nu^{n-k}) \leq \big(I_K(\mu)\big)^{\frac{k}{n}} \cdot \big(I_K(\nu)\big)^{\frac{n-k}{n}}. \tag{5.22}$$

Then I_K is convex at μ. If (5.21) *or* (5.22) *holds for all $\mu \in \mathbb{P}(\Omega)$, then I_K is convex on $\mathbb{P}(\Omega)$.*

Proof. Assume that (5.21) holds. For all $t \in [0,1]$, we have

$$\begin{aligned}
I_K((1-t)\mu + t\nu) &= \sum_{k=0}^{n} (1-t)^k t^{n-k} \binom{n}{k} I_K(\mu^k, \nu^{n-k}) \\
&\leq \sum_{k=0}^{n} (1-t)^k t^{n-k} \binom{n}{k} \left(\frac{k}{n} I_K(\mu) + \frac{n-k}{n} I_K(\nu) \right) \\
&= \sum_{k=1}^{n} (1-t)^k t^{n-k} \binom{n-1}{k-1} I_K(\mu) + \sum_{k=0}^{n-1} (1-t)^k t^{n-k} \binom{n-1}{k} I_K(\nu) \\
&= (1-t)I_K(\mu) + tI_K(\nu),
\end{aligned}$$

which proves convexity of the energy functional. The multiplicative inequality (5.22) implies (5.21) by the arithmetic-geometric mean inequality, leading to convexity of K in this case. ∎

Lemma 5.3.2 with $\mu_1 = \cdots = \mu_k = \mu$ and $\mu_{k+1} = \cdots = \mu_n = \nu$ shows that inequality (5.21) holds, if K is conditionally n-positive definite. This leads to the following corollary.

Corollary 5.3.3.1. *If K is conditionally n-positive definite, then I_K is convex on $\mathbb{P}(\Omega)$.*

To prove the converse implication for $n = 2$, we start by observing that Proposition 5.3.3 admits a partial converse:

Lemma 5.3.4. *Suppose $\mu \in \mathbb{P}(\Omega)$ is such that I_K is convex at μ. Then for all $\nu \in \mathbb{P}(\Omega)$,*

$$I_K(\mu^{n-1}, \nu) \leq \frac{n-1}{n} I_K(\mu) + \frac{1}{n} I_K(\nu). \tag{5.23}$$

Proof. Let $\nu \in \mathbb{P}(\Omega)$. Assume $t \in (0,1)$ such that (5.20) holds. Then

$$\begin{aligned} tI_K(\nu) + (1-t)I_K(\mu) &\geq I_K(t\nu + (1-t)\mu) \\ &= \sum_{j=0}^{n} (1-t)^j t^{n-j} \binom{n}{j} I_K(\mu^j, \nu^{n-j}). \end{aligned}$$

Clearly then

$$\Big(t - t^n\Big) I_K(\nu) + \Big((1-t) - (1-t)^n\Big) I_K(\mu) \geq \sum_{j=1}^{n-1} (1-t)^j t^{n-j} \binom{n}{j} I_K(\mu^j, \nu^{n-j}),$$

and dividing by $t(1-t)$, we obtain

$$\Big(\sum_{k=0}^{n-2} t^k\Big) I_K(\nu) + \Big(\sum_{l=0}^{n-2} (1-t)^l\Big) I_K(\mu) \geq \sum_{j=1}^{n-1} (1-t)^{j-1} t^{n-j-1} \binom{n}{j} I_K(\mu^j, \nu^{n-j}).$$

If I_K is convex at μ, then we may take the limit as t goes to 0, which gives us

$$I_K(\nu) + (n-1)I_K(\mu) \geq nI_K(\mu^{n-1}, \nu).$$

■

Observe that if I_K is convex (in particular, convex at ν), switching the roles of μ and ν we obtain

$$(n-1)I_K(\nu) + I_K(\mu) \geq nI_K(\mu, \nu^{n-1}).$$

Therefore, in the case $n = 2, 3$, Lemma 5.3.4 provides the converse of Proposition 5.3.3, in other words, I_K is convex if and only if it satisfies the arithmetic mean inequalities (5.21). We are now ready to demonstrate the equivalence of the conditional positive definiteness of K and the convexity of K for the two-input case.

Proposition 5.3.5. *Suppose $K : \Omega^2 \to \mathbb{R}$ is continuous and symmetric. Then K is conditionally positive definite if and only if I_K is convex.*

Proof. Corollary 5.3.3.1 gives us one direction. For the other, assume that I_K is convex. Let $\mu \in \mathcal{M}(\Omega)$ satisfy $\mu(\Omega) = 0$. Then there exist $\mu_+, \mu_- \in \mathbb{P}(\Omega)$ and some constant $c \geq 0$ such that $\mu = c(\mu_+ - \mu_-)$. Lemma 5.3.4 with $n = 2$ implies that $I_K(\mu_+, \mu_-) \leq \frac{1}{2}\big(I_K(\mu_+) + I_K(\mu_-)\big)$ and therefore

$$I_K(\mu) = c^2 \Big(I_K(\mu_+) - 2I_K(\mu_+, \mu_-) + I_K(\mu_-)\Big) \geq 0,$$

i.e. K is conditionally positive definite. ■

It is not completely clear whether this equivalence holds for $n \geq 3$, but evidence suggests that it doesn't. Indeed, Proposition 5.6.5 provides an example of a three-input kernel with $\Omega = \mathbb{S}^{d-1}$, which is not conditionally 3-positive definite, but at the same time the energy functional is convex at σ (although we don't know if it is convex at *all* measures in $\mathbb{P}(\mathbb{S}^{d-1})$) and is minimized by σ.

In this regard, we would also like to point out that a number of our results about energy minimizers do not require the full power of convexity of I_K on $\mathbb{P}(\Omega)$, but rather just the convexity at the presumptive minimizer μ. In particular, condition (5.26), which appears in Theorems 5.4.2 and 5.4.5, is implied by inequality (5.23) of Lemma 5.3.4, and hence it holds if I_K is convex at μ.

Using convexity of the energy functional, one can draw a connection between minimizing the n-input energy I_K and the $(n-1)$-input energy $I_{U_K^\mu}$, thus obtaining our first result about minimizers of multi-input energies.

Proposition 5.3.6. *Let $n \geq 3$. Assume that $K : \Omega^n \to \mathbb{R}$ is continuous and symmetric, I_K is convex, and that $\mu \in \mathbb{P}(\Omega)$ is a minimizer of $I_{U_K^\mu}$. Then μ is a minimizer of I_K.*

Proof. We first prove that if the energy I_K is convex and μ, $\nu \in \mathbb{P}(\Omega)$, then

$$I_K(\nu) - I_K(\mu) \geq \frac{n}{n-1}\Big(I_{U_K^\mu}(\nu) - I_{U_K^\mu}(\mu)\Big). \tag{5.24}$$

Indeed, we have $I_{U_K^\mu}(\mu) = I_K(\mu)$ and, by Lemma 5.3.4, $I_{U_K^\mu}(\nu) = I_K(\mu, \nu^{n-1}) \leq \frac{1}{n} I_K(\mu) + \frac{n-1}{n} I_K(\nu)$. Thus,

$$\begin{aligned} I_K(\nu) - I_K(\mu) - n\Big(I_{U_K^\mu}(\nu) - I_{U_K^\mu}(\mu)\Big) &= I_K(\nu) - nI_K(\mu, \nu^{n-1}) + (n-1)I_K(\mu) \\ &\geq (n-2)\Big(I_K(\mu) - I_K(\nu)\Big), \end{aligned}$$

which implies inequality (5.24).

Inequality (5.24), together with the fact that μ is a minimizer of $I_{U_K^\mu}$, implies that for all $\nu \in \mathbb{P}(\Omega)$, we have

$$I_K(\nu) - I_K(\mu) \geq \frac{n}{n-1}\Big(I_{U_K^\mu}(\nu) - I_{U_K^\mu}(\mu)\Big) \geq 0,$$

hence μ minimizes I_K. ■

Proposition 5.3.6 can be viewed as a precursor of some of our more advanced results from Section 5.4 which show that there is a strong relation between μ minimizing the n-input energy I_K and the energy functional $I_{U_K^{\mu^k}}$ with a lower number of inputs. In fact, Theorem 5.4.5 contains Proposition 5.3.6 as a special case. We have nevertheless decided to include this proposition, as it admits a very transparent and elementary proof, which also provides a quantitative relation between the minimization of I_K and $I_{U_K^\mu}$.

5.4. Minimizers of the energy functional

We finally turn to some of the general results about minimizers of energies with multivariate kernels. In the classical two-input case, properties of minimizing measures are closely related to their potentials, see e.g. [Bjö56; BHS19]. Direct analogues of such statements can be obtained for multi-input energies. We start with the necessary condition, which states that the potential of a minimizer is constant on its support. As before, in all of the statements of this section we assume that $K : \Omega^n \to \mathbb{R}$ is continuous and symmetric, even if not explicitly stated.

Theorem 5.4.1. *Let $K : \Omega^n \to R$ be continuous and symmetric. Suppose that μ is a minimizer of I_K over $\mathbb{P}(\Omega)$. Then $U_K^{\mu^{n-1}}(x) = I_K(\mu)$ on* $\mathrm{supp}(\mu)$ *and $U_K^{\mu^{n-1}}(x) \geq I_K(\mu)$ on Ω.*

Proof. The proof is a simple extension of the proof of Theorem 2 in [Bjö56], and we include it for the sake of completeness. Let $\nu \in \mathcal{M}(\Omega)$ be such that $\nu(\Omega) = 0$ and $\mu(A) + \varepsilon\nu(A) \geq 0$ for all Borel subsets $A \subseteq \Omega$ and $0 \leq \varepsilon \leq 1$. This clearly means that $\mu + \varepsilon\nu \in \mathbb{P}(\Omega)$, so

$$I_K(\mu) \leq I_K(\mu + \varepsilon\nu) = \sum_{k=0}^{n} \binom{n}{k} \varepsilon^k I_K(\mu^{n-k}, \nu^k).$$

Thus, for $0 \leq \varepsilon \leq 1$,

$$0 \leq \varepsilon \left(\sum_{k=1}^{n} \binom{n}{k} \varepsilon^{k-1} I_K(\mu^{n-k}, \nu^k) \right).$$

This means that $I_K(\mu^{n-1}, \nu) \geq 0$.

Suppose, indirectly, that there exist $a, b \in \mathbb{R}$, $z \in \mathrm{supp}(\mu)$ and $y \in \Omega$ such that

$$a = U_K^{\mu^{n-1}}(z) > U_K^{\mu^{n-1}}(y) = b.$$

Let B be a ball centered at z, small enough so that $y \notin B$ and oscillation of $U_K^{\mu^{n-1}}(x)$ is at most $\frac{a-b}{2}$, and let $m = \mu(B)$. Let ν be defined by

$$\nu(A) = m\delta_y(A) - \mu(A \cap B). \tag{5.25}$$

Then

$$I_K(\mu^{n-1}, \nu) = U_K^{\mu^{n-1}}(y) \cdot m - \int_B U_K^{\mu^{n-1}}(x) d\mu(x) \leq bm - \left(a - \frac{a-b}{2} \right) m < 0,$$

which is a contradiction. Thus, if $U_K^{\mu^{n-1}}(z) = a$ for some $z \in \mathrm{supp}(\mu)$, then $U_K^{\mu^{n-1}}(x) \geq a$ for all $x \in \Omega$. Our claim then follows. ■

Definition 5.4.1. We shall say that μ is a *local minimizer* of I_K if it is a local minimizer in every direction, in other words, if for each $\nu \in \mathbb{P}(\Omega)$, there exists $t_\nu \in (0, 1]$ such that for all $t \in [0, t_\nu]$ we have

$$I_K\big((1-t)\mu + t\nu\big) \geq I_K(\mu).$$

Observe that this definition differs from the definition of local minimizers with respect to some metric, such as the Wasserstein d_∞ metric or the total variation norm (the difference is similar to that between the Gateaux and Fréchet derivatives).

Analyzing the proof of Theorem 5.4.1, we find that for ν defined in (5.25), we can write $\mu+\varepsilon\nu = (1-\varepsilon)\mu+\varepsilon\widetilde{\nu}$ with $\widetilde{\nu} = \mu+\nu \in \mathbb{P}(\Omega)$. Hence, one arrives at a contradiction even if μ is just a local minimizer.

Corollary 5.4.1.1. *The statement of Theorem 5.4.1 remains true if we only assume that μ is a local (not global) minimizer of I_K.*

In general, the converse to Theorem 5.4.1 is not true. However, with some additional convexity assumptions, the necessary condition also becomes sufficient.

Theorem 5.4.2. *Let $K : \Omega^n \to \mathbb{R}$ be symmetric and continuous. Suppose that for some $\mu \in \mathbb{P}(\Omega)$, there exists a finite constant M such that $U_K^{\mu^{n-1}}(x) \geq M$ on Ω and $U_K^{\mu^{n-1}}(x) = M$ on $\operatorname{supp}(\mu)$. Suppose further that for all $\nu \in \mathbb{P}(\Omega)$, there exists some $\alpha \in (0,1)$, possibly depending on ν, such that*

$$I_K(\mu^{n-1},\nu) \leq \alpha I_K(\nu) + (1-\alpha) I_K(\mu). \tag{5.26}$$

Then μ is a minimizer of I_K.

Proof. For any $\nu \in \mathbb{P}(\Omega)$, for some $\alpha \in (0,1)$, we have

$$I_K(\mu) = \int_\Omega U_K^{\mu^{n-1}}(x)d\mu(x) \leq \int_\Omega U_K^{\mu^{n-1}}(x)d\nu(x) = I_K(\mu^{n-1},\nu) \leq \alpha I_K(\nu)+(1-\alpha)I_K(\mu),$$

hence $I_K(\mu) \leq I_K(\nu)$. ■

Some remarks concerning the assumptions of Theorem 5.4.2, i.e. condition (5.26), are in order. Due to Lemma 5.3.4, convexity of the energy functional I_K at μ implies condition (5.26) with $\alpha = \frac{1}{n}$. In turn, if K is conditionally n-positive definite, Corollary 5.3.3.1 states that I_K is convex, and hence again condition (5.26) is satisfied (alternatively, Lemma 5.3.2 shows directly that conditional n-positive definiteness of K implies the convexity condition (5.26) of Theorem 5.4.2 with $\alpha = \frac{1}{n}$). The hierarchy of these conditions can be summarized in the following diagram:

$$\begin{aligned} &K \text{ is } n\text{-positive definite} \implies K \text{ is conditionally } n\text{-positive definite} \implies \\ &\implies I_K \text{ is convex} \implies I_K \text{ is convex at } \mu \implies \text{condition (5.26) holds.} \end{aligned} \tag{5.27}$$

Therefore, Theorem 5.4.2 (as well as other statements relying on (5.26), e.g. Lemma 5.4.3 or Theorem 5.4.5) may be applied under the assumptions that K is (conditionally) n-positive definite or that I_K is convex at μ.

We also make the following remark: in the case when μ has full support, i.e. $\operatorname{supp}(\mu) = \Omega$, if the first condition of Theorem 5.4.2 holds, i.e. $U_K^{\mu^{n-1}}(x) = M$ for all $x \in \Omega$, then

$I_K(\mu^{n-1}, \nu) = I_K(\mu)$, and the assumption (5.26) is obviously the same as the conclusion of Theorem 5.4.2. This does not, however, render this case of the theorem useless – on the contrary, if one replaces (5.26) with one of the stronger conditions in (5.27), one obtains an interesting and meaningful statement. (This shows that the most of the content is hidden in the implications presented in (5.27).) We summarize this case in a separate corollary, as it will be of use later.

Corollary 5.4.2.1. *Let $K : \Omega^n \to \mathbb{R}$ be symmetric and continuous. Suppose that $\mu \in \mathbb{P}(\Omega)$ has full support* $\operatorname{supp}(\mu) = \Omega$ *and that there exists a constant M such that $U_K^{\mu^{n-1}}(x) = M$ on Ω. Assume also that any of the conditions in* (5.27) *holds (e.g., K is n-positive definite or I_K is convex). Then μ is a minimizer of I_K.*

We also observe that Corollary 5.4.1.1 and Theorem 5.4.2 imply the following local-to-global principle for minimizers of I_K under convexity assumptions.

Lemma 5.4.3. *Let $n \geq 2$ and let μ be a local minimizer of the energy functional I_K. Assume also that condition* (5.26) *is satisfied. Then μ is a global minimizer of I_K over $\mathbb{P}(\Omega)$.*

Proof. Corollary 5.4.1.1 shows that the first condition of Theorem 5.4.2 holds. Together with condition (5.26), this implies that μ is a global minimizer of I_K. ■

Naturally, the set of minimizers of a convex functional is convex. By Corollary 5.3.3.1, for conditionally n-positive definite kernels, the energy I_K is convex, i.e. minimizers of I_K form a convex set in this case.

Proposition 5.4.4. *Let K be a conditionally n-positive definite kernel. Then the set of minimizers of the energy I_K is convex.*

While Theorems 5.4.1 and 5.4.2 are straightforward generalizations of the corresponding facts for the classical two-input energies, they lead to some interesting consequences for energies with multivariate kernels. In particular, we start by showing that under condition (5.26), if μ minimizes the lower input energy with the kernel $U_K^{\mu^k}$, then it also minimizes the original n-input energy I_K.

Theorem 5.4.5. *Let $K : \Omega^n \to \mathbb{R}$, $n \geq 3$, be symmetric and continuous. Assume that for some $1 \leq k \leq n-2$, the measure $\mu \in \mathbb{P}(\Omega)$ (locally) minimizes the $(n-k)$-input energy $I_{U_K^{\mu^k}}$. Assume also that μ satisfies condition* (5.26) *of Theorem 5.4.2. Then μ minimizes the n-input energy I_K.*

Proof. Theorem 5.4.1 (or Corollary 5.4.1.1) applied to the kernel $U_K^{\mu^k}$ implies that for all $x \in \Omega$

$$U_K^{\mu^{n-1}}(x) = U_{U_K^{\mu^k}}^{\mu^{n-k-1}}(x) \geq I_{U_K^{\mu^k}}(\mu) = I_K(\mu)$$

with equality for $x \in \operatorname{supp}(\mu)$. Condition (5.26) then allows one to invoke Theorem 5.4.2, which shows that μ minimizes I_K. ■

The converse to Theorem 5.4.5 holds for $k = n - 2$ even without any convexity assumptions: in this case, if μ locally minimizes I_K, it also locally minimizes the two-input energy $I_{U_K^{\mu^{n-2}}}$. Moreover, under the additional condition that μ has full support, one can deduce that the measure μ is a *global* minimizer of $I_{U_K^{\mu^{n-2}}}$, see parts (1)-(2) of Theorem 5.4.6 below. Furthermore, this implication may be reversed, if one additionally assumes that μ *uniquely* minimizes $I_{U_K^{\mu^{n-2}}}$. Observe that, unlike Theorem 5.4.5, part (3) of Theorem 5.4.6 does not require any of the conditions of (5.27) and, unlike part (2), it does not require the condition $\mathrm{supp}(\mu) = \Omega$.

Theorem 5.4.6. *Let $K : \Omega^n \to \mathbb{R}$, $n \geq 3$, be symmetric and continuous and let $\mu \in \mathbb{P}(\Omega)$.*

1. *Let μ be a local minimizer of I_K. Then μ is a local minimizer of the two-input energy $I_{U_K^{\mu^{n-2}}}$.*
2. *Let μ be a local minimizer of I_K and assume, in addition, that μ has full support, i.e. $\mathrm{supp}(\mu) = \Omega$. Then μ minimizes the two-input energy $I_{U_K^{\mu^{n-2}}}$ over $\mathbb{P}(\Omega)$.*
3. *If μ is the unique minimizer of $I_{U_K^{\mu^{n-2}}}$ in $\mathbb{P}(\Omega)$, then μ is a local minimizer of I_K.*

Proof. Fix an arbitrary measure $\nu \in \mathbb{P}(\Omega)$. For $t \in [0, 1]$, let us define two functions $g_\nu(t) = I_K\big((1-t)\mu + t\nu\big)$ and $h_\nu(t) = I_{U_K^{\mu^{n-2}}}\big((1-t)\mu + t\nu\big) = I_K\big(\mu^{n-2}, \big((1-t)\mu + t\nu\big)^2\big)$. We have

$$g_\nu(t) = (1-t)^n I_K(\mu) + nt(1-t)^{n-1} I_K(\mu^{n-1}, \nu) + \binom{n}{2} t^2 (1-t)^{n-2} I_K(\mu^{n-2}, \nu^2) + R_\nu(t),$$

where each term in $R_\nu(t)$ contains a factor of the form t^k with $k \geq 3$ and, therefore, $R_\nu'(0) = R_\nu''(0) = 0$,

$$h_\nu(t) = (1-t)^2 I_K(\mu) + 2t(1-t) I_K(\mu^{n-1}, \nu) + t^2 I_K(\mu^{n-2}, \nu^2).$$

A direct (elementary, but lengthy) computation, which we omit, shows that

$$h_\nu'(0) = \frac{2}{n} g_\nu'(0) = 2\big(I_K(\mu^{n-1}, \nu) - I_K(\mu)\big), \tag{5.28}$$

$$h_\nu''(0) = \frac{2}{n(n-1)} g_\nu''(0) = 2\big(I_K(\mu) - 2I_K(\mu^{n-1}, \nu) + I_K(\mu^{n-2}, \nu^2)\big). \tag{5.29}$$

We now start by proving (1). Let μ be a local minimizer of I_K. According to Corollary 5.4.1.1, we have that $U_K^{\mu^{n-1}}(x) \geq I_K(\mu)$ on Ω and therefore, $I_K(\mu^{n-1}, \nu) \geq I_K(\mu)$ for any $\nu \in \mathbb{P}(\Omega)$. Since g_ν has a local minimum at $t = 0$, either $g_\nu'(0) > 0$, or $g_\nu'(0) = 0$ and $g_\nu''(0) \geq 0$. In the first case, we also have $h_\nu'(0) > 0$. In the second case, $h_\nu'(0) = 0$ and $h_\nu''(0) \geq 0$, and since h_ν is quadratic, this implies that $h_\nu(t) = at^2 + b$ with $a \geq 0$. Thus, h_ν has a local minimum at $t = 0$ for each $\nu \in \mathbb{P}(\Omega)$, i.e. μ is a local minimizer of $I_{U_K^{\mu^{n-2}}}$.

If in addition μ has full support, then Corollary 5.4.1.1 implies that for any $\nu \in \mathbb{P}(\Omega)$, we have $I_K(\mu^{n-1}, \nu) = I_K(\mu)$. Therefore, relations (5.28)-(5.29), together with the fact that g_ν has a local minimum at $t = 0$, show that $g'_\nu(0) = 0$, hence $g''_\nu(0) \geq 0$, and at the same time

$$g''_\nu(0) = n(n-1)\big(I_K(\mu^{n-2}, \nu^2) - I_K(\mu)\big) = n(n-1)\Big(I_{U_K^{\mu^{n-2}}}(\nu) - I_{U_K^{\mu^{n-2}}}(\mu)\Big). \quad (5.30)$$

Hence, μ is a global minimizer of $I_{U_K^{\mu^{n-2}}}$, which proves part (2).

To prove (3), assume that μ is the unique global minimizer of $I_{U_K^{\mu^{n-2}}}$. Observe that, since the potential of $U_K^{\mu^{n-2}}$ with respect to μ is $U_K^{\mu^{n-1}}$, Theorem 5.4.1 applied to $U_K^{\mu^{n-2}}$ implies that, just like in part (1), we have $I_K(\mu^{n-1}, \nu) \geq I_K(\mu)$. Thus, $g'_\nu(0) \geq 0$ by (5.28). If $g'_\nu(0) > 0$, there is a local minimum at $t = 0$. If, however, $g'_\nu(0) = 0$, then $I_K(\mu^{n-1}, \nu) = I_K(\mu)$ and relation (5.30) holds. Since μ uniquely minimizes $I_{U_K^{\mu^{n-2}}}$, this proves that $g''_\nu(0) > 0$ for $\nu \neq \mu$. Hence, in each case, g_ν has a local minimum at $t = 0$, i.e. μ is a local minimizer of I_K. ■

For classical pairwise interaction energies, it is well known that the kernel is conditionally positive definite on the support of the minimizer (see, e.g., [FS13]), therefore, we obtain the following corollary to part (2) Theorem 5.4.6:

Corollary 5.4.6.1. *Assume that $\mu \in \mathbb{P}(\Omega)$ with* $\operatorname{supp}(\mu) = \Omega$ *is a local minimizer of I_K. Then the $(n-2)$-fold potential of K with respect to μ, i.e. the two-variable function $U_K^{\mu^{n-2}}(x, y)$, is conditionally positive definite on Ω.*

Observe that, if the kernel K is conditionally n-positive definite, then, according to Lemma 5.2.1, $U_K^{\mu^{n-2}}(x, y)$ is conditionally positive definite. Moreover, Theorem 5.4.5 applies for conditionally positive definite kernels K. Therefore, the statement of Corollary 5.4.6.1 may be viewed as a partial converse of Theorem 5.4.5 for conditionally positive definite kernels. This interplay will manifest itself in an even stronger fashion on the sphere, the situation to be explored in Section 5.5.

5.5. Multi-input energy on the sphere

We now restrict our attention to the case when Ω is the unit sphere, i.e. $\Omega = \mathbb{S}^{d-1} \subset \mathbb{R}^d$, where the symmetries and structure of the domain allow one to deduce additional information about energy minimization.

We shall denote by σ the normalized uniform surface measure on the sphere. One of the most natural questions is whether σ minimizes the energy functional over $\mathbb{P}(\mathbb{S}^{d-1})$, or, in other words, whether energy minimization induces uniform distribution.

In this section, we shall be interested in kernels, which (in addition to being continuous and symmetric) are *rotationally invariant*, i.e. have the form

$$K(x_1, \ldots, x_n) = F\Big(\big(\langle x_i, x_j \rangle\big)_{i,j=1}^n\Big), \quad (5.31)$$

in other words, they depend only on the Gram matrix of $\{x_1, \dots, x_n\} \subset \mathbb{S}^{d-1}$.

When $n = 2$, one obtains classical pairwise interaction kernels of the form $K(x, y) = F(\langle x, y\rangle)$. The theory of both discrete and continuous energies with such kernels on the sphere is very rich and goes back at least to Schoenberg [Sch41].

In the case $n = 3$ rotationally invariant kernels are functions of the form

$$K(x, y, z) = F(\langle x, y\rangle, \langle y, z\rangle, \langle z, x\rangle) = F(u, v, t), \tag{5.32}$$

where we set $u = \langle x, y\rangle$, $v = \langle y, z\rangle$, $t = \langle z, x\rangle$, and we shall keep this notation throughout the text (the slightly non-alphabetic order is inherited from [CW12]).

Observe that, if the n-input kernel K is rotationally invariant, its potential with respect to σ is again rotationally invariant. Indeed, for any $V \in SO(d)$, we have

$$U_K^\sigma(Vx_1, \dots, Vx_{n-1}) = U_K^\sigma(x_1, \dots, x_{n-1}), \tag{5.33}$$

which easily follows from (5.31) and the facts that $\langle x_i, x_j\rangle = \langle Vx_i, Vx_j\rangle$ and $\langle Vx_i, x_n\rangle = \langle x_i, V^{-1}x_n\rangle$, $1 \le i, j \le n-1$, together with the rotational invariance of σ, i.e. $d\sigma(x_n) = d\sigma(V^{-1}x_n)$. Iterating this observation, one finds that all k-fold potentials of K with respect to σ, i.e. functions $U_K^{\sigma^k}$ with $1 \le k \le n-1$, are rotationally invariant. In particular, when $k = n-2$, the two-input kernel $U_K^{\sigma^{n-2}}$ depends only on the inner product of the inputs, and for $k = n-1$, the potential $U_K^{\sigma^{n-1}}$ is just a constant:

$$U_K^{\sigma^{n-2}}(x, y) = G(\langle x, y\rangle) = G(u) \quad \text{and} \quad U_K^{\sigma^{n-1}}(x) = \text{const} = I_K(\sigma). \tag{5.34}$$

Recall that Theorem 5.4.1 would guarantee the latter condition in the case when σ is a minimizer of I_K. However, for rotationally invariant kernels, this is automatically satisfied, which facilitates the application of the results of Section 5.4 and will play an important role later, in Theorem 5.5.1.

Turning to the primary task of understanding when σ minimizes I_K, we first remind ourselves that in the classical case of a two-input energy with a rotationally invariant kernel $G(\langle x, y\rangle)$ on $\mathbb{S}^{d-1}$, the answer to this question is well understood. In particular, the following three conditions are equivalent, see e.g. [BDM18]:

1. The uniform surface measure σ minimizes I_G over $\mathbb{P}(\mathbb{S}^{d-1})$.
2. The kernel G is conditionally positive definite on $\mathbb{S}^{d-1}$.
3. The kernel G is positive definite on $\mathbb{S}^{d-1}$ up to a constant term, i.e. there exists a constant $c \in \mathbb{R}$ such that $G + c$ is positive definite on $\mathbb{S}^{d-1}$ (in fact, one can take $c = -I_G(\sigma)$).

Our goal is to generalize these statements (at least partially) to the case of multi-input energies. We observe that, if a symmetric rotationally invariant kernel K is conditionally n-positive definite on $\mathbb{S}^{d-1}$, then, according to Lemma 5.2.1, the potential

$G(u) = U_K^{\sigma^{n-2}}(x,y)$ is also conditionally positive definite, and hence, by the discussion above, σ is a minimizer of the two-input energy $I_{U_K^{\sigma^{n-2}}}$. Therefore, since conditionally n-positive definite kernels satisfy condition (5.26), Theorem 5.4.5 with $k = n-2$ applies and we obtain the following statement:

Theorem 5.5.1. *Suppose that $K : (\mathbb{S}^{d-1})^n \to \mathbb{R}$ is continuous, symmetric, rotationally invariant, and conditionally n-positive definite on $\mathbb{S}^{d-1}$. Then σ is a minimizer of I_K over $\mathbb{P}(\Omega)$.*

This theorem also easily follows from Theorem 5.4.2 and the remarks thereafter (or, more precisely, from Corollary 5.4.2.1), since, as explained above, the potential $U_K^{\sigma^{n-1}}$ is constant on $\mathbb{S}^{d-1}$.

Notice that, unlike some statements of Section 5.4, e.g. Theorem 5.4.5, for rotationally invariant kernels in the theorem above one does not need to assume anything about energies with a lower number of inputs – conditional positive definiteness alone suffices.

Theorem 5.5.1 immediately yields some interesting examples:

Corollary 5.5.1.1. *Let $f : [-1,1] \to \mathbb{R}$ be a real-analytic function with nonnegative Taylor coefficients and let $F(u,v,t) = f(uvt)$. Then, for K defined as in* (5.32)*, the uniform surface measure σ minimizes the energy I_K over $\mathbb{P}(\mathbb{S}^{d-1})$.*

Proof. Observe first that in this setup, if K_z is positive definite for one point $z \in \mathbb{S}^{d-1}$, it is also positive definite for each $z \in \mathbb{S}^{d-1}$ due to rotational invariance, i.e. Definition 5.2.2 only needs to be checked at one point. Consider first $F(u,v,t) = uvt$ and fix any $z \in \mathbb{S}^{d-1}$, e.g., $z = e_1$. Then for any $\nu \in \mathcal{M}(\mathbb{S}^{d-1})$,

$$I_{K_{e_1}}(\nu) = \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \langle x, y\rangle x_1 y_1 d\nu(x) d\nu(y) = \sum_{i=1}^{d} \left(\int_{\mathbb{S}^{d-1}} x_1 x_i \, d\nu(x) \right)^2 \geq 0,$$

i.e. the kernel $K(x,y,z) = \langle x, y\rangle \langle y, z\rangle \langle z, x\rangle = uvt$ is 3-positive definite, and hence, by Lemma 5.2.2, so are all of its integer powers, positive linear combinations and their limits. The conclusion now follows from Theorem 5.5.1. ∎

This Corollary provides a whole array of examples: for instance, three-input energies with kernels $K(x,y,z) = uvt$, or $(uvt)^n$, or e^{uvt} are all minimized by σ. We remark that, while for $K = uvt$ this statement could be proved using semi-definite programing, for higher powers $(uvt)^n$ this would be extremely difficult technically, and for kernels like e^{uvt} almost impossible.

For even exponents, the energies with the kernels $K = (uvt)^{2k}$ can be viewed as three-input generalizations of the well-known p-frame potentials [Bil+21c; EO12], which are closely related to tight frames and projective designs [**SG**; BF03]. We also point out that Proposition 5.6.2 provides a more general class of n-positive definite kernels, which contains $K = uvt$ as a special case.

Unfortunately, unlike the classical two-input case, the converse to Theorem 5.5.1 is not true: Propositions 5.6.5, 5.6.9, and 5.6.10 show that some kernels, naturally arising in semidefinite programming and geometry, fail to be conditionally n-positive definite, even though σ minimizes corresponding energies (see Theorems 5.6.6 and 5.6.7). In other words, conditional n-positive definiteness of the kernel is not equivalent to the fact that σ minimizes the energy.

We suspect that the property that σ minimizes I_K is equivalent to the fact that $U_K^{\sigma^{n-2}}$ is conditionally positive definite, i.e. the two-input energy $I_{U_K^{\sigma^{n-2}}}$ is minimized by σ. This conjecture is supported by all the examples known to us. Conditional positive-definiteness of $U_K^{\sigma^{n-2}}$ obviously follows from conditional n-positive definiteness of K, due to Lemma 5.2.1, but the converse implication is not true, see e.g. Proposition 5.6.5. In fact, all the kernels discussed in Section 5.6.3 (Propositions 5.6.5, 5.6.9, and 5.6.10) possess this property: they are not 3-positive definite, but their potentials U_K^{σ} with respect to σ are (conditionally) positive definite, and the correspondig energies I_K are minimized by σ.

Theorem 5.5.2 below (which is essentially a restatement of Theorem 5.4.6 for the spherical case, along with the fact that σ has full support) shows that conditional positive definiteness of $U_K^{\sigma^{n-2}}$ is implied if σ is a *local* minimizer of I_K, and a partial converse to this statement also holds. Observe that, if the conjecture above is true, then being a local and global minimizer are equivalent for σ: this fact is indeed true for the two-input energies, see [Bil+21c].

Theorem 5.5.2. *Let $K : (\mathbb{S}^{d-1})^n \to \mathbb{R}$ be a continuous, symmetric, and rotationally invariant kernel.*

1. *Assume that σ is a local minimizer of I_K in $\mathbb{P}(\mathbb{S}^{d-1})$. Then the uniform measure σ is a global minimizer of the two-input energy $I_{U_K^{\sigma^{n-2}}}$, or, equivalently, $U_K^{\sigma^{n-2}}$ is conditionally positive definite on the sphere $\mathbb{S}^{d-1}$.*

2. *Assume that σ is the unique global minimizer of $I_{U_K^{\sigma^{n-2}}}$ over $\mathbb{P}(\mathbb{S}^{d-1})$. Then σ is a local minimizer of the n-input energy I_K.*

Theorem 5.5.2 above shows that if σ is a global minimizer of I_K, then the potential $U_K^{\sigma^{n-2}}$ is conditionally positive definite. We do not know whether the converse of this statement holds. One can show, however, at least for $n = 3$ that if σ minimizes $I_{U_K^{\sigma}}$, but fails to minimize I_K, then the minimizer of I_K cannot be supported on the whole sphere.

Lemma 5.5.3. *Let $K : (\mathbb{S}^{d-1})^3 \to \mathbb{R}$ be a continuous, symmetric, and rotationally invariant three-input kernel. Assume that U_K^{σ} is conditionally positive definite on the sphere $\mathbb{S}^{d-1}$ (i.e. σ minimizes $I_{U_K^{\sigma}}$), but at the same time σ is not a minimizer of I_K. Let μ be a minimizer of I_K. Then $\mathrm{supp}(\mu) \subsetneq \mathbb{S}^{d-1}$.*

Proof. Assume, by contradiction, that $\mathrm{supp}(\mu) = \mathbb{S}^{d-1}$. Then, by Theorem 5.4.1, $U_K^{\mu^2}(x) = I_K(\mu)$ for every $x \in \mathbb{S}^{d-1}$, and therefore,

$$I_{U_K^{\sigma}}(\mu) = I_K(\mu, \mu, \sigma) = \int_{\mathbb{S}^{d-1}} U_K^{\mu^2}(x)\, d\sigma(x) = I_K(\mu).$$

On the other hand, obviously, $I_K(\sigma) = I_{U_K^{\sigma}}(\sigma)$. Since μ is a minimizer of I_K, and σ is not, we have $I_K(\mu) < I_K(\sigma)$. This implies that $I_{U_K^{\sigma}}(\mu) < I_{U_K^{\sigma}}(\sigma)$, which contradicts the conditional positive definiteness of U_K^{σ}. ■

5.6. Positive definite kernels

Corollary 5.5.1.1 of the previous section already provided a class of 3-positive definite functions. In this section we provide several other classes of kernels that are (conditionally) n-positive definite.

5.6.1. General classes of (conditionally) n-positive definite kernels

We start with some very natural examples, which show how to construct (conditionally) n-positive definite kernels from kernels with fewer inputs. In particular, we show that an n-input kernel can be constructed from m-input ones, $m < n$, by considering the sum or product over all m-element subsets of inputs. We first deal with the statement about the sum.

Proposition 5.6.1. *Let $2 \leq m \leq n-1$, and suppose $H : \Omega^m \to \mathbb{R}$ is continuous, symmetric, and conditionally m-positive definite. Then*

$$K(z_1, ..., z_n) := \sum_{1 \leq j_1 < j_2 < \cdots < j_m \leq n} H(z_{j_1}, z_{j_2}, ..., z_{j_m})$$

is conditionally n-positive definite.

Proof. Let ν be a finite signed Borel measure on Ω such that $\nu(\Omega) = 0$. Then for any fixed $z_1, ..., z_{n-2} \in \Omega$, since H is conditionally m-positive definite, we have

$$\begin{aligned}
&\int_\Omega \int_\Omega K(z_1, ..., z_{n-2}, x, y) d\nu(x) d\nu(y) \\
&\quad = \int_\Omega \int_\Omega \sum_{1 \leq j_1 < \cdots < j_{m-2} \leq n-2} H(z_{j_1}, ..., z_{j_{m-2}}, x, y) d\nu(x) d\nu(y) \\
&\qquad + \int_\Omega \int_\Omega \sum_{1 \leq k_1 < \cdots < k_{m-1} \leq n-2} \Big(H(z_{k_1}, ..., z_{k_{m-1}}, x) + H(z_{k_1}, ..., z_{k_{m-1}}, y) \Big) d\nu(x) d\nu(y) \\
&\qquad + \int_\Omega \int_\Omega \sum_{1 \leq l_1 < \cdots < l_m \leq n-2} H(z_{l_1}, ..., z_{l_m}) d\nu(x) d\nu(y) \\
&\quad = \sum_{1 \leq j_1 < \cdots < j_{m-2} \leq n-2} \int_\Omega \int_\Omega H(z_{j_1}, ..., z_{j_{m-2}}, x, y) d\nu(x) d\nu(y) \geq 0,
\end{aligned}$$

which shows that K is conditionally n-positive definite. ■

We can also prove an analogue of Proposition 5.6.1 for products of positive definite functions.

Proposition 5.6.2. *Let $2 \leq m \leq n-1$ and assume that $H : \Omega^m \to \mathbb{R}$ is continuous, symmetric, and m-positive definite. If H is a nonnegative function or $m = n-1$, then*

$$K(z_1, ..., z_n) = \prod_{1 \leq j_1 < \cdots < j_m \leq n} H(z_{j_1}, ..., z_{j_m})$$

is n-positive definite.

Proof. Fix $z_1, \dots, z_{n-2} \in \Omega$. We can write

$$K(z_1, \dots, z_{n-2}, x, y) = \prod_{1 \le j_1 < \dots < j_m \le n-2} H(z_{j_1}, \dots, z_{j_m}) \tag{5.35}$$

$$\times \prod_{1 \le j_1 < \dots < j_{m-1} \le n-2} H(z_{j_1}, \dots, z_{j_{m-1}}, x) \tag{5.36}$$

$$\times \prod_{1 \le j_1 < \dots < j_{m-1} \le n-2} H(z_{j_1}, \dots, z_{j_{m-1}}, y) \tag{5.37}$$

$$\times \prod_{1 \le j_1 < \dots < j_{m-2} \le n-2} H(z_{j_1}, \dots, z_{j_{m-2}}, x, y). \tag{5.38}$$

Observe that the product in line (5.35) is non-negative when $H \ge 0$ or if $m = n-1$ (the product is empty in the latter case). The product of lines (5.36) and (5.37) is positive definite as a function of x and y: indeed, it has the form $F(x, y) = \phi(x)\phi(y)$ and hence

$$I_F(\mu) = \left(\int_\Omega \phi(x) d\mu(x) \right)^2 \ge 0$$

for any $\mu \in (\Omega)$. Finally, every factor in the product in line (5.38) is positive definite as a function of x and y, because H is m-positive definite. Thus, Schur's product theorem (see Lemma 5.2.2) ensures that the whole product is positive definite as a function of x and y, therefore, K is n-positive definite. ■

Propositions 5.6.1 and 5.6.2 provide us with large classes of n-positive definite kernels. However, these constructions do not exhaust all such kernels. In the following subsection, we provide examples of three-positive definite kernels, which are not obtained from two-input kernels by the methods described above.

5.6.2. Three-positive definite kernels on the sphere

We also provide some examples of kernels on the unit sphere $\mathbb{S}^{d-1}$. We use the same notation as in Section 5.5: for $x, y, z \in \mathbb{S}^{d-1}$, we set $u = \langle x, y \rangle$, $v = \langle y, z \rangle$, and $t = \langle z, x \rangle$.

In Corollary 5.5.1.1, we showed that $K = uvt$ is 3-positive definite on the sphere. Observe that this is a specific case of Proposition 5.6.2 above, since $\langle x, y \rangle$ is a positive definite function on $\mathbb{S}^{d-1}$. More generally, Proposition 5.6.2 implies that any kernel of the form $K(x, y, z) = h(u)h(v)h(t)$ is 3-positive definite, as long as h is a positive definite function on the sphere.

The kernels considered in Lemmas 5.6.3 and 5.6.4 are closely related to the parallelepiped spanned by the vectors x, y, and $z \in \mathbb{S}^{d-1}$. Indeed, setting $a = 2$ in (5.39), one obtains negative volume squared of this parallelepiped: this kernel is not conditionally 3-positive definite according to Proposition 5.6.9, even though σ is a minimizer of the corresponding energy, as shown in Theorem 5.6.6. However, positive definiteness does hold for other values of the parameter a.

Lemma 5.6.3. *For $a < 1$,*

$$K(x,y,z) = t^2 + u^2 + v^2 - auvt + \frac{1}{1-a} \tag{5.39}$$

is 3-positive definite.

Proof. Due to rotational invariance, we only need to check one value of z. Let $z = e_1$. We have that

$$\begin{aligned}
&K(x,y,e_1) = \langle x,y\rangle^2 + x_1^2 + y_1^2 - ax_1y_1\langle x,y\rangle + \frac{1}{1-a}\\
&= \langle x,y\rangle^2 - ax_1y_1\langle x,y\rangle - (1-a)x_1^2y_1^2 + (1-a)x_1^2y_1^2 + x_1^2 + y_1^2 + \frac{1}{1-a}\\
&= \langle x,y\rangle^2 - ax_1y_1\langle x,y\rangle - (1-a)x_1^2y_1^2 + \Big(x_1^2\sqrt{1-a} + \frac{1}{\sqrt{1-a}}\Big)\Big(y_1^2\sqrt{1-a} + \frac{1}{\sqrt{1-a}}\Big)\\
&= \sum_{j,k=2}^{d} x_jy_jx_ky_k + (2-a)\sum_{m=2}^{d} x_1y_1x_my_m + \Big(x_1^2\sqrt{1-a} + \frac{1}{\sqrt{1-a}}\Big)\Big(y_1^2\sqrt{1-a} + \frac{1}{\sqrt{1-a}}\Big).
\end{aligned}$$

We quickly see that for any finite signed Borel measure $\nu \in \mathcal{M}(\mathbb{S}^{d-1})$,

$$\begin{aligned}
\int_{\mathbb{S}^{d-1}}\int_{\mathbb{S}^{d-1}} K(x,y,e_1)d\nu(x)d\nu(y) &= \sum_{j,k=2}^{d}\Big(\int_{\mathbb{S}^{d-1}} x_jx_kd\nu(x)\Big)^2\\
&+ (2-a)\sum_{m=2}^{d}\Big(\int_{\mathbb{S}^{d-1}} x_1x_md\nu(x)\Big)^2 + \Big(\int_{\mathbb{S}^{d-1}}\Big(x_1^2\sqrt{1-a} + \frac{1}{\sqrt{1-a}}\Big)d\nu(x)\Big)^2,
\end{aligned}$$

hence, K is 3-positive definite. ■

Lemma 5.6.4. *For $a \leq 1$, $K(x,y,z) = t^2 + u^2 + v^2 - auvt$ is conditionally 3-positive definite.*

Proof. For $a < 1$, according to Lemma 5.6.3, $K + \frac{1}{1-a}$ is 3-positive definite. Thus, for any fixed $z \in \mathbb{S}^{d-1}$ and any $\nu \in \mathcal{M}(\mathbb{S}^{d-1})$ with $\nu(\mathbb{S}^{d-1}) = 0$,

$$I_{K_z}(\nu) = I_{K_z + \frac{1}{1-a}}(\nu) \geq 0,$$

i.e. K is conditionally 3-positive definite. Lemma 5.2.2 then gives the result for $a = 1$. ■

5.6.3. Some counterexamples

While our results provide new and less complicated means to determine minimizers for a wide range of kernels, it is clear that more general ideas are necessary to categorize all kernels on the sphere for which σ is a minimizer. In this subsection, we present naturally arising kernels on the sphere which are *not* 3-positive definite on the sphere, but yet the three-input energies generated by these kernels are minimized by the uniform measure σ.

The semidefinite programming methods of Bachoc and Vallentin [BV08] are more computationally difficult than ours, and would likely be infeasible for non-polynomial kernels in the context relevant to this paper. At the same time, they apply to certain functions which are not covered by our methods from Section 5.5. In particular, an appropriate version of semidefinite programming implies that the energies with the following kernels (we keep the notation introduced in [BV08])

$$S^d_{0,1,1}(x,y,z) = uv + vt + tu \tag{5.40}$$

and

$$S^d_{1,0,0}(x,y,z) = (t-uv) + (u-vt) + (v-tu) \tag{5.41}$$

are both minimized by σ, see [Bil+21a]. However, neither function is conditionally 3-positive definite, as we demonstrate below. This implies that the converse to Theorem 5.5.1 does not hold. In addition, the potential of both kernels with respect to σ is a positive definite two-input kernel, which provides evidence that this might indeed be the correct necessary and sufficient condition for σ to minimize the three-input energy (see the discussion before Theorem 5.5.2).

The former example (5.40) is particularly interesting, since the energy functional with this kernel is convex at the minimizer σ, which suggests that conditional n-positive definiteness and convexity of the energy functional are perhaps not equivalent for $n \geq 3$, unlike the two-input case (see Proposition 5.3.5). We summarize these properties in the following proposition:

Proposition 5.6.5. *Let $\Omega = \mathbb{S}^{d-1}$ and set*

$$K(x,y,z) = S^d_{0,1,1}(x,y,z) = uv + vt + tu.$$

The kernel K satisfies the following:

1. *the uniform measure σ minimizes the energy I_K,*
2. *the energy functional I_K is convex at σ,*
3. *$U^\sigma_K(x,y)$ is positive definite,*
4. *K is not conditionally 3-positive definite.*

Proof. As mentioned above, part (1) follows from the semidefinite programming method [Bil+21a], however, there is also a simple direct proof of this fact. Observe that by symmetry, for any $\nu \in \mathbb{P}(\mathbb{S}^{d-1})$,

$$I_K(\nu) = 3\int_{\mathbb{S}^{d-1}} \left(\int_{\mathbb{S}^{d-1}} \langle x,y\rangle d\nu(x) \right)^2 d\nu(y) \geq 0 = I_K(\sigma). \tag{5.42}$$

We now turn to parts (2)–(3). We first note that

$$U_K^\sigma(x,y) = \int_{\mathbb{S}^{d-1}} \langle z, x\rangle \langle y, z\rangle d\sigma(z) = \frac{1}{d}\langle x, y\rangle,$$

which can be proved using the Funk–Hecke formula or by a direct computation (see, e.g., [BDM18]). Hence, the kernel $U_K^\sigma(x,y)$ is positive definite, i.e. (3) holds. Therefore σ minimizes the two-input energy with this kernel, i.e., for any $\nu \in \mathbb{P}(\mathbb{S}^{d-1})$,

$$I_{U_K^\sigma}(\nu) = I_K(\nu,\nu,\sigma) \geq I_{U_K^\sigma}(\sigma) = I_K(\sigma) = 0.$$

Observe also that $U_K^{\sigma^2}(x) = 0$ and thus $I_K(\sigma,\sigma,\nu) = 0$.

For an arbitrary $\nu \in \mathbb{P}(\mathbb{S}^{d-1})$ and $t \in [0,1]$, define $\sigma_t = (1-t)\sigma + t\nu$. Then

$$\begin{aligned} I_K(\sigma_t) &= (1-t)^3 I_K(\sigma) + 3(1-t)^2 t I_K(\sigma,\sigma,\nu) + 3(1-t)t^2 I_K(\nu,\nu,\sigma) + t^3 I_K(\nu) \\ &= 3(1-t)t^2 I_K(\nu,\nu,\sigma) + t^3 I_K(\nu). \end{aligned}$$

If $I_K(\nu) > 0$, we can choose t_ν so small that for all $t \in (0, t_\nu)$, we have $I_K(\nu,\nu,\sigma) \leq \frac{1+t}{3t} I_K(\nu)$, since the right-hand side goes to $+\infty$ as $t \to 0$. Then

$$I_K(\sigma_t) \leq (1-t^2)t I_K(\nu) + t^3 I_K(\nu) = t I_K(\nu) = t I_K(\nu) + (1-t) I_K(\sigma).$$

It remains to consider the case $I_K(\nu) = 0$. According to (5.42), in this situation, $\int_{\mathbb{S}^{d-1}} \langle x, y\rangle d\nu(x) = 0$ for ν-a.e. $y \in \mathbb{S}^{d-1}$, and therefore

$$\int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \langle x, y\rangle d\nu(x) d\nu(y) = 0.$$

But this implies that

$$I_K(\nu,\nu,\sigma) = I_{U_K^\sigma}(\nu) = \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \frac{1}{d}\langle x, y\rangle d\nu(x) d\nu(y) = 0.$$

Thus, when $I_K(\nu) = 0$, we have

$$I_K(\sigma_t) = 3(1-t)t^2 I_K(\nu,\nu,\sigma) + t^3 I_K(\nu) = 0 = (1-t) I_K(\sigma) + t I_K(\nu)$$

for all $t \in [0,1]$. This finishes the proof that I_K is convex at σ.

Finally, we show that I_K is not conditionally 3-positive definite, i.e. part (4). Taking $\mu = \delta_{e_2} - \delta_{-e_1}$ and $z = e_1$, a straightforward computation shows that

$$I_{K_z}(\mu) = I_K(\delta_{e_1}, \mu, \mu) = -1 < 0,$$

which proves our claim. ■

The behavior of the kernel $S^d_{1,0,0}$ is somewhat different. Since

$$I_{S^d_{1,0,0}}(\delta_{e_1}, \delta_{e_1}, \sigma) = \int_{\mathbb{S}^{d-1}} (1 - z_1^2) d\sigma(z) > 0 = I_{S^d_{1,0,0}}(\sigma) = I_{S^d_{1,0,0}}(\sigma, \sigma, \delta_{e_1}) = I_{S^d_{1,0,0}}(\delta_{e_1}),$$

we see that for all $t \in (0, 1)$,

$$I_{S^d_{1,0,0}}(t\delta_{e_1} + (1-t)\sigma) = 3t^2(1-t) I_{S^d_{1,0,0}}(\delta_{e_1}, \delta_{e_1}, \sigma) > t I_{S^d_{1,0,0}}(\delta_{e_1}) + (1-t) I_{S^d_{1,0,0}}(\sigma),$$

so $I_{S^d_{1,0,0}}$ is not convex at σ, and therefore not conditionally 3-positive definite, according to Corollary 5.3.3.1.

In the next subsection we introduce, two more three-input kernels with a geometric flavor, which have similar properties: they also fail to be 3-positive definite, yet the corresponding energies are minimized by the uniform measure σ.

5.6.4. Energies with geometric kernels, which are optimized by the uniform surface measure.

Riesz energies with the kernel $K(x, y) = \|x - y\|^\alpha$ are one of the most important classes of two-input energies. In particular, when $\alpha = 1$, maximizing the sum of distances between points or the corresponding distance integrals is a classical optimization problem of metric geometry [AS74; Bjö56; Tót56] . One can construct interesting multi-input analogues of Riesz energies by replacing the distance with other geometric characteristics which depend on n points, such as area and volume. For $n = 3$, some of the most natural examples include the area of the triangle generated by three points or the volume of the tetrahedron (or the parallelepiped) spanned by three vectors. This can be generalized to higher values of n by considering volumes of various simplices or polytopes generated by n points or vectors.

It is reasonable to conjecture that on the sphere, energy integrals with these three-input kernels (namely, the area of the triangle and the volume of the parallelepiped) are *maximized* by the uniform measure σ. Probabilistically, this can be reformulated in the following way: assume that three random points are chosen on the sphere $\mathbb{S}^{d-1}$ independently according to a probability distribution μ. The conjecture then states that the expected value of these geometric quantities is maximized when the distribution μ is uniform, i.e. $\mu = \sigma$. The question was posed in this form in [Rom19].

This conjecture is supported, among other reasons, by the fact that for the classical case $n = 2$, the analogous kernels $|\sin(\arccos\langle x, y\rangle)| = \sqrt{1 - u^2}$ and $\|x - y\| = \sqrt{2 - 2u}$ (i.e. the area of the parallelogram and the Euclidean distance, respectively) are both negative definite kernels on the sphere (up to an additive constant), and hence the corresponding two-input energies are maximized by σ.

In this section, we verify the conjecture above for slightly different, yet closely related kernels V^2 and A^2: the *squares* of the said volume and area. In these cases, the kernels are multivariate polynomials, which substantially simplifies the analysis. Theorems 5.6.6

and 5.6.7 show that the three-input energies I_{V^2} and I_{A^2} are maximized by the uniform surface measure σ.

Despite the fact that σ is a minimizer of I_{-V^2} and I_{-A^2}, we shall show in Propositions 5.6.9 and 5.6.10 that both kernels $-V^2$ and $-A^2$ fail to be conditionally 3-positive definite, which provides yet another proof that the converse to Theorem 5.5.1 does not hold, unlike in the two-input case.

While in the present paper we only touch upon these questions tangentially, a much more thorough investigation of such geometric problems is undertaken in our paper [Bil+21a].

Volume of the tetrahedron/parallelepiped

Let $V(x, y, z)$ denote the three-dimensional volume of the parallelepiped spanned by the vectors x, y, $z \in \mathbb{S}^{d-1}$. (Observe that the volume of the tetrahedron with vertices at x, y, z, and the origin is $\frac{1}{6}V(x, y, z)$.) The square of the volume $V(x, y, z)$ is given by the determinant of the Gram matrix. Thus we consider the kernel

$$V^2(x, y, z) = \det \begin{pmatrix} 1 & u & v \\ u & 1 & t \\ v & t & 1 \end{pmatrix} = 1 - u^2 - v^2 - t^2 + 2uvt, \tag{5.43}$$

where, as before, we set $u = \langle x, y\rangle$, $v = \langle y, z\rangle$, $t = \langle z, x\rangle$. We have the following statement.

Theorem 5.6.6. *Assume that $d \geq 3$ and $\Omega = \mathbb{S}^{d-1}$. Let $V^2(x, y, z) = 1-t^2-u^2-v^2+2uvt$ be the square of the volume of the parallelepiped spanned by the vectors x, y, $z \in \mathbb{S}^{d-1}$. Then σ is a maximizer of I_{V^2} over $\mathbb{P}(\mathbb{S}^{d-1})$.*

In fact, this theorem also holds for the n-input kernel $K(x_1, \dots, x_n)$ defined as the determinant of the Gram matrix of the set of vectors $\{x_1, \dots, x_n\} \subset \mathbb{S}^{d-1}$ with $d \geq n \geq 3$. This statement is essentially contained in the works of Rankin [Ran56, p. 1956] ($n = d$) and of Cahill and Casazza [CC19] (for $d \geq n$). A comprehensive exposition is presented in our paper [Bil+21a].

Area of the triangle

We now turn to the discussion of the area $A(x, y, z)$ of the triangle with vertices x, y, and $z \in \mathbb{S}^{d-1}$. It is a standard geometrical fact that

$$A^2(x, y, z) = \frac{1}{4}\big(\|y - x\|^2 \cdot \|z - x\|^2 - \langle y - x, z - x\rangle^2\big). \tag{5.44}$$

A straightforward computation then shows that

$$A^2(x, y, z) = \frac{3}{4} - \frac{1}{2}(u + v + t) + \frac{1}{2}(uv + vt + tu) - \frac{1}{4}(u^2 + v^2 + t^2). \tag{5.45}$$

One could also deduce this identity from Heron's formula. We are now ready to prove that the expectation of the area of the triangle *squared* is maximized by the uniform surface measure σ on the sphere $\mathbb{S}^{d-1}$.

Theorem 5.6.7. *Suppose $d \geq 2$, and let $A^2(x,y,z)$ be the square of the area of the triangle with vertices at x, y, $z \in \mathbb{S}^{d-1}$. Then the uniform surface measure σ maximizes $I_{A^2}(\mu)$ over $\mathbb{P}(\mathbb{S}^{d-1})$.*

Proof. Fix an arbitrary measure $\mu \in \mathbb{P}(\mathbb{S}^{d-1})$. Observe that

$$I_u(\mu) = \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \langle x, y \rangle d\mu(x) d\mu(y) = \left\| \int_{\mathbb{S}^{d-1}} x \; \mu(x) \right\|^2. \tag{5.46}$$

Furthermore, applying the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} I_{uv}(\mu) &= \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \langle x, y \rangle \langle z, x \rangle d\mu(x) d\mu(y) d\mu(z) = \int_{\mathbb{S}^{d-1}} \left\langle x, \int_{\mathbb{S}^{d-1}} y d\mu(y) \right\rangle^2 d\mu(x) \\ &\leq \int_{\mathbb{S}^{d-1}} \|x\|^2 \cdot \left\| \int_{\mathbb{S}^{d-1}} y d\mu(y) \right\|^2 d\mu(x) = \left\| \int_{\mathbb{S}^{d-1}} y d\mu(y) \right\|^2 = I_u(\mu). \end{aligned} \tag{5.47}$$

This inequality implies that the contribution of the two middle terms in the representation (5.45) is non-positive, i.e. $I_{\frac{1}{2}(uv+vt+tu)-\frac{1}{2}(u+v+t)}(\mu) \leq 0$. Finally, we have a well-known estimate

$$I_{u^2}(\mu) = \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \langle x, y \rangle^2 d\mu(x) d\mu(y) \geq \frac{1}{d}. \tag{5.48}$$

The two-input energy appearing above is known as the *frame energy*. Its discrete version was introduced in [BF03] in connection to *finite unit norm tight frames* (FUNTF's), for the continuous analogue, see e.g. [BM19]. Putting it all together, we find that

$$I_{A^2}(\mu) \leq \frac{3}{4} - \frac{1}{4} I_{u^2+v^2+t^2}(\mu) \leq \frac{3}{4} - \frac{3}{4d} = \frac{3}{4} \frac{d-1}{d},$$

and it is easy to check that equality holds if $\mu = \sigma$. ∎

Numerous generalizations and refinements of Theorems 5.6.6 and 5.6.7 (including characterizations of minimizers) can be obtained. An in-depth discussion of such geometric problems can be found in our follow-up paper [Bil+21a].

Lack of 3-positive definiteness.

It now remains to show that the kernels $-V^2$ and $-A^2$ are not conditionally 3-positive definite. We first recall the following lemma:

Lemma 5.6.8 (Chp. 3, Lemma 2.1, [BCR84]). *Let Ω be a nonempty set, $x_0 \in \Omega$, $\psi : \Omega^2 \to \mathbb{C}$ be a Hermitian kernel, i.e. $\psi(x,y) = \overline{\psi(y,x)}$, and define*

$$\phi(x,y) := \psi(x,y) + \psi(x_0,x_0) - \psi(x,x_0) - \psi(x_0,y).$$

Then ϕ is positive definite if and only if ψ is conditionally positive definite. If $\psi(x_0,x_0) \leq 0$ and

$$\phi_0(x,y) := \psi(x,y) - \psi(x,x_0) - \psi(x_0,y),$$

then ϕ_0 is positive definite if and only if ψ is conditionally positive definite.

We shall now use this lemma to show that our two geometric kernels are not conditionally 3-positive definite.

Proposition 5.6.9. *Assume that $d \geq 3$, and let $V(x, y, z)$ be the volume of the parallelepiped spanned by the vectors $x, y, z \in \mathbb{S}^{d-1}$. Define the kernel $K(x, y, z) = -V^2(x, y, z)$. Then K is not conditionally 3-positive definite.*

Proof. Using the representation $V^2(x, y, z) = 1 - u^2 - v^2 - t^2 + 2uvt$ and fixing $z = e_1$, we find that $K_{e_1}(x, y) = u^2 + y_1^2 + x_1^2 - 2ux_1y_1 - 1$. It is easy to check that $K_{e_1}(e_1, e_1) = K_{e_1}(e_1, y) = K_{e_1}(x, e_1) = 0$ and hence

$$K_{e_1}(x, y) + K_{e_1}(e_1, e_1) - K_{e_1}(e_1, y) - K_{e_1}(x, e_1) = K_{e_1}(x, y). \tag{5.49}$$

Taking $\nu = \delta_{e_2} + \delta_{e_3}$, one can compute

$$I_{K_{e_1}}(\nu) = -2 < 0,$$

i.e. K_{e_1} is not positive definite. Lemma 5.6.8 and (5.49) then tell us that K_{e_1} is not conditionally positive definite and thus K is not conditionally 3-positive definite. ∎

We now turn to area squared of a triangle and prove an analogous statement.

Proposition 5.6.10. *Assume that $d \geq 2$. Let $A(x, y, z)$ be the area of the triangle with vertices at $x, y, z \subset \mathbb{S}^{d-1}$ and set $K(x, y, z) = -A^2(x, y, z)$. Then K is not conditionally 3-positive definite.*

Proof. As computed in (5.45),

$$A^2(x, y, z) = \frac{3}{4} - \frac{1}{2}(u + v + t) + \frac{1}{2}(uv + vt + tu) - \frac{1}{4}(u^2 + v^2 + t^2).$$

Fixing $z = e_1$, we find that

$$4K_{e_1}(x, y) = u^2 + x_1^2 + y_1^2 + 2u + 2x_1 + 2y_1 - 2x_1y_1 - 2ux_1 - 2uy_1 - 3.$$

The rest of the argument almost repeats the proof of Proposition 5.6.9: we have that

$$K_{e_1}(x, y) + K_{e_1}(e_1, e_1) - K_{e_1}(e_1, y) - K_{e_1}(x, e_1) = K_{e_1}(x, y), \tag{5.50}$$

as well as

$$I_{K_{e_1}}(\delta_{e_2} + \delta_{-e_1}) = -2 < 0,$$

and an application of Lemma 5.6.8 finishes the proof. ∎

5.7. Kernels of the form $F(u, v, t) = h(u)h(v)h(t)$

In this section we show that the potential of F with respect to σ might be positive definite, even though h is not. Further we show that the product of two kernels with positive definite potentials each, does not have necessarily this property.

5.7.1. Zonal harmonics

This short presentation follows Chapter 1 of [DX13] and will be needed in the next subsection.

We denote by $Z_m(x,y)$ the reproducing kernel associated to the orthogonal projection of square integrable functions on the sphere to harmonic, homogeneous polynomials of degree m:

$$\begin{array}{rcl} \mathrm{proj}_m : L^2(\mathbb{S}^d, \sigma) & \to & \mathcal{H}_m^d \\ f & \mapsto & \int_{\mathbb{S}^d} f(y) Z_m(x,y) \, \mathrm{d}\sigma(y). \end{array}$$

Theorem 5.7.1. *The kernels Z_m satisfy following relations:*

- *for $x, y \in \mathbb{S}^d$ we have*

$$\int_{\mathbb{S}^d} Z_m(x,z) Z_m(z,y) \, \mathrm{d}\sigma(z) = Z_m(x,y),$$

- *for $m, k \in \mathbb{N}_0$ with $m \neq k$ and $x \in \mathbb{S}^d$ we have*

$$\int_{\mathbb{S}^d} Z_m(x,y) Z_k(x,y) \, \mathrm{d}\sigma(y) = 0,$$

- *$Z_m(x,y)$ only depends on $\langle x, y \rangle$,*
- *for $d \geq 2$ and with $2\lambda = d - 1$, we have the identity $Z_m(x,y) = \frac{m+\lambda}{\lambda} \mathcal{C}_m^\lambda\big(\langle x, y \rangle\big)$.*

The kernels Z_m are called zonal harmonics, and because of the last relation, Gegenbauer polynomials are also called ultra spherical polynomials.

5.7.2. Counterexamples galore

If our 3-input kernel F is as in the title of the section, then we can integrate once with respect to σ and say that the resulting 2-input kernel is positive definite, i.e. has a Gegenbauer expansion with non-negative coefficients save the constant term. Can we deduce that h has to have a Gegenbauer expansion with non-negative coefficients (save the constant term)? After all, if it does, then so does the potential of F with respect to σ by Proposition 5.6.2. The answer is no, as following lemma will show, but first we define

$$H(u) = \int h(v) h(t) d\sigma(z),$$

where we again used the notation $u = \langle x, y \rangle, \ v = \langle y, z \rangle, \ t = \langle z, x \rangle$, so that

$$h(u) H(u) = \int F(u, v, t) \sigma(z).$$

Lemma 5.7.2. *Let $2\lambda = d-1$ and $u = \langle x, y\rangle$ for $x, y \in \mathbb{S}^d$ with $d \geq 2$. We define $h(u)$ for $n \geq 2$ as*

$$h(u) = \frac{1}{3} - C_1^\lambda(u) + C_2^\lambda(u) + \sum_{j=3}^{n+1} a_j C_j^\lambda(u), \tag{5.51}$$

where $\frac{\lambda+j}{\lambda+1} \geq a_j \geq 0$ for $j \geq 3$ in case $d > 2$, and if $d = 2$ we will require additionally $a_3 = 1$. Then $h(u)H(u)$ will have a Gegenbauer expansion with non-negative coefficients, save the constant term.

Proof. First we will need following formula with $n \leq m$, found in [SRD01, Equ. (29)], which uses the Pochhammer symbol $(\lambda)_n = \prod_{j=0}^{n-1}(\lambda + j)$ for $n > 0$ and $(\lambda)_0 = 1$:

$$\begin{aligned} &C_n^\lambda(t) C_m^\lambda(t) = \\ &\sum_{k=0}^{n} \frac{n+m-2k+\lambda}{n+m-k+\lambda} \frac{(n+m-2k)!(\lambda)_k(\lambda)_{m-k}(\lambda)_{n-k}(2\lambda)_{n+m-k}}{k!(n-k)!(m-k)!(\lambda)_{n+m-k}(2\lambda)_{n+m-2k}} C_{n+m-2k}^\lambda(t). \end{aligned} \tag{5.52}$$

Note that the coefficients are all positive. Further we obtain for $j \geq 1$

$$C_j^\lambda(t) C_1^\lambda(t) = \frac{(j+1)\lambda}{\lambda+j} C_{j+1}^\lambda(t) + \frac{\lambda}{j+\lambda}(2\lambda + j - 1) C_{j-1}^\lambda(t).$$

Now, with some non-negative reminder terms R, R', we have as identity for $h(u)H(u)$:

$$\begin{aligned} &\Big(\frac{1}{3} - C_1^\lambda(u) + \sum_{j=2}^{n+1} a_j C_j^\lambda(u)\Big)\Big(\frac{1}{9} + \sum_{j=1}^{n+1} \frac{\lambda a_j^2}{\lambda+j} C_j^\lambda(u)\Big) \\ &= -\frac{1}{9} C_1^\lambda(u) - \sum_{j=1}^{n+1} \frac{\lambda a_j^2}{\lambda+j} C_j^\lambda(u) C_1^\lambda(u) + \frac{1}{3}\sum_{j=1}^{n+1} \frac{\lambda a_j^2}{\lambda+j} C_j^\lambda(u) + \frac{1}{9}\sum_{j=2}^{n+1} a_j C_j^\lambda(u) \\ &+ \sum_{j=2}^{n+1} a_j C_j^\lambda(u) \frac{\lambda}{\lambda+1} C_1^\lambda(u) + R \\ &= C_1^\lambda(u)\Big(\frac{1}{3}\frac{\lambda}{\lambda+1} - \frac{1}{9}\Big) - \frac{\lambda}{\lambda+1} C_1^\lambda(u)^2 + \sum_{j=2}^{n+1}\Big(\frac{\lambda a_j}{\lambda+1} - \frac{\lambda a_j^2}{\lambda+j}\Big) C_j^\lambda(u) C_1^\lambda(u) + R'. \end{aligned}$$

For this to be non-negative, we only need to make sure that the term

$$-\frac{\lambda}{\lambda+1} C_1^\lambda(u)^2 = -\frac{2\lambda^2}{(\lambda+1)^2} C_2^\lambda(u) + *$$

is compensated ("$*$" denotes some constant that we are not interested in). We already have

$$\frac{1}{3}\frac{\lambda}{\lambda+2} C_2^\lambda(u) + \frac{1}{9} C_2^\lambda(u)$$

for compensation, but we also see that R contains following terms, of which we calculate the coefficient of $C_2^\lambda(u)$ in the Gegenbauer linearization as in (5.52):

$$\begin{aligned}\frac{\lambda}{\lambda+2}C_2^\lambda(u)C_2^\lambda(u) &= *C_4^\lambda(u) + \frac{\lambda}{\lambda+2}C_2^\lambda(u)\frac{2+\lambda}{3+\lambda}\frac{2\lambda^3(2\lambda)_3}{(\lambda)_3(2\lambda)_2} + * \\ &= *C_4^\lambda(u) + C_2^\lambda(u)\frac{\lambda^3}{3+\lambda}\frac{4}{\lambda+2} + *;\end{aligned}$$

and $C_3^\lambda(u)C_1^\lambda(u) = *C_4^\lambda(u) + C_2^\lambda(u)\frac{2\lambda(\lambda+1)}{3+\lambda}$. Thus it would be sufficient if

$$\frac{1}{3}\frac{\lambda}{\lambda+2} + \frac{1}{9} + \frac{\lambda^3}{3+\lambda}\frac{4}{\lambda+2} - \frac{2\lambda^2}{(\lambda+1)^2} + \delta_2^d\Big(\frac{\lambda}{\lambda+1} - \frac{\lambda}{\lambda+3}\Big)\frac{2\lambda(\lambda+1)}{3+\lambda} \geq 0,$$

which is true. ∎

Computer experiments suggest that more counterexamples are given by

$$h(u) = \frac{m_n}{d} + m_n\sum_{j=1}^{n-1} C_j^\lambda(u) - m_n C_n^\lambda(u) + C_{n+1}^\lambda(u) \tag{5.53}$$

for $m_n = \frac{n+\lambda}{n+1+\lambda}$ with $\lambda = \frac{d-1}{2}$; where $u = \langle x, y\rangle$ for $x, y \in \mathbb{S}^d$.

5.7.3. Product of potentials

A natural question that arises for functions in u, v, t that have positive definite potentials, is if the potential of their product is positive definite. This is not necessarily the case as following counter-example demonstrates: Set $F_i(u,v,t) = h_i(u)h_i(v)h_i(t)$ for $i \in \{1,2\}$ with

$$\begin{aligned}h_1(x) &= \frac{1}{5} - \frac{3}{5}C_1^{\frac{1}{2}}(x) + C_2^{\frac{1}{2}}(x),\\ h_2(x) &= \frac{5}{21} + \frac{5}{7}C_1^{\frac{1}{2}}(x) - \frac{5}{7}C_2^{\frac{1}{2}}(x) + C_3^{\frac{1}{2}}(x).\end{aligned}$$

F_1 and F_2 will have positive definite potentials as the h_i are of the form (5.53), but the Gegenbauer expansion of their product's potential has negative coefficients as computer calculations show.

5.8. Acknowledgement

The authors thankfully acknowledge the support of this research: NSF grant DMS 1665007 and the Simons Foundation collaboration grant for mathematicians 712810 (D. Bilyk), the Austrian Science Fund (FWF): F5503 “Quasi-Monte Carlo Methods” and FWF: W1230 “Doctoral School Discrete Mathematics”, and the Austrian Marshall Plan Foundation (D. Ferizović), the NSF Graduate Fellowship 00039202 and the UMN Doctoral Dissertation Fellowship (R. Matzke), AMS-Simons Travel Grant and Postdoctoral Travel Award from the FSU Office of Postdoctoral Affairs (O. Vlasiuk), NSF grant no. CCF-1934904 and NSF grant DMS-1600693 (J. Park).

Bibliography

[Arv12] J. Arvo. “Fast Random Rotation Matrices”. In: *Graphics Gems III*. Ed. by David Kirk. AP Professional, 2012.

[AS72] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Applied Mathematics Series 55. 1972.

[AS74] R. Alexander and K.B. Stolarsky. “Extremal Problems of Distance Geometry Related to Energy Integrals”. In: *Transaction of the Amer. Math. Soc.* 193 (1974).

[AT43] B. M. Axilrod and E. Teller. “Interaction of the van der Waals Type Between Three Atoms”. In: *The Journal of Chem. Physics* 11.6 (1943).

[Aub98] T. Aubin. *Some Nonlinear Problems in Riemannian Geometry*. Springer Monographs in Mathematics. Springer Berlin Heidelberg, 1998.

[AZ15] K. Alishashi and M.S. Zamani. “The spherical ensemble and uniform distribution of points on the sphere”. In: *Electron. J. Probab.* 20.23 (2015), p. 27.

[AZ98] M.H. Annaby and A.I. Zayed. “On the use of Green’s function in Sampling Theory”. In: *J. Integral Equations Applications* 10.2 (1998).

[BC17] E. van den Ban and M. Crainic. *Analysis on Manifolds*. Course Notes. 2017.

[BCC19] C. Beltrán, J.G. Criado del Rey, and N. Corral. “Discrete and Continuous Green Energy on Compact Manifolds”. In: *J. Approx. Theory* 237 (2019).

[BCR84] C. Berg, J.P.R. Christensen, and P. Ressel. *Harmonic Analysis on Semi Groups - Theory of Positive Definite and Related Functions*. Vol. 100. Graduate Texts in Mathematics. Springer, 1984.

[BDM18] D. Bilyk, F. Dai, and R. Matzke. “Stolarsky principle and energy optimization on the sphere”. In: *Constr. Approx.* 48 (2018).

[BE18] C. Beltrán and U. Etayo. “The Projective Ensamble and Distribution of Points in Odd-Dimensional Spheres”. In: *Constr. Approx.* 48.1 (2018), pp. 163–182.

[BF03] J. Benedetto and M. Fickus. “Finite normalized tight frames”. In: *Advances in Computational Mathematics* 18 (2003).

[BF20] C. Beltrán and D. Ferizović. “Approximation to uniform distribution in SO(3)”. In: *Constr Approx* 52 (2020), pp. 283–311.

[BHS19] S.V. Borodachov, D.P. Hardin, and E.B. Saff. *Discrete Energy on Rectifiable Sets*. Springer Monographs in Mathematics. Springer, 2019.

[Bil+21a] D. Bilyk, D. Ferizović, A. Glazyrin, R. Matzke, J. Park, and O. Vlasiuk. "Optimal measures for multivariate geometric potentials". 2021.

[Bil+21b] D. Bilyk, D. Ferizović, A. Glazyrin, R. Matzke, J. Park, and O. Vlasiuk. "Potential Theory with Multivariate Kernels". 2021.

[Bil+21c] D. Bilyk, A. Glazyrin, R. Matzke, J. Park, and O. Vlasiuk. "Optimal measures for p-frame energies on spheres". 2021.

[Bjö56] B. Björck. "Distributions of positive mass, which maximize a certain generalized energy integral". In: *Arkiv för Matematik* 3.21 (1956).

[BM19] D. Bilyk and R. Matzke. "On the Fejes Tóth problem on the sum of acute angles". In: *Proceeding of the Amer. Math. Soc.* 147 (2019).

[BMOC16] C. Beltrán, J. Marzo, and J. Ortega-Cerdà. "Energy and discrepancy of rotationally invariant determinantal point processes in high dimensional spheres". In: *J. Complexity* 37 (2016), pp. 76–109.

[BMV21] D. Bilyk, R. Matzke, and O. Vlasiuk. "Positive definiteness and the Stolarsky invariance principle". 2021.

[BMZ07] H.P. Büchler, A. Micheli, and P. Zoller. "Three-body Interactions with Cold Polar Molecules". In: *Nature Physics* 3 (2007).

[BV08] C. Bachoc and F. Vallentin. "New Upper Bounds for Kissing Numbers from Semidefinite Programming". In: *J. of the Amer. Math. Soc.* 21.3 (2008).

[Can13] Y. Canzani. *Notes for Analysis on Manifolds via the Laplacian*. Course Notes. 2013.

[Car92] M. P. do Carmo. *Riemannian Geometry*. Birkhäuser Boston, 1992.

[CC19] J. Cahill and P.G. Casazza. "Optimal Parseval frames: Total coherence and total volume". https://arxiv.org/pdf/1910.01733. 2019.

[CDS10] B. Colbois, E. B. Dryden, and A. El Soufi. "Bounding the eigenvalues of the Laplace–Beltrami operator on compact submanifolds". In: *Bull. London Math. Soc.* 42 (2010).

[Coc72] J.A. Cochran. *The Analysis of Linear Integral Equations*. McGraw-Hill Book Company, 1972.

[Col+11] K.D. Cole, J.V. Beck, A. Haji-Sheikh, and B. Litkouhi. *Heat Conduction Using Green's Functions*. Ed. by W.J. Minkowycz and E.M. Sparrow. 2nd Edition. Series in Computational and Physical Processes in Mechanics and Thermal Sciences. CRC Press, 2011.

[Cri19] J.G. Criado del Rey. "On the Separation Distance of Minimal Green Energy Points on Compact Riemannian Manifolds". 2019.

[CW12] H. Cohn and J. Woo. "Three-Point Bounds for Energy Minimization". In: *Journal of the AMS* 25.4 (2012).

[Dav99] G. David. “Analytic capacity, Calderón-Zygmund operators, and rectifiability”. In: *Publicacions Matemàtiques* 43.1 (1999).

[Det93] H. Dette. “New identities for orthogonal polynomials on a compact interval”. In: *J. Math. Anal. Appl.* 179 (1993), pp. 547–573.

[Die72] J. Dieudonné. *Treatise on Analysis*. Vol. III. Pure and Applied Mathematics. Academic Press, 1972.

[DS87] P. Diaconis and M. Shahshahani. “The subgroup algorithm for generating uniform random variables”. In: *Prob. Eng. Inf. Sc.* 1 (1987), pp. 15–32.

[DX13] F. Dai and Y. Xu. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer Monographs in Mathematics. Springer, 2013.

[EO12] M. Ehler and K.A. Okoudjou. “Minimization of the probabilistic p-frame potential”. In: *Journal of Statistical Planning and Inference* 142.3 (2012), pp. 645–659. ISSN: 0378-3758. DOI: `10.1016/j.jspi.2011.09.001`.

[Eva10] L. C. Evans. *Partial Differential Equations*. 2nd Edition. American Mathematical Society, 2010.

[Fer21] D. Ferizović. “On the L^2-norm of Gegenbauer polynomials”. In: *Mathematical Sciences* (2021). URL: `https://doi.org/10.1007/s40096-021-00398-1`.

[FS13] F. Finster and D. Schiefeneder. “On the support of minimizers of causal variational principles”. In: *Arch Rational Mech Anal* 210.2 (2013).

[FT15] L.C. Flatley and T. Theil. “Face-Centered Cubic Crystallization in Atomistic Configurations”. In: *Arch Rational Mech Anal* 218 (2015).

[Fug60] B. Fuglede. “On the theory of potentials in locally compact spaces”. In: *Acta Math.* 103 (1960).

[Gra+00] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey, and D. Zwillinger. *Table of Integrals, Series, and Products*. 6th Edition. Academic Press, 2000.

[Gri01] A. Grigor’yan. *Analysis on Manifolds*. Course Notes. 201?

[Gri09] A. Grigor’yan. *Heat Kernel and Analysis on Manifolds*. Ed. by S.-T. Yau. Vol. 47. Studies in Advanced Mathematics. Amer. Math. Soc., 2009.

[Gri83] A. Grigor’yan. “On the existence of a Green’s function on a manifold”. In: *Russ. Math. Surv.* 38.190 (1983).

[HKP16] A. Hassannezhad, G. Kokarev, and I. Polterovich. “Eigenvalue inequalities on Riemannian manifolds with a lower Ricci curvature bound”. In: *J. Spectr. Theory* 6 (2016).

[Hou+09] J. Ben Hough, M. Krishnapur, Y. Peres, and V. Virág. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. Providence, RI: American Mathematical Society, 2009.

[HS04] D.P. Hardin and E.B. Saff. “Discretizing manifolds via minimum energy points”. In: *Notices of the Amer. Math. Soc.* 51.10 (2004).

[HS11] T. Hangelbroek and D. Schmid. “Surface Spline Approximation on SO(3)”. In: *Appl. Comput. Harmon. Anal.* 31.2 (2011), pp. 169–184.

[Hua63] L. K. Hua. *Harmonic analysis of functions of several complex variables in the classical domains*. Translated from the Russian by L. Ebner and A. Korányi. Providence, R.I.: American Mathematical Society, 1963.

[Jos11] J. Jost. *Riemannian geometry and geometric analysis*. 6th Edition. Universitext. Heidelberg: Springer, 2011.

[Jos73] A. W. Joshi. *Elements of Group Theory for Physicists*. New Delhi: Wiley Eastern Private Limited, 1973.

[KS97] A.B.J. Kuijlaars and E.B. Saff. “Distributing many points on a sphere”. In: *The Mathematical Intelligencer* 19 (1997).

[Laa+19] D. de Laat, F.C. Machado, F.M. Oliveira de Filho, and F. Vallentin. “k-point semidefinite programming bounds for equiangular lines”. In: (2019). https://arxiv.org/abs/1812.06045.

[Lan88] S. Lang. *Introduction to Arakelov Theory*. New York: Springer-Verlag, 1988.

[Lee13] J. M. Lee. *Introduction to Smooth Manifolds*. 2nd Edition. Graduate Texts in Mathematics. New York: Springer Science+Business, 2013.

[Li+18] S. Li, P. Roy, A. Travesset, and R. Zandi. “Why large icosahedral viruses need scaffolding proteins”. In: *PNAS* 43.115 (2018).

[Lio06] P.-L. Lions. *Équations aux dérivées partielles et applications*. 2006. URL: `https://www.college-de-france.fr/site/pierre-louis-lions/course-2006-2007.htm`.

[Min53] S. Minakshisundaram. “Eigenfunctions on Riemannian Manifolds”. In: *J. Indian Math. Soc.* 17.4 (1953).

[MMV96] P. Mattila, M.S. Melnikov, and J. Verdera. “The Cauchy integral, analytic capacity, and uniform rectifiability”. In: *Ann. of Math.* 144.1 (1996).

[MOC18] J. Marzo and J. Ortega-Cerdà. “Expected Riesz energy of some determinantal processes on flat tori”. In: *Constructive Approximation* 47.1 (2018), pp. 75–88.

[MS14] E. Mainini and U. Stefanelli. “Crystallization in Carbon Nanostructures”. In: *Communications in Mathematical Physics* 328 (2014).

[Mus14] O.R. Musin. “Multivariate positive definite functions on spheres”. In: *Discrete Geometry and Algebraic Combinatorics*. Contemporary Mathematics. Amer. Math. Soc., 2014.

[Qi05] L. Qi. “Eigenvalues of a real supersymmetric tensor”. In: *Journal of Symbolic Computation* 40 (2005).

[Ran56] R.A. Rankin. “On the Minimal Points of Positive Definite Quadratic Forms”. In: *Mathematika* 3.1 (1956).

[Rom19] Romeo. math.stackexchange post. 2019. URL: `https://math.stackexchange.com/questions/3114400/maximum-expectation-of-volume-of-a-regular-tetrahedron-whose-angles-are-i-i-d`.

[RR04] M. Renardy and R.C. Rogers. *An Introduction to Partial Differential Equations.* Ed. by J.E. Marsden, L. Sirovich, and S.S. Antman. 2nd Edition. Vol. 13. Texts in Applied Mathematics. Springer, 2004.

[RV07] B. Rider and B. Virág. "Complex determinantal processes and H^1 noise". In: *Electronic Journal of Probability* 12 (2007), pp. 1238–1257.

[RWP06] M. Reuter, F.-E. Wolter, and N. Peinecke. "Laplace–Beltrami spectra as 'Shape-DNA' of surfaces and solids". In: *Computer-Aided Design* 38.4 (2006).

[San15] F. Santambrogio. *Optimal transport for applied mathematicians. Calculus of variations, PDEs, and modeling.* Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, 2015.

[Sch13] R.E. Schwartz. "The 5-Electron case of Thomson's Problem". In: *J. Experimental Math* 22.2 (2013).

[Sch41] I.J. Schoenberg. "Positive definite functions on spheres". In: *Duke Mathematical Journal* 9 (1941).

[Skr20] M. Skriganov. "Stolarsky's invariance principle for projective spaces". In: *Journal of Complexity* 56 (2020).

[Sma98] S. Smale. "Mathematical Problems for the Next Century". In: *Math. Intelligencer* 20.2 (1998), pp. 7–15.

[Smi38] F. Smithies. "The Eigenvalues and Singular Values of Integral Equations". In: *Proceedings of the London Mathematical Society* s2-43.1 (1938).

[SRD01] J. Sánchez-Ruiz and J. S. Dehesa. "Some connection and linearization problems for polynomials in and beyond Askey scheme". In: *Journal of Computational and Applied Mathematics* 133 (2001).

[SS93] Eyssette F. and Galligo A., eds. *Complexity of Bezout's theorem II – Volumes and probabilities.* Progress in Mathematics, vol 109. Boston, MA: Birkhäuser, 1993, pp. 267–285.

[Sto73] K.B. Stolarsky. "Sums of distances between points on a sphere II". In: *Proceeding of the Amer. Math. Soc.* 41 (1973).

[SW85] F.H. Stillinger and T.A. Weber. "Computer simulation of local order in condensed phases of silicon". In: *Physical Review B* 31 (1985).

[Sze39] G. Szegö. *Orthogonal Polynomials.* Amer. Math. Soc., 1939.

[TE51] F. Tricomi and A. Erdélyi. "The Asymptotic Expansion of a Ration of Gamma Functions". In: *Pacific Journal of Mathematics* 1.1 (1951), pp. 133–142.

[Th] *Thomson Problem.* URL: `https://en.wikipedia.org/wiki/Thomson_problem`.

[Tor10] S. Torquato. “Reformulation of the covering and quantizer problems as ground states of interacting particles”. In: *Phys. Rev. E* 82 (2010).

[Tót56] L. Fejes Tóth. “On the sum of distances determined by a point set”. In: *Acta Mathematica Academiae Scientiarum Hungarica* 7 (1956).

[Ura93] H. Urakawa. “Geometry of Laplace-Beltrami Operator on a Complete Riemannian Manifold”. In: *Progress in Differential Geometry.* Vol. 22. Advanced Studies in Pure Mathematics. Mathematical Society of Japan, 1993, pp. 347–406.

[Ver87] C. de Verdière. “Construction de laplaciens dont une partie finie (avec multi-plicités) du spectre est donnée”. In: *Annales scientifiques de l’École Normale Supérieure* Série 4, Tome 20.4 (1987).

[Vol10] A. Vollrath. “The Nonequispaced Fast SO(3) Fourier Transform, Generalisations and Applications”. PhD thesis. University of Lübeck, 2010.

[Wig59] E. P. Wigner. *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra.* New York, London: Academic Press, 1959.

[Zel09] V.G. Zelevinsky. “Three-Body Forces and Many-Body Dynamics”. In: *Physics of Atomic Nuclei* 72 (2009).

[ZST16] G. Zhang, F.H. Stillinger, and S. Torquato. “The Perfect Glass Paradigm: Disordered Hyperuniform Glasses Down to Absolute Zero”. In: *Scientific Reports* 6 (2016).