A Walk through Energy, Discrepancy, Numerical Integration and Group Invariant Measures on Measurable Subsets of Euclidean Space

S. B. Damelin
in collaboration with P. Grabner, F. Hickernell, J. Levesley, D. Ragozin, X. Sun and Z. Zeng

February 10, 2008

Abstract

(A) The celebrated Gaussian quadrature formula on finite intervals tells us that the Gauss nodes are the zeros of the unique solution of an extremal problem. We announce recent results of Damelin, Grabner, Levesley, Ragozin and Sun which derive quadrature estimates on compact, homogenous manifolds embedded in Euclidean spaces, via energy functionals associated with a class of group-invariant kernels which are generalizations of zonal kernels on the spheres or radial kernels in euclidean spaces. Our results apply, in particular, to weighted Riesz kernels defined on spheres and certain projective spaces. Our energy functionals describe both uniform and perturbed uniform distribution of quadrature point sets.

(B) Given $\mathcal{X}$, some measurable subset of Euclidean space, one sometimes wants to construct, a design, a finite set of points, $\mathcal{P} \subset \mathcal{X}$, with a small energy or discrepancy. We announce recent results of Damelin, Hickernell, Ragozin and Zeng which show that these two measures of quality are equivalent when they are defined via positive definite kernels $K : \mathcal{X}^2 (= \mathcal{X} \times \mathcal{X}) \to \mathbb{R}$. The error of approximating the integral $\int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x})$ by the sample average of f over $\mathcal{P}$ has a tight upper bound in terms the energy or discrepancy of $\mathcal{P}$. The tightness of this error bound follows by requiring f to lie in the Hilbert space with reproducing kernel K. The theory presented here provides an interpretation of the best design for numerical integration as one with minimum energy, provided that the μ defining the integration problem is the equilibrium measure or charge distribution corresponding to the energy kernel, K.

(C) Let $\mathcal{X}$ be the orbit of a compact, possibly non Abelian group, $\mathcal{G}$, acting as measurable transformations of $\mathcal{X}$ and the kernel K is invariant under the group action. We announce recent results of Damelin, Hickernell, Ragozin and Zeng which show that the equilibrium measure is the normalized measure on $\mathcal{X}$ induced by Haar measure on $\mathcal{G}$. This allows us to calculate explicit representations of equilibrium measures.

There is an extensive literature on the topics (A-C). We emphasize that this paper surveys recent work of Damelin, Grabner, Levesley, Hickernell, Ragozin, Sun and Zeng and does not mean to serve as a comprehensive survey of all recent work covered by the topics (A-C).

AMS(MOS) Classification: Primary 41A05; Secondary 41A15, 41A63

Keywords: Compact Homogeneous Manifold, Discrepancy, Group, Energy, Invariant Kernels, Invariant Polynomial, Numerical Integration, Projection Kernels, Projective Space, Quadature, Reflexive Manifold, Riesz Kernel, Spherical Harmonic, Sphere, Torus, Weight, Capacity, Cubature, Discrepancy, Distribution, Group invariant kernel, Group invariant measure, Energy Minimizer, Equilibrium measure, Numerical integration, Positive definite, Potential field, Reproducing Hilbert space, Signed measure.

1 Introduction

The problem of uniformly distributing points on a sphere or other compact set in Euclidean space with positive d dimensional Hausdorf measure is an interesting and difficult problem. It was discussed already by Carl Friedrich Gauss in his famous Disquisitiones arithmaticae in the case of a sphere, although it is most likely that similar problems appeared in mathematical writings even before that time. For $d \geq 1$, let S^d denote the d-dimensional unit sphere in $\mathbb{R}^{d+1}$. For $d = 1$, the problem is reduced to uniformly distributing $n \geq 1$ points on a circle, and equidistant points provide an obvious answer. For $d \geq 2$, the problem becomes much more difficult, yet a physically motivated solution is to treat the points as electrostatic charges and place them so that an electrostatic *energy* is minimized. In the time since Gauss the problem has expanded to include general definitions of the energy, E, defined for charge distributions, μ, on general subsets of Euclidean space, $\mathcal{X}$. The problem remains as to how to arrange a set of points $\mathcal{P} = \{\boldsymbol{z}_i\}_{i=1}^n \subseteq \mathcal{X}$ so that its energy, $E(\mathcal{P})$, is minimized.

Another approach to spreading points uniformly, developed initially for the d-dimensional unit cube, $[0,1]^d$, is the *discrepancy* of Weyl. This original discrepancy is defined as the sup-norm of the difference between the uniform

distribution and the empirical distribution of the points, $\mathcal{P}$. It is known in the statistics literature as a Kolmogorov-Smirnov statistic.

The equivalence between energy and discrepancy benefits the energy community by providing tight upper bounds on numerical integration error via the energy of the design:

$$\sup_{\|f\|_{\mathcal{H}(K)}\leq 1}\left|\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\mu_{\mathrm{e}}(\boldsymbol{x})-\frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{z}_i)\right|=\sqrt{E(\mathcal{P})-E(\mu_{\mathrm{e}})}=D(\mathcal{P};\mu_{\mathrm{e}}),$$

where μ_{e} is the equilibrium charge distribution. The error bound above differs from most of those found in the energy literature in that it is *tight*. A kernel, K, is thus used to define both energy and discrepancy in the above. In the definition of energy, $K(\boldsymbol{x},\boldsymbol{y})$ represents the energy of unit charges placed at the points $\boldsymbol{x}$ and $\boldsymbol{y}$. In the discrepancy literature K defines a reproducing kernel Hilbert space of integrands, $\mathcal{H}(K)$, and this interpretation of K leads to the above worst-case numerical integration error.

This equivalence between discrepancy and energy provides the discrepancy community with a physical interpretation of discrepancy, namely as the square root of an energy, see Section 2. The space of measures for which the energy is finite, $\mathcal{M}(K)$, may be interpreted as a space of charge distributions, and the energy is the square norm of those charge distributions. The reproducing kernel Hilbert space of integrands, $\mathcal{H}(K)$, may be interpreted as a space of potential functions induced by the charge distributions in $\mathcal{M}(K)$. The discrepancy of a design, $D(\mathcal{P};\mu)$, depends on a target distribution defining the integration problem, μ, however, when μ coincides with the equilibrium charge distribution the expression for the discrepancy simplifies to the root difference of two energies.

In Section 2, of this article, we announce results of Damelin, Hickernell, Ragozin and Zeng dealing with the equivalence of discrepancy and energy and their relationship to quadrature estimates over measurable subsets $\mathcal{X}$ of Euclidean space and for integrands $f:\mathcal{X}\to\mathbb{R}$. Given $\mathcal{X}$, some measurable subset of Euclidean space, one sometimes wants to construct a design, a finite set of points, $\mathcal{P}\subset\mathcal{X}$, with a small energy or discrepancy. We announce recent results of Damelin, Hickernell, Ragozin and Zeng which show that these two measures of quality are equivalent when they are defined via positive definite kernels $K:\mathcal{X}^2(=\mathcal{X}\times\mathcal{X})\to\mathbb{R}$. The error of approximating the integral $\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x})$ by the sample average of f over $\mathcal{P}$ has a tight upper bound in terms the energy or discrepancy of $\mathcal{P}$. The tightness of this error bound follows by requiring f to lie in the Hilbert space with reproducing kernel K. The theory presented here provides an interpretation of the best

design for numerical integration as one with minimum energy, provided that the μ defining the integration problem is the equilibrium measure or charge distribution corresponding to the energy kernel, K.

The celebrated Gaussian quadrature formula on finite intervals tells us that the Gauss nodes are the zeros of the unique solution of an extremal problem. In Section 4, we announce recent results of Damelin, Grabner, Levesley, Ragozin and Sun which derive quadrature estimates on compact, homogenous manifolds embedded in Euclidean spaces, via energy functionals associated with a class of group-invariant kernels which are generalizations of zonal kernels on the spheres or radial kernels in euclidean spaces. Our results apply, in particular, to weighted Riesz kernels defined on spheres and certain projective spaces. Our energy functionals describe both uniform and perturbed uniform distribution of quadrature point sets. In the results of this section, the discrepancy bound splits into two parts. The first depending on the nodal set and the second depending on the function space. This in contrast to the results of Section 2 where the function space depends on the kernel. Thus the price we pay for a tight error bound is typically a smaller class of functions f.

Section 3, tackles the problem of calculating equilibrium measures, which play a crucial role in the distribution of quadrature, interpolatory and minimal energy points, zeroes of orthogonal and other extremal polynomials, as well as suitably scaled eigenvalues in random matrix theory. In this section, we announce recent results of Damelin, Hickernell, Ragozin and Zeng dealing with the following. Let $\mathcal{X}$ be the orbit of a compact, possibly non Abelian group, $\mathcal{G}$, acting as measurable transformations of $\mathcal{X}$ and the kernel K is invariant under the group action. We show that the equilibrium measure is the normalized measure on $\mathcal{X}$ induced by Haar measure on $\mathcal{G}$. This allows us to calculate explicit representations of equilibrium measures

There is an extensive literature on the topics (A-C). We emphasize that this paper surveys recent work of Damelin, Levesley, Hickernell, Ragozin, Sun an Zeng and does not mean to serve as comprehensive survey of recent work covered by the topics (A-C). See the references at the end of Section 4.

The remainder of this paper is devoted to the announcement of results in Sections 2-4.

2 On Discrepancy and Energy

Here and throughout, let $\mathcal{X}$ be a measurable subset of Euclidean space with Hausdorf dimension d and let $\|\cdot\|_2$ denote the Euclidean norm. Let $\mathcal{B}$ denote the space of all finite signed Borel measures (*charge distributions*) μ on $\mathcal{X}$. The measure of the whole set will be denoted $Q(\mu) := \mu(\mathcal{X}) = \int_{\mathcal{X}} \mathrm{d}\mu(\boldsymbol{x})$. In the energy interpretation, μ is a charge distribution and $Q(\mu)$ is the *total charge*.

2.1 Kernels and Energy

Let $K : \mathcal{X}^2 \to \mathbb{R} \cup \{+\infty\}$ be a function which is Borel measurable, symmetric, bounded below and (strictly) *positive definite*. That is, $\forall \mu \in \mathcal{B}$

$$\begin{gathered} K(\boldsymbol{x}, \boldsymbol{y}) = K(\boldsymbol{y}, \boldsymbol{x}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}, \\ \exists L_K \in \mathbb{R} \text{ with } K(\boldsymbol{x}, \boldsymbol{y}) \geq L_K, \\ E(\mu) := \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \mathrm{d}\mu(\boldsymbol{y}) > 0 \text{ for all } \mu \neq 0 \text{ with } E(\mu) < \infty. \end{gathered} \tag{1}$$

A positive definite kernel implies that the energies of all nonzero charge distributions are positive. In some cases, the kernel is allowed to be only *conditionally positive definite*. This means that that $E(\mu) > 0$ if $\mu \neq 0$ and $Q(\mu) = 0$. The kernel itself may be infinite for some values, e.g., one important kernel defined on $\mathbb{R}$ arising from electrostatic energy is $-\log(|x - y|)$.

From an energy perspective, the function $K(\cdot, \boldsymbol{y})$ is the *potential field* induced by a unit point charge placed at $\boldsymbol{y}$, and $K(\boldsymbol{x}, \boldsymbol{y})$ is then the *potential energy* of a unit test charge placed at $\boldsymbol{x}$ under this field. Generalizing from this, we may identify $E(\mu)$ defined above as the *energy* of the charge distribution μ. It must be remembered that the definition of energy depends on the kernel, K, although this dependence is suppressed in the notation for simplicity's sake.

Example 2.1 (Riesz Kernel) *If $\mathcal{X} = [0, 1]$, the logarithmic potential, mentioned above, and the Coulombic potential, $\|\boldsymbol{x} - \boldsymbol{y}\|_2^{-1}$, are special cases of the* Riesz kernel*:*

$$K_s(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} \operatorname{sign}(s) \, \|\boldsymbol{x} - \boldsymbol{y}\|_2^{-s}, & s \neq 0, \\ -\log(\|\boldsymbol{x} - \boldsymbol{y}\|_2), & s = 0. \end{cases} \tag{2}$$

The Riesz kernel is symmetric, and is positive definite for $0 \leq s < 1$. It is conditionally positive definite *for $-1 \leq s < 0$. The Riesz kernel can be*

made strictly positive definite for $-1 \leq s < 0$ where $\mathcal{X}$ is bounded by adding a large enough constant. For $s \geq 1$, the energy is infinite for nonzero charge distributions.

Example 2.2 (Symmetric, Positive Definite Matrix) *If $\mathcal{X} = \{1, \ldots, N\}$, then the kernel, K, may be represented as an $N \times N$ symmetric, positive definite matrix, $\mathsf{K} = (K(x,y))_{x,y=1}^{N}$. In this case, the space of signed measures is $\mathcal{B} = \mathbb{R}^N$, and the energy may be represented as the vector-matrix product $E(\boldsymbol{\mu}) = \boldsymbol{\mu}^T \mathsf{K} \boldsymbol{\mu}$.*

Note that the definition of energy in (1) deviates somewhat from the usual physical definition of energy for a total unit charge distributed over the $n \geq 1$ points in $\mathcal{P} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$:

$$E_{\text{self}}(\mathcal{P}) = \frac{1}{n^2} \sum_{1 \leq i < j \leq n} K(\boldsymbol{z}_i, \boldsymbol{z}_j). \tag{3}$$

Let $\delta_{\boldsymbol{z}} \in \mathcal{B}$ be the Dirac delta measure that corresponds to a unit charge at the point $\boldsymbol{z} \in \mathcal{X}$, i.e., $\int_{\mathcal{Y}} \mathrm{d}\delta_{\boldsymbol{z}}(\boldsymbol{x}) = 1$ for all measurable sets $\mathcal{Y} \subseteq \mathcal{X}$ with $\boldsymbol{z} \in \mathcal{Y}$. The *empirical distribution* of the set $\mathcal{P}$, defined as

$$\mu_{\mathcal{P}} := \frac{1}{n} \sum_{i=1}^{n} \delta_{\boldsymbol{z}_i}, \tag{4}$$

assigns equal charge n^{-1} to each point in the set $\mathcal{P}$ for a total charge of unity. When we talk of minimizing the energy of a set of points $\mathcal{P}$, we mean minimizing

$$E(\mathcal{P}) := E(\mu_{\mathcal{P}}) = \frac{1}{n^2} \sum_{i,j=1}^{n} K(\boldsymbol{z}_i, \boldsymbol{z}_j). \tag{5}$$

Here we abuse the notation to allow the argument of E to be a set of points or a charge distribution.

2.2 Linear Spaces of Charge Distributions and Discrepancy

The quadratic form defining the energy may be used to define an inner product. Let $\mathcal{M}(K) \subseteq \mathcal{B}$ be the linear space of measures with finite energy, $E(\mu)$. For this space one has an inner product:

$$\langle \mu, \nu \rangle_{\mathcal{M}(K)} := \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \mathrm{d}\nu(\boldsymbol{y}), \tag{6}$$

and the energy then corresponds to the square norm of the measure, i.e.,

$$E(\mu) = \|\mu\|^2_{\mathcal{M}(K)}\,. \tag{7}$$

The *discrepancy* of the measure ν with respect to the measure μ is defined as

$$D(\nu;\mu) := \|\mu - \nu\|_{\mathcal{M}(K)}\,. \tag{8}$$

The definition of the discrepancy depends on the choice of kernel as well as of the target distribution, μ. The definitions of energy and discrepancy immediately yield the following equivalence theorem:

Theorem 2.3 *For energy and discrepancy defined above, it follows that* $D(\nu;\mu) = \sqrt{E(\mu-\nu)}$.

As with energy, the discrepancy of a set of points, $\mathcal{P} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$, is defined as the discrepancy of the empirical distribution function of that set of points, namely,

$$D^2(\mathcal{P};\mu) := D^2(\mu_{\mathcal{P}};\mu) = \int_{\mathcal{X}^2} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{y}) - \frac{2}{n}\sum_{i=1}^{n}\int_{\mathcal{X}} K(\boldsymbol{z}_i,\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) + \frac{1}{n^2}\sum_{i,j=1}^{n} K(\boldsymbol{z}_i,\boldsymbol{z}_j), \tag{9}$$

where the last term in this expression is $E(\mu_{\mathcal{P}})$. A generalization of the distribution $\mu_{\mathcal{P}}$ places a charge of magnitude q_i at each point $\boldsymbol{z}_i$, i.e., $\mu_{\mathcal{P},\boldsymbol{q}} := \sum_{i=1}^{n} q_i \delta_{z_i}$. Its discrepancy is

$$D^2(\mu_{\mathcal{P},\boldsymbol{q}};\mu) = \int_{\mathcal{X}^2} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{y}) - \sum_{i=1}^{n} q_i \int_{\mathcal{X}} K(\boldsymbol{z}_i,\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) + \sum_{i,j=1}^{n} q_i q_j K(\boldsymbol{z}_i,\boldsymbol{z}_j), \tag{10}$$

where again the last term in this expression is $E(\mu_{\mathcal{P},\boldsymbol{q}})$

Note that the definition of discrepancy in (8) does not strictly include the discrepancy of Weyl, also called the L_∞ star discrepancy and defined over $\mathcal{X} = [0,1]^d$ as

$$D_{*,\infty}(\nu;\mu) = \|\mu - \nu\|_\infty = sup_{\boldsymbol{x}\in[0,1]^d}\,|\mu(\boldsymbol{x}) - \nu(\boldsymbol{x})|\,,$$

where $\boldsymbol{x} = (x_1, \ldots, x_d)$, $\mu(\boldsymbol{x}) := \mu(\prod_{i=0}^{d}(-\infty, x_i])$. However, the closely related L_2 star discrepancy,

$$D_{*,2}(\nu;\mu) = \|\mu - \nu\|_2 = \left[\int_{[0,1]^d} |\mu(\boldsymbol{x}) - \nu(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x}\right]^{1/2}.$$

Example 2.4 (Product Kernels) *The L_2 star discrepancy for $[0,1]$ is the discrepancy associated to the kernel $K(x,y) = 1 - |x-y|$ which is a slight modification of the Riesz kernel in Example 1 for $s = -1$. An analogous discrepancy for the d-dimensional unit cube in dimensions $d > 1$ is typically defined using a product kernel, such as in:*

$$K(\boldsymbol{x}, \boldsymbol{y}) = \prod_{k=1}^{d} [1 - |x_k - y_k|], \quad \boldsymbol{x}, \boldsymbol{y} \in [0,1]^d. \tag{11}$$

2.3 Linear Spaces of Potential Fields and Numerical Integration Error

For every signed measure $\mu \in \mathcal{M}(K)$, there exists a function f_μ defined by

$$f_\mu(\boldsymbol{x}) = \int_{\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{y}).$$

In the energy literature, this function is the potential field induced by the charge distribution μ. Let $\mathcal{H}(K)$ denote the linear space of all such potential fields. The inner product on the linear space of charge distributions, $\mathcal{M}(K)$, induces an inner product on the space of potential fields as follows:

$$\langle f_\mu, f_\nu \rangle_{\mathcal{H}(K)} := \langle \mu, \nu \rangle_{\mathcal{M}(K)} \quad \forall f_\mu, f_\nu \in \mathcal{H}(K). \tag{12}$$

When we need to denote the charge distribution that induces the field, the subscript is used. Otherwise, we simply denote elements of $\mathcal{H}(K)$ by the letters f, g, etc.

The space of potential fields, $\mathcal{H}(K)$, arises in the study of numerical integration error. Note that for any $\mu \in \mathcal{M}(K)$, integration against this measure gives a continuous linear functional, L_μ, on $\mathcal{H}(K)$:

$$L_\mu(f_\nu) = \int_{\mathcal{X}} f_\nu(\boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{y}) = \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\nu(\boldsymbol{x}) \mathrm{d}\mu(\boldsymbol{y}) = \langle f_\nu, f_\mu \rangle_{\mathcal{H}(K)}. \tag{13}$$

This says that the linear functional $L_\mu(\cdot) = \langle \cdot, f_\mu \rangle_{\mathcal{H}(K)}$ whose *representer* is the potential field f_μ is just integration against the measure μ. The Cauchy-Schwartz inequality then implies that the largest possible difference between

the integrals of f with respect to two different measures has a tight bound in terms of the discrepancy.

Theorem 2.5 *For any $\mu, \nu \in \mathcal{M}(K)$ it follows that*

$$\sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) - \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{x}) \right| = D(\nu; \mu).$$

For numerical integration we approximate an integral $\int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x})$ by the sample average of the integrand values on a set $\mathcal{P}$ of n points, $\frac{1}{n} \sum_{\boldsymbol{z} \in \mathcal{P}} f(\boldsymbol{z}) = \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu_{\mathcal{P}}(\boldsymbol{x})$. A sufficient condition for the measure $\mu_{\mathcal{P}}$ to lie in $\mathcal{M}(K)$ is that all the point charges lie in $\mathcal{M}(K)$, i.e.,

$$\delta_{\boldsymbol{z}} \in \mathcal{M}(K) \quad \forall \boldsymbol{z} \in \mathcal{X}.$$

For kernels, K, satisfying this condition, the linear space $\mathcal{H}(K)$ contains all potential fields of the form $K(\cdot, \boldsymbol{y})$ for $\boldsymbol{y} \in \mathcal{X}$. Moreover, the space of potential fields may be completed. The kernel K is then called the reproducing kernel, and $K(\cdot, \boldsymbol{z}) = \int_{\mathcal{X}} K(\cdot, \boldsymbol{y}) \, \mathrm{d}\delta_{\boldsymbol{z}}(\boldsymbol{y})$ is the representer for function evaluation at the point $\boldsymbol{z}$. We denote the completion of $\mathcal{H}(K)$ by $\mathcal{H}(K)$ as well. Similarly, if $\mathcal{M}(K)$ can be completed, we denote this completion by $\mathcal{M}(K)$ as well. The worst case numerical integration error is given by the discrepancy, as specified in the corollary below.

Corollary 2.6 *Suppose that the kernel K defines a reproducing kernel Hilbert space, $\mathcal{H}(K)$, μ is some fixed measure in $\mathcal{M}(K)$, and $\mathcal{P} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\} \subseteq \mathcal{X}$. Then it follows that*

$$\sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) - \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{z}_i) \right| = D(\mathcal{P}; \mu) = D(\mu_{\mathcal{P}}; \mu) = \sqrt{E(\mu - \mu_{\mathcal{P}})}.$$

Note that reproducing kernels require that $K(\boldsymbol{x}, \boldsymbol{x})$ be finite, which is the case for the Riesz kernel with $s < 0$ but not for $s \geq 0$. Associating energy to numerical integration error for kernels that are infinite along the diagonal has been studied by Damelin and Grabner considering these kernels as limits of finite-valued kernels K_ϵ that approach K in the limit $\epsilon \downarrow 0$.

The essence of this Corollary is that optimal point sets, $\mathcal{P}$, for numerical integration are those that minimize $D(\mathcal{P}; \mu)$. How small the discrepancy can be made for a given n depends on the difficulty of the numerical integration problem, which depends on what functions are allowed in the space of integrands, $\mathcal{H}(K)$. The reproducing kernel, K, defines the degree of smoothness

and other properties of the integrands lying in $\mathcal{H}(K)$. The Kernel defines a space of integrands that have square integrable mixed partial derivatives of order up to one in each variable. It is known that the discrepancy for this kernel when μ is the uniform measure decays at best as $\mathcal{O}(n^{-1+\epsilon})$ for any $\epsilon > 0$ and this rate of decay is observed for certain low discrepancy sequences.

3 Equilibrium Measures and Minimizers

Although we have seen that numerical integration error relates to the energy, we now show that minimum energy point sets, i.e., $\mathcal{P}$ minimizing $E(\mathcal{P})$, yield the smallest numerical integration error. To make a direct connection between minimum energy points and numerical integration error we need to look deeper into the energy paradigm, in particular at equilibrium measures and minimizers.

3.1 Equilibrium Measures Minimize the Energy

Definition 3.1 *Consider a symmetric, positive definite kernel K on $\mathcal{X}$ and the energy, E, defined in* (1)*. Let $\mathcal{Y}$ be a measurable subset of $\mathcal{X}$ and let* $\operatorname{supp}(\mu)$ *denote the support of a measure μ. The* capacity *of a measurable set $\mathcal{Y} \subseteq \mathcal{X}$ is defined as the inverse of the infimum of the energies for charge distributions with support in $\mathcal{Y}$ and unit total charge:*

$$C_K(\mathcal{Y}) = \left[\inf_{\substack{\mu \in \mathcal{M}(K) \\ \operatorname{supp}(\mu) \subseteq \mathcal{Y} \\ Q(\mu)=1}} E(\mu) \right]^{-1}.$$

A minimizer *on $\mathcal{Y}$, $\mu_{\min,K,\mathcal{Y}}$, when it exists, is defined as the charge distribution with unit total charge that attains this minimum energy:*

$$E(\mu_{\min,K,\mathcal{Y}}) = \frac{1}{C_K(\mathcal{Y})}.$$

The minimizer denoted on $\mathcal{X}$ is denoted simply by $\mu_{\min,K}$, and the corresponding capacity is denoted simply C_K. An equilibrium measure *on $\mathcal{Y}$, $\mu_{\mathrm{e},K,\mathcal{Y}}$, when it exists, is defined as the charge distribution with unit total charge that induces a constant potential field on $\mathcal{Y}$:*

$$f_{\mu_{\mathrm{e},K,\mathcal{Y}}}(\boldsymbol{x}) = \int_{\mathcal{Y}} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu_{\mathrm{e},K,\mathcal{Y}}(\boldsymbol{y}) = \frac{1}{C_{K,e}(\mathcal{Y})} \quad \forall \boldsymbol{x} \in \mathcal{Y}$$

for some constant $C_{K,e}(\mathcal{Y})$. The equilibrium measure on $\mathcal{X}$ is denoted simply $\mu_{\mathrm{e},K}$. See Theorem 3.3 below for uniqueness.

Note that the capacity is an increasing function, i.e, $\mathcal{Y} \subseteq \mathcal{Z}$ implies $C_K(\mathcal{Y}) \leq C_K(\mathcal{Z})$. Physically, this is because as the support of the charge distribution increases, the charges may spread out to give a smaller energy. To facilitate the proof of the uniqueness of the equilibrium measure we observe a lemma that follows from the above definition:

Lemma 3.2 *Let $\mu_{\mathrm{e},K}$ denote any equilibrium measure on $\mathcal{X}$ (when it exists) and $\mu, \nu \in \mathcal{M}(K)$ denote any measures. It follows that the energy of μ may be decomposed as the sum of two parts:*

$$E(\mu) = E\left(\mu - Q(\mu)\mu_{\mathrm{e},K}\right) + [Q(\mu)]^2 E(\mu_{\mathrm{e},K}).$$

Moreover, the inner product for measures defined in (6)*, $\langle \mu, \nu \rangle_{\mathcal{M}(K)}$, may also be decomposed:*

$$\langle \mu, \nu \rangle_{\mathcal{M}(K)} = \langle \mu - Q(\mu)\mu_{\mathrm{e},K}, \nu - Q(\nu)\mu_{\mathrm{e},K} \rangle_{\mathcal{M}(K)} + Q(\mu)Q(\nu)E(\mu_{\mathrm{e},K}). \quad (14)$$

Note that this lemma is also true if the support of the measures is restricted to some $\mathcal{Y} \subseteq \mathcal{X}$ and $\mu_{\mathrm{e},K}$ is replaced by $\mu_{\mathrm{e},K,\mathcal{Y}}$, an equilibrium measure on $\mathcal{Y}$. This lemma is now used to show the uniqueness of the equilibrium measure.

Theorem 3.3 *If a minimizer exists on $\mathcal{Y}$, it is unique. If an equilibrium measure on $\mathcal{Y}$ exists, it is unique. When $\mu_{\mathrm{e},K,\mathcal{Y}}$ exists, it is the same as $\mu_{\min,K,\mathcal{Y}}$, and $f_{\mu_{\mathrm{e},K,\mathcal{Y}}}$ is the constant inverse of the capacity, $C_K(\mathcal{Y})^{-1}$.*

Note that the above definition may also be extended to conditionally positive definite kernels, K, and the above also hold in this case because their proofs only require that the kernel be conditionally positive definite, i.e., that $E^{1/2}$ is a norm on the space where $Q(\mu) = 0$. However, for conditionally positive definite kernels the capacity may be infinite or negative. Thus, the capacity may no longer be an increasing function, but the inverse capacity (minimum energy of a unit charge distribution) remains a decreasing function.

3.2 Discrepancy Involving Equilibrium Measures

Let ν be any measure with unit total charge. We have that $E(\nu) = E(\nu - \mu_{\mathrm{e},K}) + E(\mu_{\mathrm{e},K})$, or equivalently, $E(\mu_{\mathrm{e},K} - \nu) = E(\nu) - E(\mu_{\mathrm{e},K})$. Thus,

choosing ν to make $E(\mu_{\mathrm{e},K} - \nu)$ small is equivalent to choosing ν with small energy. This relationship, together with the theorems and corollary above shows that *measures or points with small energy are the best for approximating integrals with respect to the equilibrium measure.*

Corollary 3.4 *Let $\nu \in \mathcal{M}(K)$ be some distribution with unit total charge, a particular case of which is the empirical distribution, $\mu_{\mathcal{P}}$, for the design $\mathcal{P} = \{\boldsymbol{z}_i\}_{i=1}^n \subseteq \mathcal{X}$. It follows that*

$$D(\nu; \mu_{\mathrm{e},K}) = \sqrt{E(\nu) - E(\mu_{\mathrm{e},K})}, \qquad D(\mathcal{P}; \mu_{\mathrm{e},K}) = \sqrt{E(\mathcal{P}) - E(\mu_{\mathrm{e},K})}.$$

It then follows that

$$\sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu_{\mathrm{e},K}(\boldsymbol{x}) - \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{x}) \right| = \sqrt{E(\nu) - E(\mu_{\mathrm{e},K})},$$

$$\sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu_{\mathrm{e},K}(\boldsymbol{x}) - \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{z}_i) \right| = \sqrt{E(\mathcal{P}) - E(\mu_{\mathrm{e},K})}.$$

Unfortunately, the measure defining the integration problem of interest may not be the equilibrium measure for the kernel of interest, K. In this case, a simple modification of K yields a kernel with the desired equilibrium measure.

Let K be a symmetric, conditionally positive definite kernel on $\mathcal{X}$ with equilibrium measure $\mu_{\mathrm{e},K}$ and capacity C_K. Let $\langle \mu, \nu \rangle_{\mathcal{M}(K)} = \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{y})$ as in (6), even though K may not be strictly positive definite and C_K is not necessarily positive. The above inner product decomposition suggests the definition of a new inner product in terms of an arbitrary measure $\tilde{\mu}$ with unit total charge:

$$\begin{aligned}
\langle \mu, \nu \rangle_{\mathcal{M}(\tilde{K})} &:= \langle \mu - Q(\mu)\tilde{\mu}, \nu - Q(\nu)\tilde{\mu} \rangle_{\mathcal{M}(K)} + \frac{Q(\mu)Q(\nu)}{C} \\
&= \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{y}) - Q(\mu) \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\tilde{\mu}(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{y}) \\
&\qquad - Q(\nu) \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \, \mathrm{d}\tilde{\mu}(\boldsymbol{y}) \\
&\qquad + Q(\mu)Q(\nu) \int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\tilde{\mu}(\boldsymbol{x}) \, \mathrm{d}\tilde{\mu}(\boldsymbol{y}) + \frac{Q(\mu)Q(\nu)}{C} \\
&= \int_{\mathcal{X}^2} \tilde{K}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{y}),
\end{aligned} \tag{15}$$

where this new inner product is defined in terms of the kernel $\tilde{K}$, defined as follows:

$$\begin{aligned}\tilde{K}(\boldsymbol{x},\boldsymbol{y}) &:= \int_{\mathcal{X}^2} K(\boldsymbol{t},\boldsymbol{z})\,\mathrm{d}[\delta_{\boldsymbol{x}}-\tilde{\mu}](\boldsymbol{t})\,\mathrm{d}[\delta_{\boldsymbol{y}}-\tilde{\mu}](\boldsymbol{z}) + \frac{1}{C} \\ &= K(\boldsymbol{x},\boldsymbol{y}) - \int_{\mathcal{X}^2} K(\boldsymbol{x},\boldsymbol{z})\,\mathrm{d}\tilde{\mu}(\boldsymbol{z}) - \int_{\mathcal{X}^2} K(\boldsymbol{t},\boldsymbol{y})\,\mathrm{d}\tilde{\mu}(\boldsymbol{t}) \\ &\qquad + \int_{\mathcal{X}^2} K(\boldsymbol{t},\boldsymbol{z})\,\mathrm{d}\tilde{\mu}(\boldsymbol{t})\,\mathrm{d}\tilde{\mu}(\boldsymbol{z}) + \frac{1}{C}. \end{aligned} \tag{16}$$

This new kernel $\tilde{K}$ is symmetric by definition. Its (strict) positive definiteness may be verified by checking that the energy defined $\tilde{K}$ of a nonzero charge distribution, μ, is positive.

The discussion above is summarized in the following proposition. The corollary that follows shows that the best design, $\mathcal{P}$ for numerical integration of potential functions in $\mathcal{H}(K)$ is the one that minimizes the energy defined by kernel $\tilde{K}$.

Proposition 3.5 *Let K be a symmetric, conditionally positive definite kernel on $\mathcal{X}$ with equilibrium measure $\mu_{\mathrm{e},K}$ and capacity C_K. Let $\tilde{\mu}$ be an arbitrary measure in $\mathcal{M}(K)$ with unit total charge, and let C be any positive number. Then, the kernel $\tilde{K}$ defined by* (16) *is symmetric and positive definite kernel with capacity $C_{\tilde{K}} = C$ and equilibrium measure $\mu_{\mathrm{e},\tilde{K}} = \tilde{\mu}$.*

Corollary 3.6 *Let K be a symmetric, positive definite kernel on $\mathcal{X}$. Let $\tilde{\mu}$ be an arbitrary measure in $\mathcal{M}(K)$ with unit total charge, and let the symmetric, positive definite kernel $\tilde{K}$ be given. Let D, and E denote the discrepancy and energy, respectively, defined by kernel K, and let $\tilde{D}$ and $\tilde{E}$ denote the analogous quantities for kernel $\tilde{K}$. Let $\nu \in \mathcal{M}(K)$ be any distribution with unit total charge, a particular case of which is the empirical distribution, $\mu_{\mathcal{P}}$, for the design $\mathcal{P} = \{\boldsymbol{z}_i\}_{i=1}^n \subseteq \mathcal{X}$. It follows that*

$$\begin{aligned}\sup_{\|f\|_{\mathcal{H}(K)}\leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\tilde{\mu}(\boldsymbol{x}) - \int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{x}) \right| &= D(\nu;\tilde{\mu}) = \sqrt{E(\tilde{\mu}-\nu)} \\ \sup_{\|f\|_{\mathcal{H}(\tilde{K})}\leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\tilde{\mu}(\boldsymbol{x}) - \int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{x}) \right| &= \tilde{D}(\nu;\tilde{\mu}) = \sqrt{\tilde{E}(\tilde{\mu}-\nu)} \\ &= \sqrt{\tilde{E}(\nu) - \tilde{E}(\tilde{\mu})}. \end{aligned}$$

$$\sup_{\|f\|_{\mathcal{H}(K)}\leq 1}\left|\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\tilde{\mu}(\boldsymbol{x})-\frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{z}_i)\right| = D(\mathcal{P};\tilde{\mu}) = \sqrt{E(\tilde{\mu}-\mu_{\mathcal{P}})}$$

$$\sup_{\|f\|_{\mathcal{H}(\tilde{K})}\leq 1}\left|\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\tilde{\mu}(\boldsymbol{x})-\frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{z}_i)\right| = \tilde{D}(\mathcal{P};\tilde{\mu}) = \sqrt{\tilde{E}(\tilde{\mu}-\mu_{\mathcal{P}})}$$

$$= \sqrt{\tilde{E}(\mathcal{P})-\tilde{E}(\tilde{\mu})}.$$

4 Group Invariance and Equilibrium Measures

Sometimes the domain, $\mathcal{X}$, and the kernel, K, are invariant under a group of transformations. This invariance may be used to facilitate finding the equilibrium measure and the minimum energy or discrepancy points. Specifically, suppose that one has a compact, perhaps non-Abelian, group $\mathcal{G}$ of measurable maps of $\mathcal{X}$ and $\mathcal{G}$ acts transitively on $\mathcal{X}$. The 'transitive' condition means that for any point $\boldsymbol{\eta} \in \mathcal{X}$, its orbit, $G\boldsymbol{\eta} = \{g\boldsymbol{\eta} \ : \ g \in \mathcal{G}\}$ is all of $\mathcal{X}$. When $\mathcal{X}$ is viewed as the orbit of some point $\boldsymbol{\eta}$ that point is often referred to as a pole. By convention, given a pole $\boldsymbol{\eta}$, $g_{\boldsymbol{x}}$ will denote any element of $\mathcal{G}$ that maps the pole into $\boldsymbol{x}$, i.e., $\boldsymbol{x} = g_{\boldsymbol{x}}\boldsymbol{\eta}$. Some natural examples of spaces with transitive groups continuously are:

i) The unit d-sphere, $\mathcal{X} = S^d \subset \mathbb{R}^{d+1}$, which is the orbit of any unit vector under the action of $SO(d+1)$, the group of $d+1$ dimensional orthogonal matrices of determinant 1. The standard 2-sphere, S^2, with its 'north' pole, $(0,0,1)^T$, i.e. $x = 0, y = 0, z = 1$, is the inspiration for the pole terminology.

ii) The flat d-torus, $\mathbb{T}^d = (S^1)^d \subset (\mathbb{R}^2)^d$, which is the orbit of the point $((1,0),(1,0),\ldots,(1,0))$ under rotation by $(\theta_1,\theta_2,\ldots,\theta_d)$. Since $\mathbb{T}$ is just the compact quotient group, $\mathbb{R}/2\pi\mathbb{Z}$, the flat d-torus is the product group $\mathcal{G} = (\mathbb{R}/2\pi\mathbb{Z})^d$.

iii) A non-flat 2-torus in $\mathbb{R}^3$ given by

$$\{(x,y,z) \ : \quad x = (r_1 + r_2\sin(\theta_2))\cos(\theta_1),\ y = (r_1 + r_2\sin(\theta_2))\sin(\theta_1),$$
$$z = r_2\cos(\theta_2),\ 0 \leq \theta_1 < 2\pi, 0 \leq \theta_2 < 2\pi\},$$

with fixed $r_1 > r_2 > 0$. The group $\mathcal{G} = (\mathbb{R}/2\pi\mathbb{Z})^2$ acts transitively via translation in the θ_1, θ_2 coordinates.

iv) An important variant of the flat d-torus is the half-open unit d-cube, $[0,1)^d$. This admits a transitive measurable action of the compact group $(\mathbb{R}/\mathbb{Z})^d$ given by translation modulo one. To understand that this is a measurable action it suffices to note first that the composite map $[0,1) \to \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ is a continuous bijection with a measurable inverse. Then applying this inverse to the second factor and the image in the continuous product map $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ yields a measurable action $\mathbb{R}/\mathbb{Z} \times [0,1) \to [0,1)$ which is just translation modulo one. The d-fold product of this action is the desired transitive measurable action.

Definition 4.1 *Suppose $\mathcal{X}$ admits a transitive, measurable action by a group $\mathcal{G}$. Then $\mathcal{G}$ actions on finite signed Borel measures, measurable functions and measurable kernels on $\mathcal{X}$ are defined for any $g \in \mathcal{G}$, any $\mu \in \mathcal{B}$, measurable function f or measurable kernel K by:*

i) $(g \cdot \mu)(\mathcal{Y}) := \mu(g\mathcal{Y})$ *for all Borel sets,* $\mathcal{Y} \subseteq \mathcal{X}$.

ii) $(g \cdot f)(\boldsymbol{x}) := f(g^{-1}\boldsymbol{x})$ *for all* $\boldsymbol{x} \in \mathcal{X}$.

iii) $(g \cdot K)(\boldsymbol{x}, \boldsymbol{y}) := K(g^{-1}\boldsymbol{x}, g^{-1}\boldsymbol{y})$ *for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.

Definition 4.2 *A $\mathcal{G}$-invariant measure on $\mathcal{X}$, $\mu_{\mathcal{G}}$, is a measure with* total unit charge, *$Q(\mu_{\mathcal{G}}) = 1$, for which $g \cdot \mu_{\mathcal{G}} = \mu_{\mathcal{G}}$ for all $g \in \mathcal{G}$. Any kernel K is called $\mathcal{G}$-invariant iff $g \cdot K = K$ for all $g \in \mathcal{G}$.*

Every compact group, $\mathcal{G}$, acts transitively on itself via the product map and has a unique $\mathcal{G}$-invariant measure, called the *normalized Haar measure* on $\mathcal{G}$, which we denote by $\lambda_{\mathcal{G}}$. When $\mathcal{X}$ admits a measurable transitive action by $\mathcal{G}$, then the Haar measure on $\mathcal{G}$ induces a $\mathcal{G}$-invariant measure $\mu_{\mathcal{G}}$ defined relative to a choice of pole $\boldsymbol{\eta} \in \mathcal{X}$ by:

$$\mu_{\mathcal{G}}(\mathcal{Y}) = \lambda_{\mathcal{G}}(\{g \in G \ : \ g\boldsymbol{\eta} \in \mathcal{Y}\})$$

for any measurable $\mathcal{Y} \subseteq \mathcal{X}$. The $\mathcal{G}$-invariance and normalization of $\mu_{\mathcal{G}}$ follows immediately from the corresponding properties of Haar measure. Similarly if $\boldsymbol{x} = g_{\boldsymbol{x}}\boldsymbol{\eta}$ is another pole then

$$\begin{aligned} \mu_{\mathcal{G}}(\mathcal{Y}) = \quad & \lambda_{\mathcal{G}}(\{g \in G \ : \ g\boldsymbol{\eta} \in \mathcal{Y}\}) \quad && = \lambda_{\mathcal{G}}(\{g \in G \ : \ g{g_x}^{-1}\boldsymbol{x} \in \mathcal{Y}\}) \\ = \quad & \lambda_{\mathcal{G}}(\{g \in G \ : \ g\boldsymbol{x} \in \mathcal{Y}\}g_x) && = \lambda_{\mathcal{G}}(\{g \in G \ : \ g\boldsymbol{x} \in \mathcal{Y}\}) \end{aligned}$$

since $\lambda_{\mathcal{G}}$, the Haar measure on a compact group, is invariant under both left and right translations. Thus the definition of $\mu_{\mathcal{G}}$ is independent of the choice of pole.

Each of the first two examples above is a(n) (algebraic) submanifold of the containing Euclidean space. As such there is a local orthogonal coordinate system for the Euclidean space around a pole, $\boldsymbol{\eta} \in \mathcal{X}$, such that the first d coordinates form an orthogonal coordinate system along the submanifold and the remaining coordinates form a coordinate system along submanifolds perpendicular to the original manifold. This local coordinate system provides a splitting of Lebesgue measure as a product $dx_T\ dx_N$ of a tangential, $dx_T = dx_1 \cdots dx_d$ and a normal component. Since $\mathcal{G}$ acts as orthogonal transformations of the containing Euclidean space, the product splitting can be transformed over $\mathcal{X} = \mathcal{G}\boldsymbol{\eta}$ and the tangential component of Lebesgue measure along $\mathcal{X}$ is $\mathcal{G}$-invariant and when normalized is the measure $\mu_{\mathcal{G}}$.

In the third example the measure $d\theta_1\ d\theta_2/(4\pi^2)$ is $\mathcal{G}$-invariant. While in the last example Lebesgue measure is invariant under translations modulo one (in each coordinate).

For $\mathcal{G}$-invariant kernels, it follows that

$$K(\boldsymbol{x}, \boldsymbol{y}) = K(g_{\boldsymbol{y}}^{-1}\boldsymbol{x}, g_{\boldsymbol{y}}^{-1}\boldsymbol{y}) = K(g_{\boldsymbol{y}}^{-1}\boldsymbol{x}, \boldsymbol{\eta}),\ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X},$$

where $g_{\boldsymbol{y}}^{-1}$ denotes the group inverse of $g_{\boldsymbol{y}}$, any element that takes the pole to $\boldsymbol{y}$. This implies that $\mathcal{G}$-invariant kernels are defined by a function of a single variable, $K(\cdot, \boldsymbol{\eta})$.

Theorem 4.3 *Let K be a $\mathcal{G}$-invariant kernel on $\mathcal{X}$ satisfying* (1)*. Then the $\mathcal{G}$ action on $\mathcal{M}(K)$ preserves the energy and inner product. Moreover, if $\mathcal{M}(K) \neq \{0\}$ then the $\mathcal{G}$-invariant measure $\mu_{\mathcal{G}}$ has finite energy and is the equilibrium measure and energy minimizer in $\mathcal{M}(K)$.*

Theorem 4.4 *Let $\lambda_{\mathcal{G}}$ be the normalized Haar($\mathcal{G}$-invariant) measure on $\mathcal{G}$ meaning that for any measurable set $\mathcal{G}_1 \subseteq \mathcal{G}$, and any $g \in \mathcal{G}$, $\lambda_{\mathcal{G}}(\mathcal{G}_1) = \lambda_{\mathcal{G}}(g\mathcal{G}_1)$'. Then the kernel obtained by filtering K as follows,*

$$K_{\mathcal{G}}(\boldsymbol{x}, \boldsymbol{y}) := \int_{\mathcal{G}} (g \cdot K)(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\lambda_{\mathcal{G}}(g) = \int_{\mathcal{G}} K(g^{-1}\boldsymbol{x}, g^{-1}\boldsymbol{y})\, \mathrm{d}\lambda_{\mathcal{G}}(g), \qquad (17)$$

is $\mathcal{G}$-invariant, and if K was $\mathcal{G}$-invariant to begin with, then $K_{\mathcal{G}} = K$.

Furthermore, if $\mathcal{M}(K_{\mathcal{G}}) \neq \{0\}$ then $\mu_{\mathcal{G}} \in \mathcal{M}(K_{\mathcal{G}})$, and it is the equilibrium and minimum energy measure for $K_{\mathcal{G}}$, i.e., $\mu_{\mathrm{e},K_{\mathcal{G}}} = \mu_{\mathcal{G}} = \mu_{\min,K_{\mathcal{G}}}$. Finally, mean energy and the root mean square discrepancy of a measure ν under

the group $\mathcal{G}$ with respect to the kernel K are the energy and discrepancy of ν with respect to the kernel $K_{\mathcal{G}}$, i.e.,

$$\int_{\mathcal{G}} E(g \cdot \nu) \, \mathrm{d}\lambda_{\mathcal{G}}(g) = E_{\mathcal{G}}(\mu), \tag{18a}$$

$$\int_{\mathcal{G}} D^2(g \cdot \nu; \mu_{\mathcal{G}}) \, \mathrm{d}\lambda_{\mathcal{G}}(g) = D^2_{\mathcal{G}}(\nu; \mu_{\mathcal{G}}) = E_{\mathcal{G}}(\nu) - E_{\mathcal{G}}(\mu_{\mathcal{G}}). \tag{18b}$$

Here $E_{\mathcal{G}}$ and $D_{\mathcal{G}}$ denote the energy and discrepancy defined using the filtered kernel $K_{\mathcal{G}}$.

Remark 4.5 *We have shown that our existence results hold when $\mathcal{X}$ is a $\mathcal{G}$-orbit of a group of* measurable *maps acting on $\mathcal{X}$ (or even on some set containing $\mathcal{X}$) and that the kernel $K_{\mathcal{G}}$ admits some non-trivial finite energy measure. These results should be compared to those of Section 4 below which hold whenever $\mathcal{X}$ is the orbit of a group of* isomorphisms *of the containing space.*

5 Energies, Group Invariant Kernels and Numerical Integration on Compact Manifolds

5.1 Harmonic analysis on compact homogeneous manifolds

In the sequel we will assume further that M^d is a compact homogeneous C^∞ d-dimensional manifold embedded as the orbit of a compact group G of isometries of $\mathbb{R}^{d+k}$; i.e. there is an $\eta \in M$ (often referred to as the pole) such that $M = \{g\eta : g \in G\}$. In fact, for any $\zeta = g\eta \in M$, since $G = Gg$, $M = G\eta = Gg\eta = G\zeta$. So any point in M can be chosen as the pole. In the special case that for each pair $x, y \in M$, there is a $g_{x,y} \in G$ with $g_{x,y}x = y$ and $g_{x,y}y = x$, M is called a *reflexive* compact homogeneous manifold. Natural reflexive examples to keep in mind are S^d, $d > 1$, the d-dimensional spheres. Each sphere is realized as the subset of $\mathbb{R}^{d+1}$ which is the orbit of any unit vector under the action of $SO(d+1)$, the group of $d+1$ dimensional orthogonal matrices of determinant 1. For $x, y \in S^d$ $g_{x,y} \in SO(d+1)$ can be chosen to be any rotation by π radians about any diameter of a great circle containing x, y which joins the antipodal bisectors of the two arcs between x, y. When $d = 1$, no such rotation of S^1 exists, so the circle realized as the orbit of the rotation group, $SO(2)$, is non-reflexive. Other non-reflexive examples are the flat tori $(S^1)^k$ realized as subsets of $\mathbb{R}^{2k} = (\mathbb{R}^2)^k$ which are the orbits of $([1\ 0]^t)^k$, the k-fold product of the

column vector $[1\ 0]^t$ under $G_k = SO(2)^k$. If the G_k are enlarged to be the groups $O(2)^k$, where $O(m)$ is the group of all m dimensional orthogonal matrices, then these homogeneous realizations of the flat tori *are* reflexive, since in the i-th plane, reflection in the bisector of the line-segment $\overline{x_i y_i}$ interchanges x_i, y_i. Henceforth, we will assume that d and k are fixed for a given M.

A kernel $\kappa : M \times M \to (0, \infty]$ is termed *zonal* (or G-invariant) if $\kappa(x, y) = \kappa(gx, gy)$ for all $g \in G$ and $x, y \in M$. Since the maps in G are isometries of Euclidean space, they preserve both Euclidean distance and the (arc-length) metric $d(\cdot, \cdot)$ induced on the components of M by the Euclidean metric. Thus the distance kernel $d(x, y)$ on M is zonal, as are all functions $\psi(d(x, y)),\ \psi\ :\ \mathbb{R} \to \mathbb{R}$. Moreover, the radial functions, $\phi(\|x - y\|)$, on Euclidean space that depend only on $\|x-y\|$, the Euclidean distance between x, y are also zonal functions. The manifold M carries a Borel *surface* (G-invariant) measure μ such that $\mu(M) = 1$, where G-invariant means

$$g \cdot \mu(B) := \mu(gB), \forall g \in G \ and \ \forall \ Borel \ sets \ B.$$

With this G-invariant measure μ, we define the inner product of real functions $f_1, f_2 : M \to \mathbb{R}$

$$(f_1, f_2) = \int_M f_1 f_2 d\mu.$$

In what follows, we will assume henceforth that a kernel κ always satisfies the following three conditions:

1. The kernel κ is continuous off the "diagonal" of $M \times M$, and is lower semi-continuous on $M \times M$. Here, the "diagonal" of $M \times M$ means the set $\{(x, y) \in M \times M : x = y\}$.

2. For each fixed $x \in M$, $\kappa(x, \cdot)$ and $\kappa(\cdot, y)$ are integrable with respect to surface measure μ; i.e., $\kappa(x, \cdot)$ and $\kappa(\cdot, y) \in L_1(\mu)$.

3. For each non-trivial finite signed measure ν on M, we have

$$\int_M \int_M \kappa(x, y) d\nu(x) d\nu(y) > 0,$$

   where the iterated integral may be infinite.

We will say that a kernel κ is admissible if κ satisfies all the three conditions above. We note that kernels satisfying (3) are referred to as "strictly positive

definite" in the literature. Examples of admissible kernels are the weighted Riesz kernels

$$\kappa(x,y) = w(x,y)\|x-y\|^{-s}, \quad 0 < s < d, \quad x,y \in M.$$

Here $w : M \times M \to (0,\infty]$ is chosen so that κ is admissible. If, in addition, w is G-invariant, then κ is zonal. Such kernels (in the case $w \equiv 1$), arise naturally in describing uniform distributions of electrons on rectifiable manifolds such as the sphere S^d. The uniformity arises because of the singularity in the kernel which forces points not to stay close to each other. If w is active, then perturbations of the distributions in the electrons are allowed. Perturbations of this type, arise for example in problems in computer modeling. For the sphere S^d, zonal type kernels were introduced into the study of discrepancy first by Damelin and Grabner.

Harmonic analysis on M, in our case, requires the construction of polynomials on M. In this regard, if Π_j is the space of all polynomials of total degree j on the space $\mathbb{R}^{d+k}$, then $P_j := \Pi_j|_M$ is the space of degree j polynomials on M. When $j < 0$, $\Pi_j = \{0\}$ and so $P_j = \{0\},\ j < 0$. We can also construct the sets $H_j := P_j \bigcap P_{j-1}^{\perp}$, where the orthogonality is with respect to the inner product $(\cdot,\cdot)$. We call H_j the *harmonic* polynomials of degree j.

It is straightforward to show that H_j is G-invariant, in the sense that $g \cdot p_j \in H_j$ for all $p_j \in H_j$ and $g \in G$, where for any function f on M, $g \cdot f$ is defined by

$$g \cdot f(x) := f(g^{-1}x), \forall g \in G \text{ and } \forall x \in M. \tag{19}$$

Moreover, each H_j has an orthogonal decomposition into irreducible G-invariant subspaces (i.e., subspaces with no proper G-invariant subspace)

$$H_j = \oplus_{l=1}^{h_j} \Xi_{j,l}.$$

The machinery above, gives the following easily proved but important fact.

Lemma 5.1 *Let M be a compact homogeneous space embedded in Euclidean space. The harmonic polynomials, $\sum_{i=0}^{\infty} H_i$ are dense in $C(M)$, where "perp", in the definition of H_i is with respect to the inner product $(.)$ induced by the tangential portion of Lebesgue measure(equivalently the measure derived from the Riemannian structure on M which is induced from the embedding).*

5.2 Group invariant kernels and smooth convolution

The *kernel operator* T_κ associated with a kernel $\kappa(x, y)$ is defined by

$$(T_\kappa f)(x) = \int_M \kappa(x, y) f(y) d\mu(y), \quad x \in M \tag{20}$$

for those Borel measurable f for which the right hand side exists. More generally, $T_\kappa(\nu)(x) := \int_M \kappa(x, y) d\nu(y)$. When κ is zonal, then T_κ is G-equivariant in the sense expressed by the following equation:

$$T_\kappa(g \cdot f) = g \cdot T_\kappa(f), \tag{21}$$

where $g \cdot f$ is defined at (19). We now form the *convolution product* of kernels κ and σ,

$$(\kappa * \sigma)(x, y) = \int_M \kappa(x, z)\sigma(z, y) d\mu(z), \ x, y \in M$$

which is the kernel whose associated operator $T_{\kappa*\sigma}$ is the product of the operators T_κ and T_σ; i.e., $T_{\kappa*\sigma} = T_\kappa T_\sigma$. When κ, σ are zonal, then it is easy to show that $\kappa * \sigma$ is itself zonal, since the product of G-equivariant operators is obviously G-equivariant. However, when κ and σ are merely symmetric, we have

$$\kappa * \sigma(x, y) = \sigma * \kappa(y, x).$$

Thus the convolution product of symmetric kernels, $\kappa * \sigma$, is symmetric when and only when κ and σ commute with respect to the convolution product, just as the product of symmetric(self-adjoint) operators is symmetric exactly when the operators commute. Now in case M is *reflexive*, then

1. Any zonal kernel κ is symmetric, since $\kappa(x, y) = \kappa(g_{x,y}x, g_{x,y}y) = \kappa(y, x)$.

2. Two zonal kernels commute since $\kappa * \sigma$ is zonal, hence symmetric.

Now let an admissible kernel κ be given. For a signed Borel measure on M, its κ-energy integral is defined by:

$$\mathcal{E}_\kappa(\nu) = \int_M \int_M \kappa(x, y) d\nu(x) d\nu(y).$$

Notice that the κ-energy is unchanged when $\kappa(x, y)$ is replaced by its symmetrized form $\frac{1}{2}(\kappa(x, y) + \kappa(y, x))$, which is also an admissible kernel. Hence, we will assume that κ is *symmetric* when dealing with questions about κ-energy.

We remark that the above integral may be infinite though from our assumptions of positive definiteness and lower semicontinuity on κ, combined with the strict convexity of the κ-energy, we know that either $\mathcal{E}_\kappa(\nu) = +\infty$ for all $\nu \neq 0$ or

$$\min_{\{\nu:\, \nu(1)=1\}} \mathcal{E}_\kappa(\nu)$$

exists and the minimizer is unique. We now show remarkably that for all compact homogeneous C^∞ manifolds M and admissible symmetric zonal kernels κ, the unique finite κ-energy minimizer above exists and is precisely the normalized surface measure μ. That this is true is by no means obvious. For the sphere, this fact was established by Damelin and Grabner for a class of zonal kernels. We have:

Lemma 5.2 *The normalized surface measure μ has finite κ-energy. Moreover, $\mathcal{E}_\kappa(\nu) > \mathcal{E}_\kappa(\nu(1)\mu)$ for all $\nu \neq \nu(1)\mu$. So among all signed ν with $\nu(1) = 1$, the κ-energy is uniquely minimized by the normalized surface measure μ.*

6 N-point discrete κ energy

Let $N \geq 1$. Let Z be a finite subset of M with $|Z| = N$. We define the N-point discrete κ-energy associated with Z by

$$E_\kappa(Z) = \frac{1}{N^2} \sum_{\substack{y,z \in Z \\ y \neq z}} \kappa(y,z).$$

Since κ is continuous off the diagonal of $M \times M$, and is lower semicontinuous on $M \times M$, the minimal N-point discrete κ-energy can be attained at some $Z^* \subset M$ with $|Z^*| = N$. That is

$$E_\kappa(Z^*) = \inf_{Z \subset M} E_\kappa(Z),$$

where the infimum is taken over all subsets Z of M with $|Z| = N$. We will simply call such a set Z^* a minimal energy configuration. It is clear that for each $g \in G$, gZ^* is also a minimal energy configuration. Heuristics suggests that probability measures supported on minimal energy configurations provide good approximation to the measure μ in the sense that the integral of a continuous $f : M \to \mathbb{R}$ with respect to μ is approximated well by a discrete sum over the points of Z. This was first shown by Damelin and Grabner for a class of unweighted Riesz kernels on the sphere S^d, $d \geq 2$, for

a class of Lipchitz functions, where μ is the rotation invariant probability measure on S^d. For the circle, S^1, it is easy to see that every minimal energy configuration corresponds to the set of vertices of a regular N-gon and are thus the best points to use for numerical integration for equally weighted quadrature rules.

We find it convenient to work with the full quadratic form

$$\sum_{y,z\in Z} \kappa(y,z).$$

However, the diagonal entries in the above quadratic form, $\kappa(x,x), x \in Z$, may not be finite. As a matter of fact, for Riesz kernels, these diagonal entries are infinity. The lower semi-continuity of the kernel κ allows us to consider approximating κ from below by a sequence of smooth kernels via convolution. To make this precise, let us fix an $\alpha_0 > 0$. Assume that, for each $0 < \alpha < \alpha_0$, σ_α is a zonal kernel such that the convolution kernel $\kappa_\alpha := \kappa * \sigma_\alpha$ is well defined and satisfies the following properties:

a. κ_α is continuous on $M \times M$.

b. κ_α is strictly positive definite.

c. $\kappa_\alpha(x,y) \le \kappa(x,y)$ for all $x, y \in M$.

d. $\lim_{\alpha\downarrow 0} \kappa_\alpha(x,y) = \kappa(x,y)$ for all $x, y \in M$.

If the above construction is possible, we say that κ is *strongly admissible*. The construction details are often delicate and entail case-by-case analysis. We offer two examples in which the criteria are all satisfied.

Example 1. Wagner studied a kind of modified Riesz kernel in the form

$$\kappa_\alpha(x,y) = (1+\alpha-xy)^{-s/2},\ x,y \in S^2,\ 0<s<1,\ \alpha>0,$$

where xy denotes the Euclidean inner product of the vectors x and y. On S^2, this kernel can be written as the convolution of the Riesz kernel $\kappa(x,y) = (1-xy)^{-1/2}$ and the smooth kernel σ_α with the Fourier-Legendre expansion

$$\sigma_\alpha(x,y) = \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} h^{2n+1} Q_n(xy),$$

in which $h = 2/(\sqrt{\alpha} + \sqrt{4+\alpha})$, and Q_n is the Legendre polynomial normalised so that $\|Q_n\|^2_{2,[-1,1]} = 2n+1$. This expansion can be found in the work of Hubbert and Baxter.

Example 2 On the 2-torus embedded in $\mathbb{R}^4$, one may form kernels as products of univariate kernels:

$$\kappa(x,y) = \rho(x_1,y_1)\rho(x_2,y_2), \quad x_1,y_1,x_2,y_2 \in S^1,$$

where

$$\rho(s,t) = |1-st|^{-1/4}.$$

The kernel

$$\rho_\alpha(s,t) = (1+\alpha-st)^{-1/4},\ s,t \in S^1,$$

can be written as a convolution of ρ with the analytic kernel

$$\sigma_\alpha(s,t) = \sum_{n=0}^{\infty} \frac{F(n+1/4,n+1/2;2n+1;\frac{4}{4+\alpha})}{F(n+1/4,n+1/2;2n+1;1)} T_n(st),$$

where T_n is the Chebyshev polynomial and $F(a,b;c;z)$ is the Gauss hypergeometric function.

6.1 Quadrature for polynomials on compact, reflexive homogeneous manifolds

In this section and henceforth, we will need to assume that M is reflexive. In this case, we need some more machinery on reflexive spaces. Firstly, if $d_{j,l} = \dim \Xi_{j,l}$ and $\{Y_{j,l}^1, \ldots, Y_{j,l}^{d_{j,l}}\}$ is any orthonormal basis for $\Xi_{j,l}$, then

$$P_{j,l}(x,y) := \sum_{m=1}^{d_{j,l}} Y_{j,l}^m(x) Y_{j,l}^m(y), \quad x,y \in M,$$

is the unique G-invariant reproducing kernel for $\Xi_{j,l}$. In other words, if $T_{j,l}$ is the orthogonal projector onto $\Xi_{j,l}$,

$$T_{j,l}f(x) := \int_M P_{j,l}(x,y) f(y) d\mu(y), \quad x \in M. \tag{22}$$

In particular,

$$T_{j,l}f(x) := (T_{j,l}f, P_{j,l}(\cdot,x)), \quad x \in M. \tag{23}$$

Moreover, for any $g \in G$, $\{g \cdot Y_{j,l}^1, \ldots, g \cdot Y_{j,l}^{d_{j,l}}\}$ is another orthonormal basis for $\Xi_{j,l}$. So the uniqueness of a reproducing kernel shows $P_{j,l}(g^{-1}x, g^{-1}y) = P_{j,l}(x,y)$; i.e., $P_{j,l}$ is zonal. From its definition, it is obvious that each $P_{j,l}$

is symmetric. Observe that the self-adjoint projectors $T_{j,l}$ associated with the G-invariant kernels $P_{j,l}$ are *kernel operators* $T_{P_{j,l}}$ defined by Equation (20). By reflexivity, we conclude that for any zonal κ T_κ commutes with each projector $T_{j,l}$ and hence $T_\kappa T_{j,l}$ is a G-equivariant self-adjoint linear transformation of $\Xi_{j,l}$. Since $\Xi_{j,l}$ is an irreducible G-invariant subspace, $T_\kappa T_{j,l} = a_{j,l} T_{j,l}$ for some scalar $a_{j,l}$, i.e. the subspace $\Xi_{j,l}$ of degree j harmonic polynomials is also a subspace of the T_κ eigenspace associated to the eigenvalue $a_{j,l}$. Moreover, if κ is admissible, so strictly positive definite, and $0 \neq p \in \Xi_{j,l}$, we have $\mathcal{E}_\kappa(p\mu) = a_{j,l}\|p\|^2_{L_2(\mu)} > 0$; i.e., all $a_{j,l} > 0$. Thus, on reflexive spaces, by the density of harmonic polynomials, each admissible zonal kernel κ has an expansion with positive coefficients (convergent in an appropriate operator norm)

$$\kappa(x,y) = \sum_{j=0}^{\infty} \sum_{l=1}^{h_j} a_{j,l}(\kappa) P_{j,l}(x,y),\ x,y \in M.$$

Let us define the *native space* $\mathcal{N}_\kappa$ for an admissible κ via the kernels defined by Equation (20) and coefficients $a_{j,l}$ defined by Equation (24).

Let

$$\mathcal{N}_\kappa := \left\{ f : \|f\|^2_{\mathcal{N}_\kappa} := \sum_{j=0}^{\infty} \sum_{l=1}^{h_j} \frac{\|T_{j,l} f\|^2}{a_{j,l}(\kappa)} < \infty \right\}.$$

We are interested in the error of integration for a class of smooth real valued functions f on M, when f is given on a point set $Z \subset M$ of finite cardinality $N \geq 1$. The error in integration is defined by

$$R(f,Z) := \int_M f(y) d\mu(y) - \frac{1}{N} \sum_{z \in Z} f(z).$$

Theorem 6.1 *Let κ be strongly admissible on M and $Z \subset M$ be a point subset of cardinality $N \geq 1$. Fix $x \in Z$. Then, for $0 < \alpha < \alpha_0$, the following estimate holds true for every $f \in \mathcal{N}_{\kappa_\alpha}$:*

$$|R(f,Z)| \quad \leq \quad \|f\|_{\mathcal{N}_{\kappa_\alpha}} \left(E_\kappa(Z) + \frac{1}{N} \kappa_\alpha(x,x) - a_{0,1}(\kappa_\alpha) \right)^{1/2}.$$

Remark The special case of a sphere S^d and a class of G invariant kernels on S^d, the above theorem was established first by Damelin and Grabner. A further elaboration on the estimate of Theorem 6.1 is also

appropriate. Notice that the right hand side of the estimate depends on both the function and the point set Z. One way to obtain tighter upper bounds in Theorem 6.1, is to link the kernel studied with the function space with a different measure of energy. This is done in Section 2 but at the price of smaller classes of functions.

7 Quadrature for smooth functions on the sphere and projective spaces

In this last section, we extend Theorem 6.1 to a class of smooth functions on projective spaces and sphere, which are examples of so-called 2-point homogeneous manifolds (see below for a definition). Let F be one of the following fields: $Q = \{r_0 + r_1 i + r_2 j + r_3 k : r_i \in \mathbb{R}\}$ (quaternions), $C = \{q \in Q : r_2 = r_3 = 0\}$ (complex) or $\mathbb{R}$. Let F have dimension m $(= 4, 2, 1$ respectively) over the reals. The length squared of an element $f \in F$ is $|f|^2 = r_0^2 + r_1^2 + r_2^2 + r_3^2$. Writing a vector $f \in F^{m+1}$ in the form $f = (f_1, f_2, \cdots, f_{m+1})$, the sphere $S(F^{m+1}) = \{f \in F^{m+1} : \sum_{i=1}^{m+1} |f_i|^2 = 1\}$. The standard definition of the projective space $P^{dm}(F)$ is the set of points on the sphere $S(F^{m+1})$, where points x and y are identified if $x = \alpha y$ for some $\alpha \in F$ with $|\alpha| = 1$.

To extend our result we require some additional machinery which we now state. We denote by $C^k(M)$, the space of k times, continuously differentiable functions $f : M \to \mathbb{R}$. It is well known that M carries an inner product and the action of G on M translates this inner product to the tangent spaces at each point in M so that M has a well defined Riemannian metric which in turn induces a well defined arc-length metric ρ on $M \times M$. To define suitable moduli of smoothness, let $\mathbf{g}$ be the natural Lie algebra on M formed by taking the set of all skew-symmetric operators D on $\mathbb{R}^{d+k}$ such that $\exp tD \in G$, $\forall t \in G$. Let G act on $C(M)$ as in (19). Then we define the space $C^1(M)$ as the space of functions $f \in C(M)$, such that for each $D \in \mathbf{g}$, there exists $D(f) \in C(M)$ such that

$$\lim_{t \to 0} ||t^{-1}(\exp tD \cdot f - f) - D(f)||_\infty = 0.$$

The space C^k, $k \in \mathbb{N}$, $k \geq 2$ is then defined inductively.

We define a *first modulus of continuity* on $C(M)$ by

$$\omega_1(f, h) := \sup \{|f(x) - f(y)| : \rho(x, y) \leq h\}$$

Similarly, if x_+, x, x_- denote equaly spaced points along a geodesic in M, then the *second modulus of continuity* on $C(M)$ is defined by way of

$$\omega_2(f,h) := \sup\left\{|f(x_+) - 2f(x) + f(x_-)| : \rho(x_+,x) \le h\right\}.$$

Now choose an orthonormal basis $D_1, ...D_j$ for $\mathbf{g}$ for some $j \ge 1$. Then define inductively for $f \in C^{k+1}$, $k \ge 0$:

$$\omega_r(f^{k+1};h) := \sum_{i=1}^{j} \omega_r\left((D_i(f))^{(k)};h\right),\ r = 1,2,$$

where $f^{(0)} = f$.

A two-point homogeneous space is one for which, given two pairs of points, x_1, y_1 and x_2, y_2 on M, with $\rho(x_1,y_1) = \rho(x_2,y_2)$, there exists a $g \in G$ such that $gx_1 = x_2$ and $gy_1 = y_2$. If this is the case then, for $\rho(x_1,y_1) = \rho(x_2,y_2)$ and a zonal kernel κ,

$$\kappa(x_2,y_2) = \kappa(gx_1,gy_1) = \kappa(x_1,y_1),$$

so that κ is a function only of the distance between the points. In this case we have a simple representation of the reproducing kernels as a univariate polynomial of an inner product.

Also, it is straightforward to see that such spaces are reflexive.

Note that for $r = 1$, for the sphere case, the construction used by Ragozin to produce q_n was introduced first by Newman and Shapiro and was used by Damelin and Grabner to prove Theorem 7.1 for $r = 1, s = 1$. The construction for the sphere uses that the reproducing kernels are essentially univariate. The univariate nature of κ also can be used to show that each H_j is irreducible, so for each j there is only one l and only one eigenvalue a_j for T_κ acting on H_j.

Theorem 7.1 *Let κ be admissible on M and $Z \subset M$ be a point subset of cardinality $N \ge 1$. There exists positive constants C, C' dependent only on d, r, s, M, such that for any $s \ge 0$, any $f \in C^s(M)$, any $n \ge 1$ and any $0 < \alpha < \alpha_0$,*

$$\begin{aligned} |R(f,Z)| \ &\le\ Cn^{-s}\omega_r\left(f^{(s)};\frac{1}{n}\right) + \\ &\qquad C'\max_{j\le n}\frac{1}{(a_j(\kappa_\alpha))^{1/2}}\|f\|_\infty\left(E_\kappa(Z) + \frac{1}{N}\kappa_\alpha(x,x) - a_0(\kappa_\alpha)\right)^{1/2}. \end{aligned}$$

Acknowledgment The work in this paper is in collaboration with P. Grabner, F. Hickernell, J. Levesley, D. Ragozin, X. Sun and Z. Zeng. It is supported, in part, by EP/C000285, NSF-DMS-0439734 and NSF-DMS-0555839. The author thanks the organisers for allowing him to write this article and for the excellent conference.

References

[1] D. Benko, *The support of the equilibrium measure*, Acta Sci Math. (Szeged), **70**(2004), pp 1-2,35-55.

[2] D. Benko, *Approximation by weighted polynomials*, J. Approx. Theory 120 (2003), no. 1, 153-182.

[3] S. B. Damelin, J. Levesley, D. Ragozin, and X. Sun, *Energies, group invariance kernels and numerical integration on compact manifolds*, submitted for publication.

[4] S. B. Damelin and P. Grabner, *Numerical integration, energy and asymptotic equidistributon on the sphere*, *Journal of Complexity*, **19** (2003), pp 231-246.

[5] S. B. Damelin, F. Hickernell, D. Ragozin and X. Zeng, *On Energy, discrepancy and $\mathcal{G}$ invariant measures on measurable subsets of Euclidean space*, submitted for publication.

[6] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New York, 1974.

[7] N. S. Lankov, *Foundations of Modern Potential Theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1972.

[8] J. Levesley and D. L. Ragozin, *Positive definite kernel interpolation on manifolds: Convergence rates*, Approximation Theory X, L. L. Schumaker and J. Stockler (eds), Vanderbilt Univ. Press (2002), pp 277-285.

[9] D. Ragozin, *Constructive polynomial approximation on spheres and projective spaces*, TAMS, **162**(1971), pp 157-170.

[10] X. Sun, *Scattered cardinal Hermite interpolation by positive definite functions*, Computers and Mathematics with Applications 40 (2000), 165-169.

[11] X. Sun, *Strictly positive definite functions on the unit circle*, Mathematics of Computation 74 (2005), 709-721.