On Energy, Discrepancy and Group Invariant Measures on Measurable Subsets of Euclidean Space

S. B. Damelin [a] F. J. Hickernell [b] D. L. Ragozin [c] X. Zeng [d]

[a] *School of Computational and Applied Mathematics, Private Bag 3, University of the Witwatersrand, 2050, South Africa,* steve.damelin@gmail.com

[b] *Department of Applied Mathematics, Illinois Institute of Technology, 10 W. 32nd Street, E1-208, Chicago, IL 60616,* hickernell@iit.edu

[c] *Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350,* rag@math.washington.edu

[d] *Argonne National Laboratory, 9700 South Cass Avenue, Building 240 Argonne, IL 60439-4844,* zeng@mcs.anl.gov

Abstract

Given $\mathcal{X}$, some measurable subset of Euclidean space, one sometimes wants to construct a finite set of points, $\mathcal{P} \subset \mathcal{X}$, called a design, with a small energy or discrepancy. Here it is shown that these two measures of design 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 measure μ defining the integration problem is the equilibrium measure or charge distribution corresponding to the energy kernel, K.

If $\mathcal{X}$ is 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, then it is shown 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.

Key words: capacity, cubature, discrepancy, distribution, group invariant kernel, group invariant measure, energy minimizer, equilibrium measure, numerical integration, positive definite, potential field, Riesz kernel, 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. See for example [10,17,24,25] and the references cited therein.

Another approach to spreading points uniformly, developed initially for the d-dimensional unit cube, $[0,1]^d$, is the *discrepancy* of Weyl [40]. 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 [6,19]. In the past decade attention has turned to cases where the discrepancy, $D(\mathcal{P};\mu)$, is defined as a *Hilbert space norm* of the difference between some arbitrary target distribution, μ, and the empirical distribution of $\mathcal{P}$ [20]. Such discrepancies appear as tight error bounds for numerical integration [18–20,29] and provide insight into the tractability of numerical integration, i.e., the dependence of the error as the dimension of the region tends to infinity [30]. The notion of discrepancy has also been extended to more arbitrary sets, $\mathcal{X}$, than the unit cube [19,21].

Besides energy and discrepancy, other distance-based measures of even spread include the fill distance (also known as the mesh norm or sphere covering radius) and the separation distance (also known as the sphere packing distance). See [14,16,23,28,33,39] and the references cited therein, for discussions of these concepts. Measures of quality placement of points arise in both the numerical analysis and statistics literatures, where $\mathcal{P}$ is known as the design. The JMP statistical package [34] offers minimum energy, minimum discrepancy, and sphere packing designs among its options for space-filling designs.

The literatures for these various kinds of even or space-filling points have developed rather independently because different quality measures may lead to different sets of points. It is shown here that *minimum energy points and*

minimum discrepancy points are the same under quite general assumptions. Although the mathematical argument supporting this statement is rather elementary, this equivalence does not appear to have been observed in either the energy or discrepancy literatures. There are some wrinkles in this equivalence that are suggested in the paragraphs below and explained more fully in the sections that follow.

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 (see Corollary 10):

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

where μ_{e} is the equilibrium charge distribution. A more general error bound involving energy for integrals with respect to an arbitrary measure, μ, appears in Corollary 6. The error bound above differs from most of those found in the energy literature in that it is *tight*. In section 2, a kernel, K, is used to define both energy in (1) and discrepancy in (8). 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 Proposition 3). 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.

Section 2 of this article establishes the equivalence of discrepancy and energy and their relationship to error estimates for quadrature over measurable subsets $\mathcal{X}$ of Euclidean space and for integrands $f:\mathcal{X}\to\mathbb{R}$. In the energy community there have been a number of results deriving quadrature error bounds in terms of energy. For $\mathcal{X}=[-1,1]$ the n nodes of the celebrated Gaussian quadrature formula are uniquely determined as the zeros of the unique monic polynomial of minimal mean-square deviation on $\mathcal{X}$. In [8], this idea is extended to spheres to obtain upper bounds for numerical integration using energy functionals of an extremal Riesz energy problem and a class of invariant

kernels defined on the spheres. In [12], upper bound quadrature error estimates on classes of compact, homogeneous spaces embedded in Euclidean space are proven via energy functionals defined by way of a class of invariant kernels which are generalizations of zonal kernels on the sphere or radial kernels in Euclidean space. In the results of [8,12], the discrepancy bound splits into two parts. The first depending on the nodal set and the second depending on the function space. In [1,2,4,35], the authors have investigated discrepancy estimates using potentials on smooth curves in the complex plane and on spheres in $d \geq 1$ Euclidean space.

The discrepancy literature has provided many tight upper bounds on numerical integration error. Niederreiter [29] a describes the seminal work of Koksma and Hlawka to obtain an error bound for the integration domain $[0,1]^d$ as the product of the star discrepancy of Weyl and the bounded variation of the integrand. Niederreiter also describes the progress made in constructing low discrepancy sets. Since the publication of [29] the Koksma-Hlawka inequality has been generalized in terms of discrepancies defined by symmetric positive definite kernels, e.g., [18–21] and for measures of integrand roughness corresponding to the semi-norms in Hilbert spaces where these kernels act as reproducing kernels. The choice of kernel reflects the smoothness, possible periodicity and other properties one wishes to assume about the integrands.

Reflecting on these parallel developments, it is observed in this article that one may construct tight quadrature error bounds on measurable subsets $\mathcal{X}$ of Euclidean space using energy functionals defined by way of a class of kernels $K : \mathcal{X}^2 (= \mathcal{X} \times \mathcal{X}) \to \mathbb{R}$. The key is to directly link the space of possible integrands to the energy kernel via a reproducing kernel Hilbert space.

Section 4, 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. See [3,7,9,11,13,22,26] and [27,32,37,38] and the references therein. Recently, in [12], the authors used the idea of group invariant kernels and measures to study equilibrium measures over various compact sets in Euclidean space. In the current paper, we show, (see Theorem 16 for complete details), that if $\mathcal{X}$ is the orbit of a compact, possibly non-Abelian, group, $\mathcal{G}$, of measurable transformations, then the equilibrium measure for the induced $\mathcal{G}$ invariant kernel $K_{\mathcal{G}}$ is the normalized group invariant measure on $\mathcal{X}$ induced by the Haar measure on $\mathcal{G}$. This allows us to calculate explicit further representations of equilibrium measures which even in one dimension are new. We note that even in one dimension, calculation of explicit representations of equilibrium measures is not well understood and in many cases, even their supports have only been recently determined. See [3,9,11,13,27,22,32,37,38]. Group invariant measures also appear in the discrepancy literature, such as the rotations of [5], and the

digital shifts and scramblings of [31]. Examples of equilibrium measures are provided in Section 5.

2 Kernels, Charge Distributions and Potential Fields

In this section, we define some central notions in the study of energy and discrepancy. 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 $Q(\mu) := \mu(\mathcal{X}) = \int_{\mathcal{X}} \mathrm{d}\mu(\boldsymbol{x})$, which in the energy interpretation, is the *total charge* distributed on $\mathcal{X}$.

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, $\forall \mu \in \mathcal{B}$

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

Here, $E(\mu)$ denotes 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. 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 $Q(\mu) = 0$, i.e., the energy is positive for distributions with zero total charge. The kernel itself may be infinite at some points in $\mathcal{X}^2$, e.g., one important kernel defined on $\mathbb{R}$ arising from electrostatic energy is $K : (x, y) \mapsto -\log(|x - y|)$. 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.

Example 1 (Riesz Kernel) *The* generalized Riesz kernels*, also known as the generalized multiquadric kernels are defined by*

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

When $\varepsilon = 0$ the term generalized *is omitted. Each kernel, $K_{s,\varepsilon}$, is symmetric in its arguments. It is (strictly) positive definite on $\mathbb{R}^d$ if either $s \geq 0$ and $\varepsilon > 0$, or alternatively $0 \leq s < d$ and $\varepsilon = 0$ [15, Sect. 4.5]. It is* conditionally positive definite *for $-2 < s < 0$ [15, Sect. 8.1–2]. For $s \geq d$ and $\varepsilon = 0$, the energy for the Riesz kernel is infinite for nonzero charge distributions.*

The logarithmic potential, mentioned in the paragraph preceding this example, and the Coulombic potential, $\|\boldsymbol{x} - \boldsymbol{y}\|_2^{-1}$, are special cases of the Riesz kernel.

Example 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 $\mathcal{P} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$:

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

The reasons that we prefer E to E_{phys} are discussed in Section 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 in (1) may be used to define an inner product. Let $\mathcal{M}(K) \subseteq \mathcal{B}$ be the set of measures with finite energy, $E(\mu)$. We assume $\mathcal{M}(K)$ is a linear space. This is always true when K is bounded and it

is true but non-trivial to prove for the singular Riesz kernels which are strictly or conditionally positive definite. 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 in [19] 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:

Proposition 3 *For energy defined as in* (1) *and discrepancy defined as in* (8) *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 [40], also called the L_∞ star discrepancy [29], 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 $\boldsymbol{x} = (x_1, \ldots, x_d)$, $\mu(\boldsymbol{x}) := \mu(\prod_{k=1}^{d}(-\infty, x_k])$. However, the closely related L_2 star discrepancy,

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

and its relatives [18] are special cases of (8).

Example 4 (Product Kernels) *The L_2 star discrepancy for $[0,1]$ is the discrepancy associated to the kernel $K(x,y) = 1 - |x-y|$ [18], 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 [18]:*

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

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 $|\mu|$ almost everywhere 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 integral $L_\mu(f_\nu)$ may be interpreted physically as the energy of the charge distribution

μ under the potential field f_ν. The Cauchy-Bunyakovsky-Schwarz 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 5 *For any $\mu, \nu \in \mathcal{M}(K)$ it follows that*

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

PROOF. By (13), the definition of the norm in $\mathcal{H}(K)$, and the Cauchy-Bunyakovsky-Schwarz inequality, one may write

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

which completes the proof. □

An immediate consequence of this theorem is that the maximum value of an integral for potential fields with norm no greater than unity is simply the square root of the energy of the charge distribution defining the integral:

$$\begin{aligned} \sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) \right| &= D(0; \mu) \\ &= \sqrt{E(\mu)} = \sqrt{\int_{\mathcal{X}^2} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\mu(\boldsymbol{x}) \mathrm{d}\mu(\boldsymbol{y})}. \quad (14) \end{aligned}$$

This implies that $|\int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x})| \leq \sqrt{E(\mu)} \, \|f\|_{\mathcal{H}(K)}$.

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, which follows from Proposition 3 and Theorem 5.

Corollary 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*

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

The essence of Corollary 6 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 $\mathcal{H}(K)$. Kernel (11) in Example 4 defines a space of integrands that have square integrable mixed partial derivatives of order up to one in each variable [18]. 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 [18,29,36].

3 Singular Kernels and Physical Energy

The tight upper bound on the numerical integration error in Corollary 6 is the discrepancy, $D(\mathcal{P}; \mu)$. The expression for $D^2(\mathcal{P}; \mu)$ in (9), consists of three terms: the first is $E(\mu)$ and the third is $E(\mathcal{P})$. In fact, in the next section in Corollary 10 it is shown that for certain important choices of μ, $D^2(\mathcal{P}; \mu) = E(\mathcal{P}) - E(\mu)$. Thus, the energy of the sample points, $E(\mathcal{P})$ plays an important role in bounding the numerical integration error. This energy can be written in terms of the physical energy as

$$E(\mathcal{P}) = \frac{1}{n^2} \sum_{i,j=1}^{n} K(\boldsymbol{z}_i, \boldsymbol{z}_j) = \frac{1}{n^2} \sum_{i=1}^{n} K(\boldsymbol{z}_i, \boldsymbol{z}_i) + 2E_{\text{phys}}(\mathcal{P}),$$

where E_{phys} is defined in (3). For the tight upper bound on the numerical integration error to be guaranteed finite, $K(\boldsymbol{x}, \boldsymbol{x})$ must be finite for all $\boldsymbol{x} \in \mathcal{X}$. This is the case for a generalized Riesz kernel with either $\varepsilon > 0$ or $s < 0$, but not for $s \geq 0$ and $\varepsilon = 0$ together.

Unfortunately, many popular kernels in the energy literature are singular, i.e., $K(\boldsymbol{x}, \boldsymbol{x}) = \infty$ for one or more $\boldsymbol{x} \in \mathcal{X}$. The Riesz kernels with $s \geq 0$ and $\varepsilon = 0$ are prime examples. For these Riesz kernels, and many other singular kernels, $E_{\text{phys}}(\mathcal{P})$ is always finite, but $E(\mathcal{P})$ is not. Thus, some may prefer E_{phys} as the

Table 1
The energy, E, and physical energy, E_{phys}, for evenly spaced distribution of charges on a circle for the generalized Riesz kernels, (2), with $s = 1$ and various values of ε.

ε	n	1	2	4	8	16	32	64	128	256
1	E	1.000	0.724	0.650	0.643	0.643	0.643	0.643	0.643	0.643
	E_{phys}	0.000	0.224	0.400	0.518	0.580	0.611	0.627	0.635	0.639
0.01	E	100.000	50.250	25.479	13.201	7.172	4.268	2.924	2.358	2.167
	E_{phys}	0.000	0.250	0.479	0.701	0.922	1.143	1.362	1.577	1.776
0	E	∞	∞	∞	∞	∞	∞	∞	∞	∞
	E_{phys}	0.000	0.250	0.479	0.701	0.922	1.143	1.364	1.584	1.805

definition of energy rather than E as defined here. There are several reponses to this potential criticism.

First of all, note that whenever one has a singular kernel, the space of potential fields, $\mathcal{H}(K)$, automatically contains singular functions, i.e., there exist $f \in \mathcal{H}(K)$ for which $f(\boldsymbol{x}) = \infty$ for some $\boldsymbol{x} \in \mathcal{X}$. If $E(\mu)$ is finite, then the integral of any potential field, $f \in \mathcal{H}(K)$ must be finite by (14). On the other hand, the estimate of this integral by using an average of the potential field f at a finite number of points may be infinite because f may be singular. Thus, once one considers singular kernels, it is impossible to guarantee a finite numerical integration error bound.

Second, many energy kernels, including the generalized Riesz kernels, are *stationary*, i.e., $K(\boldsymbol{x}, \boldsymbol{y}) = \hat{K}(\boldsymbol{x} - \boldsymbol{y})$. For a stationary kernel

$$E(\mathcal{P}) = \frac{1}{n}\hat{K}(\mathbf{0}) + 2E_{\text{phys}}(\mathcal{P}).$$

If $\hat{K}(\mathbf{0})$ is finite, then for a fixed n, choosing $\mathcal{P}$ to minimize $E(\mathcal{P})$ is equivalent to choosing $\mathcal{P}$ to minimize $E_{\text{phys}}(\mathcal{P})$.

For example, Table 1 displays $E(\mathcal{P})$ and $E_{\text{phys}}(\mathcal{P})$ defined by the generalized Riesz kernels with $s = 1$ (Coulombic case) and various ε. The positions of the charges $\mathcal{P}$ are equally spaced on the circle, which is the minimum energy configuration for either E or E_{phys}. In all cases, E decreases as n increases, but E_{phys} *increases* as n increases.

On the other hand, Table 1 illustrates a positive use of E_{phys}. For the moment, let E_ε and $E_{\text{phys},\varepsilon}$ denote the energy and physical energy with the dependence on ε, a parameter in the definition of the defining kernel, K_ε. If $K_\varepsilon(\boldsymbol{x}, \boldsymbol{y}) \leq K_0(\boldsymbol{x}, \boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, as it is for the generalized Riesz kernels, then it follows in a straightforward manner that $E_\varepsilon(\mathcal{P}) \leq E_0(\mathcal{P})$ and $E_{\text{phys},\varepsilon}(\mathcal{P}) \leq$

$E_{\text{phys},0}(\mathcal{P})$. Table 1 shows this. Thus, $E_{\text{phys},0}(\mathcal{P})$ may be used to obtain an upper bound on $E_\varepsilon(\mathcal{P})$ as follows:

$$\begin{aligned} E_\varepsilon(\mathcal{P}) &\leq \frac{1}{n^2}\sum_{i=1}^{n} K(\boldsymbol{z}_i, \boldsymbol{z}_i) + 2E_{\text{phys},0}(\mathcal{P}) \\ &= \frac{1}{n}\hat{K}(\mathbf{0}) + 2E_{\text{phys},0}(\mathcal{P}) \quad \text{if } K \text{ is stationary.} \end{aligned}$$

In Table 1 the bound of $E_{\text{phys},\varepsilon}(\mathcal{P})$ by $E_{\text{phys},0}(\mathcal{P})$ is rather tight for small enough ε and n. Associating energy to numerical integration error for singular kernels, K_0, has been studied by considering these kernels as limits of finite-valued kernels K_ε as $\varepsilon \downarrow 0$ [8,12].

4 Equilibrium Measures and Minimizers

While Corollary 6 relates numerical integration error to the energy, it does not necessarily imply 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.

4.1 Equilibrium Measures Minimize the Energy

Definition 7 *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* $\text{supp}(\mu)$ *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:*

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

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 on $\mathcal{X}$ is denoted simply by $\mu_{\min,K}$, and the corresponding capacity is denoted simply C_K. An equilibrium measure *on $\mathcal{Y}$, $\mu_{\text{e},K,\mathcal{Y}}$, when it*

exists, is defined as the charge distribution $\mu \in \mathcal{M}(K)$ *supported in* $\mathcal{Y}$ *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 9 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 8 *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. Then* $Q(\mu)\mu_{\mathrm{e},K}$ *is the orthogonal projection of* μ *on the line through* $\mu_{\mathrm{e},K}$*. 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 be decomposed as follows:*

$$\langle \mu, \nu \rangle_{\mathcal{M}(K)} = \langle \mu - Q(\mu)\mu_{\mathrm{e},K}, \nu - Q(\nu)\mu_{\mathrm{e},K} \rangle_{\mathcal{M}(K)} + Q(\mu)Q(\nu)E(\mu_{\mathrm{e},K}). \quad (15)$$

PROOF. All the claims follow if $\mu_{\mathrm{e},K}$ is orthogonal to $\omega = \mu - Q(\mu)\mu_{\mathrm{e},K}$. This orthogonality holds since the definition of an equilibrium measure shows:

$$\begin{aligned}\langle \mu_{\mathrm{e},K}, \omega \rangle_{\mathcal{M}(K)} = \int_{\mathcal{X}^2} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu_{\mathrm{e},K}(\boldsymbol{x})\mathrm{d}\omega(\boldsymbol{y}) &= \int_{\mathcal{X}} f_{\mu_{\mathrm{e},K}}(\boldsymbol{y})\mathrm{d}\omega(\boldsymbol{y}) \\ &= \int_{\mathcal{X}} \frac{1}{C_{K,\mathrm{e}}}\,\mathrm{d}\omega(\boldsymbol{y}) = \frac{Q(\omega)}{C_{K,\mathrm{e}}} = 0. \quad (16)\end{aligned}$$

□

Note that this lemma is also true if the support of the measures is restricted to some $\mathcal{Y} \subseteq \mathcal{X}$, thereby reducing $\mathcal{M}(K)$ to such measures 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 9 *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,* $C_K(\mathcal{Y})^{-1}$*.*

PROOF. We prove this result for $\mathcal{Y} = \mathcal{X}$ because the proof for arbitrary $\mathcal{Y} \subseteq \mathcal{X}$ is analogous. Given any two minimizers, μ_1 and μ_2, the function

$$h(t) = E(\mu_1 + t(\mu_2 - \mu_1)) = E(\mu_1) + 2t \langle \mu_1, \mu_2 - \mu_1 \rangle_{\mathcal{M}(K)} + t^2 E(\mu_2 - \mu_1)$$

is a quadratic function of t that obtains its minimum value at $t = 0, 1$. Thus $h(t)$ must be constant, and the coefficient of its quadratic term must be zero, i.e., $E(\mu_2 - \mu_1) = 0$. Thus, $\mu_1 = \mu_2$ since $E^{1/2}$ is a norm on the space $\mathcal{M}(K)$, and the minimizer is unique.

Consider any equilibrium measure $\mu_{\mathrm{e},K}$, and any other measure $\mu \in \mathcal{M}(K)$ with unit total charge. Lemma 8 implies that $E(\mu) = E(\mu - \mu_{\mathrm{e},K}) + E(\mu_{\mathrm{e},K}) > E(\mu_{\mathrm{e},K})$, since $\mu \neq \mu_{\mathrm{e},K}$ and so $E(\mu - \mu_{\mathrm{e},K}) > 0$. Therefore, $\mu_{\mathrm{e},K}$ is the minimizer, and $f_{\mu_{\mathrm{e},K}}(\boldsymbol{y}) = C_{K,\mathrm{e}}^{-1} = C_K^{-1}$. Since the minimizer is unique, the equilibrium measure is also unique. □

Note that Definition 7 may also be extended to conditionally positive definite kernels, K, and Lemma 8 and Theorem 9 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 subspace of $\mathcal{M}(K)$ 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.

4.2 Discrepancy Involving Equilibrium Measures

Let ν be any measure with unit total charge. Lemma 8 implies 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.* See [8,12] for further discussion of this principle.

Corollary 10 *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 from Proposition 3 that*

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

By Theorem 5 and Corollary 6, 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 inner product decomposition in Lemma 8 suggests the definition of a new inner product in terms of an arbitrary measure $\tilde{\mu}\in\mathcal{M}(K)$ with unit total charge and an arbitrary positive constant, C:

$$\begin{aligned}\langle\mu,\nu\rangle_{\mathcal{M}(\tilde{K})}&:=\langle\mu-Q(\mu)\tilde{\mu},\nu-Q(\nu)\tilde{\mu}\rangle_{\mathcal{M}(K)}+\frac{Q(\mu)Q(\nu)}{C}\\&=\int_{\mathcal{X}^2}K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{y})-Q(\mu)\int_{\mathcal{X}^2}K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\tilde{\mu}(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{y})\qquad(17)\\&\qquad-Q(\nu)\int_{\mathcal{X}^2}K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\tilde{\mu}(\boldsymbol{y})\\&\qquad+Q(\mu)Q(\nu)\int_{\mathcal{X}^2}K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\tilde{\mu}(\boldsymbol{x})\,\mathrm{d}\tilde{\mu}(\boldsymbol{y})+\frac{Q(\mu)Q(\nu)}{C}\\&=\int_{\mathcal{X}^2}\tilde{K}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{y}),\qquad(18)\end{aligned}$$

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

$$\begin{aligned}\tilde{K}(\boldsymbol{x},\boldsymbol{y})&:=\int_{\mathcal{X}^2}K(\boldsymbol{t},\boldsymbol{z})\,\mathrm{d}[\delta_{\boldsymbol{x}}-\tilde{\mu}](\boldsymbol{t})\,\mathrm{d}[\delta_{\boldsymbol{y}}-\tilde{\mu}](\boldsymbol{z})+\frac{1}{C}\\&=K(\boldsymbol{x},\boldsymbol{y})-\int_{\mathcal{X}}K(\boldsymbol{x},\boldsymbol{z})\,\mathrm{d}\tilde{\mu}(\boldsymbol{z})-\int_{\mathcal{X}}K(\boldsymbol{t},\boldsymbol{y})\,\mathrm{d}\tilde{\mu}(\boldsymbol{t})\\&\qquad+\int_{\mathcal{X}^2}K(\boldsymbol{t},\boldsymbol{z})\,\mathrm{d}\tilde{\mu}(\boldsymbol{t})\,\mathrm{d}\tilde{\mu}(\boldsymbol{z})+\frac{1}{C}.\qquad(19)\end{aligned}$$

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

$$\begin{aligned}\tilde{E}(\mu)=\langle\mu,\mu\rangle_{\mathcal{M}(\tilde{K})}=\langle\mu-Q(\mu)\tilde{\mu},\mu-Q(\mu)\tilde{\mu}\rangle_{\mathcal{M}(K)}+\frac{Q(\mu)Q(\mu)}{C}\\=E(\mu-Q(\mu)\tilde{\mu})+\frac{Q(\mu)Q(\mu)}{C}>0,\qquad(20)\end{aligned}$$

since when $Q(\mu) = 0$ then first summand in (20) is positive since K is conditionally positive definite, while if $Q(\mu) \neq 0$ then the second summand is positive.

Moreover, by definition of $\tilde{K}$ in (19), the potential field induced by $\tilde{\mu}$ for the kernel $\tilde{K}$ is $|\tilde{\mu}|$ almost everywhere constant,

$$\int_{\mathcal{X}} \tilde{K}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\tilde{\mu}(\boldsymbol{y}) = \frac{1}{C} \quad |\tilde{\mu}| a.e., \quad \text{and} \quad \tilde{E}(\tilde{\mu}) = \frac{1}{C},$$

which means that $\tilde{\mu}$ is the equilibrium measure for $\tilde{K}$ with capacity $C_{\tilde{K}} = C$. Note also that energies defined by the two kernels are related as follows:

$$\tilde{E}(\mu - Q(\mu)\tilde{\mu}) = \tilde{E}(\mu) - [Q(\mu)]^2 \tilde{E}(\tilde{\mu}) = E(\mu - Q(\mu)\tilde{\mu}). \tag{21}$$

The first equality comes from Lemma 8 and the second equality comes from (20).

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 11 *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* (19) *is symmetric and positive definite kernel with capacity $C_{\tilde{K}} = C$ and equilibrium measure $\mu_{\mathrm{e},\tilde{K}} = \tilde{\mu}$.*

Corollary 12 *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 defined by* (19)*. 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*

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

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

PROOF. The second set of equalities is a special case of the first with $\nu = \mu_{\mathcal{P}}$. The equivalence of energies $E(\tilde{\mu}-\nu)$, $\tilde{E}(\tilde{\mu}-\nu)$, and $\tilde{E}(\nu)-\tilde{E}(\tilde{\mu})$ follows from (21). The equivalence of the energies E and $\tilde{E}$ to their corresponding discrepancies and numerical integration errors follows from Proposition 3 and Theorem 5. □

This corollary implies that changing the integration problem from one with integrands that are potential functions in $\mathcal{H}(K)$ to one with integrands that are potential functions in $\mathcal{H}(\tilde{K})$, does not change the quality measure of the design. In fact, the Hilbert spaces $\mathcal{H}(K)$ and $\mathcal{H}(\tilde{K})$ contain the same functions, and the two integration problems have exactly the same difficulty. According to (18) and (19), the inner products defined by the two different kernels are related as follows:

$$\langle f_\mu, f_\nu\rangle_{\mathcal{M}(\tilde{K})} = \left\langle f_\mu - f_{Q(\mu)\tilde{\mu}}, f_\nu - f_{Q(\nu)\tilde{\mu}}\right\rangle_{\mathcal{M}(K)},$$

where f_μ, f_ν, $f_{Q(\mu)\tilde{\mu}}$, and $f_{Q(\nu)\tilde{\mu}}$ are the potential fields induced by the charge distributions μ, ν, $Q(\mu)\tilde{\mu}$, and $Q(\nu)\tilde{\mu}$, respectively, under the energy kernel K.

4.3 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, measurable group actions 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

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

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 13 *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) $(g \cdot \mu)(\mathcal{Y}) := \mu(g^{-1}\mathcal{Y})$ *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 14 *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

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

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 15 *Let K be a $\mathcal{G}$-invariant kernel on $\mathcal{X}$ satisfying* (1)*. Then the $\mathcal{G}$ action on $\mathcal{M}(K)$ preserves total charge, 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 $\mathcal{M}(K)$.*

PROOF. Let $\mu \in \mathcal{M}(K)$ be any finite energy measure. Total charge is $\mathcal{G}$-invariant since $Q(g \cdot \mu) = \mu(g^{-1}\mathcal{X}) = \mu(\mathcal{X}) = Q(\mu)$. Also, since K is $\mathcal{G}$-invariant and the action of $\mathcal{G}$ on functions or kernels is dual to its action on

measures it follows that

$$E(\mu) = \int_{\mathcal{X}^2} K(x,y)\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{y}) = \int_{\mathcal{X}^2} (g^{-1}\cdot K)(x,y)\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{y})$$
$$= \int_{\mathcal{X}^2} K(x,y)\,\mathrm{d}(g\cdot\mu)(\boldsymbol{x})\,\mathrm{d}(g\cdot\mu)(\boldsymbol{y}) = E(g\cdot\mu). \tag{22}$$

So the action of $\mathcal{G}$ conserves energy. In particular, $g\cdot\mu \in \mathcal{M}(K)$ for each $\mu \in \mathcal{M}(K)$ and $g \in \mathcal{G}$. Inner products are also unchanged under group actions, i.e., $\langle g\cdot\mu, g\cdot\nu\rangle_{\mathcal{M}(K)} = \langle \mu,\nu\rangle_{\mathcal{M}(K)}$, since inner products can be defined in terms of energies by

$$\langle \mu,\nu\rangle_{\mathcal{M}(K)} = \frac{1}{4}E(\mu+\nu) - \frac{1}{4}E(\mu-\nu).$$

Now suppose that $\mu \in \mathcal{M}(K)$ has total charge one, i.e., $Q(\mu) = 1$. The measure μ_{AVG} defined for each $\mathcal{Y} \in \mathcal{B}$ by

$$\mu_{AVG}(\mathcal{Y}) = \int_{\mathcal{G}} g\cdot\mu(\mathcal{Y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)$$

is then a unit total charge measure which is $\mathcal{G}$-invariant. So by uniqueness of the $\mathcal{G}$-invariant measure $\mu_{AVG} = \mu_{\mathcal{G}}$. Since μ_{AVG} integrates a function f via $\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\mu_{AVG}(\boldsymbol{x}) = \int_{\mathcal{G}}\int_{\mathcal{X}} f(g\cdot\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\lambda_{\mathcal{G}}(g)$, the energy of μ_{AVG} can be estimated using the isometric action of $\mathcal{G}$ and the Cauchy-Bunyakovsky-Schwarz inequality as follows.

$$\begin{aligned}
E(\mu_{AVG}) &:= \int_{\mathcal{G}^2}\int_{\mathcal{X}^2} K(g\boldsymbol{x}, h\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)\,d\lambda_{\mathcal{G}}(h)\\
&= \int_{\mathcal{G}^2}\int_{\mathcal{X}^2} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}(g\cdot\mu)(\boldsymbol{x})\,\mathrm{d}(h\cdot\mu)(\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)\,\mathrm{d}\lambda_{\mathcal{G}}(h)\\
&= \int_{\mathcal{G}^2} \langle g\cdot\mu, h\cdot\mu\rangle_{\mathcal{M}(K)}\,\mathrm{d}\lambda_{\mathcal{G}}(g)\,\mathrm{d}\lambda_{\mathcal{G}}(h)\\
&\le \int_{\mathcal{G}^2} E^{1/2}(g\cdot\mu)E^{1/2}(h\cdot\mu)\,\mathrm{d}\lambda_{\mathcal{G}}(g)\,d\lambda_{\mathcal{G}}(h) = E(\mu).
\end{aligned}$$

Hence $E(\mu_{\mathcal{G}}) = E(\mu_{AVG}) \le E(\mu)$ which shows $\mu_{\mathcal{G}}$ is in $\mathcal{M}(K)$. It also shows $\mu_{\mathcal{G}}$ is the energy minimizer since μ is an arbitrary unit total charge measure in $\mathcal{M}(K)$. The potential field induced by $\mu_{\mathcal{G}} \in \mathcal{M}(K)$ evaluated at an arbitrary point $\boldsymbol{x} = g_{\boldsymbol{x}}\boldsymbol{\eta} \in \mathcal{X}$, may be written in terms of the fixed pole $\boldsymbol{\eta}$ as

$$\int_{\mathcal{X}} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu_{\mathcal{G}}(\boldsymbol{y}) = \int_{\mathcal{X}} K(g_{\boldsymbol{x}}\boldsymbol{\eta}, g_{\boldsymbol{x}}g_{\boldsymbol{x}}^{-1}\boldsymbol{y})\,\mathrm{d}\mu_{\mathcal{G}}(\boldsymbol{y}) = \int_{\mathcal{X}} K(\boldsymbol{\eta}, g_{\boldsymbol{x}}^{-1}\boldsymbol{y})\,\mathrm{d}\mu_{\mathcal{G}}(\boldsymbol{y}),$$

since K is $\mathcal{G}$-invariant. Letting $\boldsymbol{z} = g_{\boldsymbol{x}}^{-1}\boldsymbol{y}$, the right-hand integral above can be re-written as

$$\int_{\mathcal{X}} K(\boldsymbol{\eta},\boldsymbol{z})\,\mathrm{d}(g_{\boldsymbol{x}}^{-1}\cdot\mu_{\mathcal{G}})(\boldsymbol{z}) = \int_{\mathcal{X}} K(\boldsymbol{\eta},\boldsymbol{z})\,\mathrm{d}\mu_{\mathcal{G}}(\boldsymbol{z}),$$

since $\mu_{\mathcal{G}}$ is $\mathcal{G}$-invariant. This last integral is independent of $\boldsymbol{x}$. Thus the potential field is constant and $\mu_{\mathcal{G}}$ is the equilibrium measure. □

Theorem 16 *Let $\lambda_{\mathcal{G}}$ be the normalized Haar ($\mathcal{G}$-invariant) measure on $\mathcal{G}$. Then the kernel obtained by filtering K as follows,*

$$K_{\mathcal{G}}(\boldsymbol{x},\boldsymbol{y}) := \int_{\mathcal{G}} (g\cdot K)(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g) = \int_{\mathcal{G}} K(g^{-1}\boldsymbol{x}, g^{-1}\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g), \tag{23}$$

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, the 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.,

$$\int_{\mathcal{G}} E(g\cdot\nu)\,\mathrm{d}\lambda_{\mathcal{G}}(g) = E_{\mathcal{G}}(\mu), \tag{24a}$$

$$\int_{\mathcal{G}} D^2(g\cdot\nu;\mu_{\mathcal{G}})\,\mathrm{d}\lambda_{\mathcal{G}}(g) = D^2_{\mathcal{G}}(\nu;\mu_{\mathcal{G}}) = E_{\mathcal{G}}(\nu) - E_{\mathcal{G}}(\mu_{\mathcal{G}}). \tag{24b}$$

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

PROOF. Since both K and the action of $\mathcal{G}$ are measurable, the function K is bounded below by L_K, and $\lambda_{\mathcal{G}}$ is a positive measure, the filtered version of K is a well-defined kernel satisfying the conditions in (1). Moreover, $K_{\mathcal{G}}$ is $\mathcal{G}$-invariant because the averaging is done with respect to the $\mathcal{G}$-invariant measure, $\lambda_{\mathcal{G}}$. Specifically, for any $g_1 \in \mathcal{G}$, it follows by a change of variable, Fubini's theorem and the $\mathcal{G}$-invariance of the measure that

$$\begin{aligned}(g_1\cdot K_{\mathcal{G}})(\boldsymbol{x},\boldsymbol{y}) &= \int_{\mathcal{G}} (g_1\cdot(g\cdot K))(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)\\ &= \int_{\mathcal{G}} (g\cdot K)(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}({g_1}^{-1}\cdot\lambda_{\mathcal{G}})(g) = \int_{\mathcal{G}} (g\cdot K)(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g) = K_{\mathcal{G}}(\boldsymbol{x},\boldsymbol{y}).\end{aligned}$$

Thus, $K_{\mathcal{G}}$ is $\mathcal{G}$-invariant. Moreover, if K was $\mathcal{G}$-invariant to begin with, then $K(g\boldsymbol{x}, g\boldsymbol{y}) = K(\boldsymbol{x},\boldsymbol{y})$, and so the filtering step, (23), makes no change to K.

Now suppose $0 \neq \mu \in \mathcal{M}(K_{\mathcal{G}})$. Then Theorem 15 says $\mu_{\mathcal{G}} \in \mathcal{M}(K_{\mathcal{G}})$, and $\mu_{\mathrm{e},K_{\mathcal{G}}} = \mu_{\mathcal{G}} = \mu_{\min,K_{\mathcal{G}}}$.

Finally, consider the average energy of a measure $\nu \in \mathcal{M}(K_{\mathcal{G}})$ under the group action on the left hand side of (24a). Again using the duality between the $\mathcal{G}$ actions on measures and kernels, the measurability of K and Fubini's theorem

together yields (24a):

$$\begin{aligned}\int_{\mathcal{G}} E(g\cdot\nu)\,\mathrm{d}\lambda_{\mathcal{G}}(g) &= \int_{\mathcal{G}}\int_{\mathcal{X}^2} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}(g\cdot\nu)(\boldsymbol{x})\,\mathrm{d}(g\cdot\nu)(\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)\\ &= \int_{\mathcal{G}}\int_{\mathcal{X}^2} (g^{-1}\cdot K)(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\nu(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)\\ &= \int_{\mathcal{X}^2}\int_{\mathcal{G}} (g^{-1}\cdot K)(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\lambda_{\mathcal{G}}(g)\,\mathrm{d}\nu(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{y})\\ &= \int_{\mathcal{X}^2} K_{\mathcal{G}}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\nu(\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{y}) = E_{\mathcal{G}}(\nu).\end{aligned}$$

Equation (24b) follows from the definition of discrepancy (8) in terms of energy and the fact that the $\mathcal{G}$-invariance of $\mu_{\mathcal{G}} \in \mathcal{M}(K_{\mathcal{G}})$ means that $E(g\cdot\mu_{\mathcal{G}}) = E_{\mathcal{G}}(\mu_{\mathcal{G}})$ for all $g \in \mathcal{G}$. $\square$

Remark 17 *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 [12] which hold whenever $\mathcal{X}$ is the orbit of a group of* isometries *of the containing space.*

Some designs $\mathcal{P}$ utilize group actions in their construction, e.g.,

$$\mathcal{P} = \{g\boldsymbol{\eta} : g \in \tilde{\mathcal{G}}\}, \tag{25}$$

for some $\tilde{\mathcal{G}}$ that is an n element subgroup of $\mathcal{G}$. Rank-1 integration lattices [36] are an example of such designs. Digital nets [23,29] are another example. For designs of the form (25), the formulas for the energy $E_{\mathcal{G}}(\mathcal{P})$ and the discrepancy $D_{\mathcal{G}}(\mathcal{P};\mu_{\mathcal{G}})$ may be simplified as single sums rather than double sums, namely,

$$\begin{aligned}E_{\mathcal{G}}(\mathcal{P}) &= \frac{1}{n^2}\sum_{g,h\in\tilde{\mathcal{G}}} K_{\mathcal{G}}(g\boldsymbol{\eta},h\boldsymbol{\eta}) = \frac{1}{n^2}\sum_{g,h\in\tilde{\mathcal{G}}} K_{\mathcal{G}}(h^{-1}g\boldsymbol{\eta},\boldsymbol{\eta})\\ &= \frac{1}{n}\sum_{g\in\tilde{\mathcal{G}}} K_{\mathcal{G}}(g\boldsymbol{\eta},\boldsymbol{\eta}),\\ [D_{\mathcal{G}}(\mathcal{P};\mu_{\mathcal{G}})]^2 &= E_{\mathcal{G}}(\mathcal{P}) - E_{\mathcal{G}}(\mu_{\mathcal{G}})\\ &= \frac{1}{n}\sum_{g\in\tilde{\mathcal{G}}} K_{\mathcal{G}}(g\boldsymbol{\eta},\boldsymbol{\eta}) - \frac{1}{C_{K_{\mathcal{G}}}}.\end{aligned}$$

Contrast these formulae with those in (5) and (9), where double sums are needed to evaluate the energy and discrepancy.

5 Examples

To illustrate some of the results above we consider some simple examples. First consider the unit interval domain, $\mathcal{X} = [0, 1]$, and the Riesz-like kernel

$$K(x, y) = 1 - |x - y| . \tag{26}$$

The inner product for the space of charge distributions, $\mathcal{M}(K)$, defined by this kernel may be written, after some elementary algebra, as

$$\begin{aligned}\langle \mu, \nu \rangle_{\mathcal{M}(K)} &= \int_0^1 \int_0^1 [1 - |x - y|] \, \mathrm{d}\mu(x) \mathrm{d}\nu(y) \\ &= \frac{1}{2} Q(\mu) Q(\nu) + 2 \int_0^1 \left[\mu(x) - \frac{1}{2} Q(\mu) \right] \left[\nu(x) - \frac{1}{2} Q(\nu) \right] \mathrm{d}x.\end{aligned}$$

The energy this kernel is then given as in (7) as

$$E(\mu) = \|\mu\|^2_{\mathcal{M}(K)} = \frac{1}{2} [Q(\mu)]^2 + 2 \int_0^1 \left[\mu(x) - \frac{1}{2} Q(\mu) \right]^2 \mathrm{d}x,$$

and the square discrepancy by (8) is

$$\begin{aligned}D^2(\nu; \mu) &= \|\mu - \nu\|^2_{\mathcal{M}(K)} \\ &= \frac{1}{2} [Q(\mu - \nu)]^2 + 2 \int_0^1 \left[\mu(x) - \nu(x) - \frac{1}{2} Q(\mu - \nu) \right]^2 \mathrm{d}x. \end{aligned} \tag{27}$$

When μ and ν have the same total charge, then $D(\nu; \mu) = \sqrt{2} \, \|\mu - \nu\|_2$, which is essentially the L_2 star discrepancy.

The equilibrium measure for this kernel and domain is concentrated at the endpoints of the interval, i.e., $\mu_{\mathrm{e},K} = (\delta_0 + \delta_1)/2$, since

$$f_{\mu_{\mathrm{e},K}}(x) = \int_0^1 K(x, y) \, \mathrm{d}\mu_{\mathrm{e},K}(x) = \frac{K(0, y) + K(1, y)}{2} = \frac{1}{2},$$

and so the capacity is $C_K = 2$. For any charge distribution, ν, with unit total charge, one has

$$D^2(\nu; \mu_{\mathrm{e},K}) = E(\nu) - E(\mu_{\mathrm{e},K}) = 2 \int_0^1 [\nu(x) - 1/2]^2 \, \mathrm{d}x.$$

For a distribution with support on two points, y and z, with equal charge at these points, $\nu = (\delta_y + \delta_z)/2$, the square discrepancy and energy become

$$D^2(\nu; \mu_{\mathrm{e},K}) = E(\nu) - E(\mu_{\mathrm{e},K}) = \frac{1 - |y - z|}{2}.$$

This expression is minimized, and in fact vanishes, when the two points are moved to the ends of the interval and ν replicates the equilibrium distribution,

i.e., $D^2(\nu;\mu_{\mathrm{e},K}) = E(\nu) - E(\mu_{\mathrm{e},K}) = 0$ when $\nu = (\delta_0 + \delta_1)/2 = \mu_{\mathrm{e},K}$. If ν consists of a single unit point charge, $\nu = \delta_y$, then $D^2(\nu;\mu_{\mathrm{e},K}) = E(\nu) - E(\mu_{\mathrm{e},K}) = 1/2$, independent of the placement of the point charge. This means that the minimum energy point can be placed anywhere in the interval $[0,1]$. If ν consists of a single point charge of magnitude q, $\nu = q\delta_y$, then by (27) it follows that

$$D^2(\nu;\mu_{\mathrm{e},K}) = E(\mu_{\mathrm{e},K} - \nu) = \frac{(1-q)^2}{2} + 2\int_0^1 \left[\frac{q}{2} - q\delta_z(x)\right]^2 \mathrm{d}x = q^2 - q + \frac{1}{2}.$$

Again the placement of this point charge does not affect the value of the discrepancy, but the minimum discrepancy is obtained for $q = 1/2$, not a unit charge, and $D^2(\delta_z/2;\mu_{\mathrm{e},K}) = 1/4$.

The Hilbert space of potential fields, $\mathcal{H}(K)$, based on the kernel (26) is the space of absolutely continuous functions with square integrable first derivatives, and the inner product is

$$\langle f, g\rangle_{\mathcal{H}(K)} = \frac{1}{2}\left\{[f(0)+f(1)][g(0)+g(1)] + \int_0^1 f'(x)g'(x)\,\mathrm{d}x\right\}.$$

The integral with respect to the equilibrium measure, $\mu_{\mathrm{e},K} = (\delta_0 + \delta_1)/2$, corresponds to the average of the function at the two endpoints of the integral:

$$\int_0^1 f(x)\,\mathrm{d}\mu_{\mathrm{e},K}(x) = \frac{1}{2}[f(0)+f(1)],$$

and the representer of the integration functional is $f_{\mu_{\mathrm{e},K}}(x) = 1/2$. By Theorem 5, the worst-case error for approximating this integral by $\int_0^1 f(x)\,\mathrm{d}\nu(x)$ is the discrepancy, $D(\nu;\mu_{\mathrm{e},K})$.

Based on the observations above, approximating $[f(0)+f(1)]/2$ by $[f(y)+f(z)]/2$ has a worst-case error of $1 - |y-z|/2$, which vanishes for $y = 0$ and $z = 1$. The worst-case potential field is (proportional to) the piecewise linear function

$$f_{\mu_{\mathrm{e},K}-\nu}(x) = \frac{|x-y| + |x-z|}{2} = \begin{cases} (y+z)/2 - x, & 0 \le x \le \min(y,z), \\ |y-z|/2, & \min(y,z) < x \le \max(y,z), \\ x - (y+z)/2, & \max(y,z) < x \le 1. \end{cases}$$

Approximating $[f(0)+f(1)]/2$ by the single weighted function value $qf(y)$ has a worst-case error of $q^2 - q + 1/2$, independent of y, and the worst-case potential field is the v-shaped

$$f_{\mu_{\mathrm{e},K}-\nu}(x) = \frac{1 - 2q + 2q|x-y|}{2}.$$

The worst-case error attains a minimum value of $1/4$ for $q = 1/2$.

Integration with respect to $\mu_{\mathrm{e},K} = (\delta_0 + \delta_1)/2$ is a rather simple problem, since it may be done exactly with just two well-chosen function evaluations. A more typical problem is integration with respect to a continuous distribution, such as the uniform distribution, $\tilde{\mu}(x) = x$. Modifying the Riesz kernel in Example 1 using Proposition 11 to change the equilibrium measure to the uniform measure on $[0, 1]$ yields the kernel

$$\tilde{K}(x,y) = \operatorname{sign}(s)\left[|x-y|^{-s} - \frac{x^{1-s} + (1-x)^{1-s} + y^{1-s} + (1-y)^{1-s}}{1-s} + \max\left(\frac{2}{(1-s)(2-s)}, 0\right)\right] + \frac{1}{\tilde{C}},$$

where $\tilde{C}$ is an arbitrary positive constant. For $s = 0, 1$ the above formula must be modified with certain powers replaced by logarithms. For $s \geq 1$ Proposition 11 cannot be applied directly since the integral of the Riesz kernel with respect to the uniform measure is infinite. Thus, we consider the kernel $K(x,y) = \operatorname{sign}(s)(|x-y| + \epsilon)^{-s}$, and take limits as $\epsilon \downarrow 0$. Letting $\tilde{E}$ denote the energy defined by the kernel $\tilde{K}$, it follows that $\tilde{E}(\tilde{\mu}) = 1/C_{\tilde{K}}$ for $s < 1$.

The discrepancy, $\tilde{D}$ defined by $\tilde{K}$ and the discrepancy D defined by K, are the same, so for $s = -1$ and any measure ν,

$$\tilde{D}^2(\nu;\tilde{\mu}) = D^2(\nu;\tilde{\mu}) = \frac{1}{2}[1 - Q(\nu)]^2 + 2\int_0^1 \left[x - \frac{1}{2} - \nu(x) + \frac{1}{2}Q(\nu)\right]^2 \mathrm{d}x.$$

When ν has unit total charge, this simplifies to

$$\begin{aligned} D^2(\nu;\tilde{\mu}) = \tilde{D}^2(\nu;\tilde{\mu}) &= \tilde{E}(\nu) - \tilde{E}(\tilde{\mu}) \\ &= \sup_{\|f\|_{\mathcal{H}(\tilde{K})} \leq 1} \left| \int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x) - \int_{\mathcal{X}} f(x)\,\mathrm{d}\nu(x) \right|^2 \\ = \sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x) - \int_{\mathcal{X}} f(x)\,\mathrm{d}\nu(x) \right|^2 &= 2\int_0^1 [x - \nu(x)]^2 \,\mathrm{d}x. \end{aligned}$$

When $\nu = \mu_{\mathcal{P},\boldsymbol{q}} = \sum_{i=1}^n q_i \delta_{z_i}$, the sum of n point charges of possibly differing magnitudes but total charge one, then

$$\begin{aligned} D^2(\mu_{\mathcal{P},\boldsymbol{q}};\tilde{\mu}) = \tilde{D}^2(\mu_{\mathcal{P},\boldsymbol{q}};\tilde{\mu}) &= \tilde{E}(\mu_{\mathcal{P},\boldsymbol{q}}) - \tilde{E}(\tilde{\mu}) \\ &= \sup_{\|f\|_{\mathcal{H}(K)} \leq 1} \left| \int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x) - \sum_{i=1}^n q_i f(z_i) \right|^2 \\ &= 2\left\{ \frac{1}{3} + \sum_{i=1}^n q_i z_i^2 - \sum_{i,j=1}^n q_i q_j \max(z_i, z_j) \right\}. \end{aligned}$$

and when $\nu = \mu_{\mathcal{P}} = n^{-1}\sum_{i=1}^n \delta_{z_i}$, the empirical distribution of the design $\mathcal{P}$,

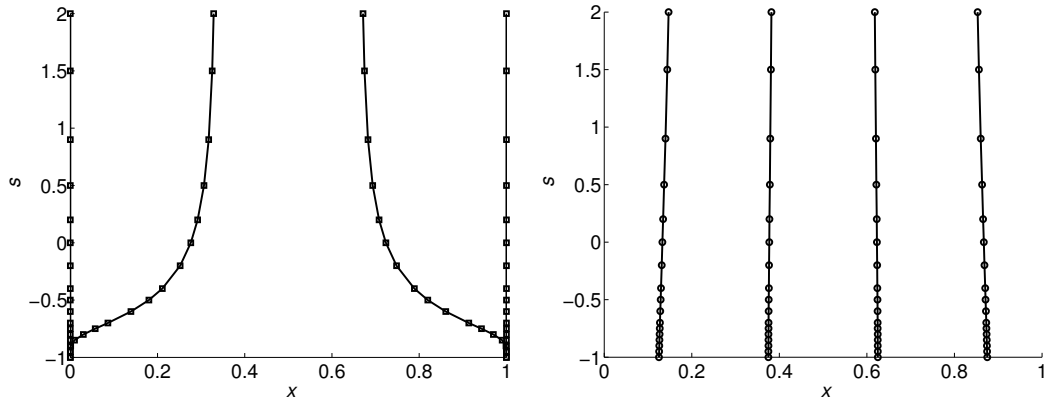

Fig. 1. Configurations of four points, $\mathcal{P}$, with minimizing the energy $E_{\text{phys}}(\mathcal{P})$ on the left, and $\tilde{E}_{\text{phys}}(\mathcal{P})$ on the right for the Riesz kernel, (2), as a function of the parameter s.

then

$$\begin{aligned} D^2(\mu_{\mathcal{P}};\tilde{\mu}) = \tilde{D}^2(\mu_{\mathcal{P}};\tilde{\mu}) &= \tilde{E}(\mu_{\mathcal{P}}) - \tilde{E}(\tilde{\mu}) \\ &= \sup_{\|f\|_{\mathcal{H}(K)}\leq 1} \left| \int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x) - \frac{1}{n}\sum_{i=1}^{n} f(z_i) \right|^2 \\ &= 2\left\{ \frac{1}{3} + \frac{1}{n}\sum_{i=1}^{n} z_i^2 - \frac{1}{n^2}\sum_{i=1}^{n} \max(z_i, z_j) \right\}. \end{aligned}$$

Figure 1 displays the minimum energy points, $E_{\text{phys}}(\mathcal{P})$ and $\tilde{E}_{\text{phys}}(\mathcal{P})$, defined above in (3) for the Riesz kernel and the modified Riesz kernel in dimension one. Because $E_{\text{phys}}(\mathcal{P})$ ignores the diagonal terms $K(z_i, z_i)$, these energies are finite even for $s \geq 0$. For $s < 0$ the points minimizing $E_{\text{phys}}(\mathcal{P})$ and $E(\mu_{\mathcal{P}})$ are the same since the diagonal terms are constant. The point configurations on the right also minimize the discrepancy $D(\mu_{\mathcal{P}};\tilde{\mu}) = \tilde{D}(\mu_{\mathcal{P}};\tilde{\mu})$ for $s < 0$.

The minimum discrepancy points for $s = -1$ are the evenly spaced points $\mathcal{P} = \{1/8, 3/8, 5/8, 7/8\}$, and as s increases these points move towards the center. This can be understood by looking at the modified Riesz kernel $\tilde{K}$, and noting that its definition adds the external fields placed at 0 and 1, which act to repel the point charges away from the boundaries. In contrast, we see that the original Riesz kernel pushes the charges towards the boundaries. For $s = -1$, we have that the points end up at each end of the interval, while for other values of s, at least one point charge is at each endpoint.

If one considers the group of measure preserving bijections $\mathcal{G} = \{g_\Delta x := x + \Delta \mod 1 : 0 \leq \Delta < 1\}$, then the filtered Riesz kernel defined in (23) is

$$K_{\mathcal{G}}(x, y) = \text{sign}(s)\left[|x-y|^{-s}\left(1 - |x-y|\right) + |x-y|\left(1 - |x-y|\right)^{-s} \right].$$

For this kernel, the set of evenly spaced points with shift Δ, $0 \leq \Delta \leq 1$, i.e., $\mathcal{P} = \{\Delta, 1/n + \Delta, \ldots, (n-1)/n + \Delta\}$, has an energy and discrepancy independent of the Δ. For $s = -1$, it is known to be the minimum energy point set.

6 Discussion and Conclusion

We conclude by highlighting some further open questions arising from our investigations here.

i. Many symmetric positive definite kernels possess a *maximum principle*. This means that the potential field induced by the equilibrium measure, $\mu_{\mathrm{e},K,\mathcal{Y}}$, attains its maximum on its support, $\mathcal{Y}$. However, this maximum principle does not hold in general, even under what might seem to be some reasonable additional conditions. Returning to Example 2, let $\mathcal{X} = \{1, 2, 3\}$ and write the kernel, K, in the form of a 3×3 symmetric, positive definite matrix:

$$\mathsf{K} = \begin{pmatrix} 9 & 1 & 6 \\ 1 & 9 & 6 \\ 6 & 6 & 14 \end{pmatrix} = \mathsf{V} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 22 \end{pmatrix} \mathsf{V}^T, \qquad \mathsf{V} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix},$$

which is even diagonally dominant. For $\mathcal{Y} = \{1, 2\}$ the equilibrium potential is $(0.5, 0.5)^T$, and the potential field is $(5, 5, 6)^T$ which, on $\mathcal{Y}$ is the constant 5, but attains its maximum value of 6 outside $\mathcal{Y}$. It would be nice to know under what general conditions a kernel possesses a maximum principle.

ii. For some kernels $K_{\mathcal{G}}$ it is known that minimum energy or discrepancy points are of the form in (25), but it is not known in general.

iii. In Theorem 5 and Corollary 10 the energy based on a kernel is identified as the tight numerical integration error bound for a Hilbert space of integrands defined by that same kernel. For some energy kernels with a simple form, such as the Riesz kernel, the corresponding Hilbert space of integrands cannot be simply described. On the other hand, for some spaces of integrands with a simple description, the corresponding kernel may not have a simple expression. A greater understanding is needed of how well minimum energy or discrepancy points defined using one kernel are good for use in approximating an integral for a space of integrands defined by a similar kernel.

iv. There exist other distance-based criteria for constructing designs, such as minimizing the covering radius or maximizing the separation distance. For example, the design minimizing the energy $\tilde{E}_{\mathrm{phys}}$ in Figure 1 for

$s = -1$ is a minimum covering radius design, and the design minimizing (fill distance) the energy E_{phys} in Figure 1 for $s \uparrow \infty$ is a maximum separation distance (also called a sphere-packing) design. For some situations these correspond to minimum energy or discrepancy designs, but in other situations they are different. It would be helpful to understand how well a design that optimizes one criterion measures up under another criterion.

Acknowledgements

The authors thank three referees for their constructive and useful comments which helped us improve this paper. S. B. Damelin was supported, in part, by EP/C000285, NSF-DMS-0439734 and NSF-DMS-0555839. F. J. Hickernell and X. Zeng were supported in part by NSF-DMS-0713848, the Illinois Institute of Technology, Hong Kong Baptist University grant FRG/03-04/II-49, and Hong Kong Research Council Grant HKBU/2009/04P. Part of this work was completed during visits by S. B. Damelin to the Illinois Institute of Technology and by F. J. Hickernell and, separately, D. L. Ragozin to Georgia Southern University. The authors thank these institutions for their financial support for these visits.

References

[1] Andrievskii, V.V., Blatt, H.P., Discrepancy of Signed Measures and Polynomial Approximation. Springer Monographs in Mathematics. Springer, New York (2002)

[2] Andrievskii, V.V., Blatt, H.P., Götz, M.: Discrepancy estimates on the sphere. Monatsh. Math. **128**, 179–188 (1999)

[3] Benko, D., Damelin, S.B., Dragnev, P.D.: On the support of the equilibrium measure for arcs of the unit circle and real intervals. Electron. Trans. Numer. Anal. **25**, 27–40 (2006)

[4] Blatt, H.P., Mhaskar, H.N.: A general discrepancy theorem. Ark. Mat. **31**, 219–246 (1993)

[5] Cranley, R., Patterson, T.N.L.: Randomization of number theoretic methods for multiple integration. SIAM J. Numer. Anal. **13**, 904–914 (1976)

[6] D'Agostino, R.B., Stephens, M.A. (eds.): Goodness-of-Fit Techniques. Dekker, New York (1986)

[7] Damelin, S.B.: Marcinkiewicz-Zygmund inequalities and the numerical approximation of singular integrals for exponential weights: methods,results and open problems, some new, some old. J. Complexity **19**, 406–415 (2003)

[8] Damelin, S.B., Grabner, P.: Numerical integration, energy and asymptotic equidistributon on the sphere. J. Complexity **19**, 231–246 (2003).

[9] Damelin, S.B., Kuijlaars, A.B.J.: The support of the equilibrium measure for monomial external fields on $[-1, 1]$. Trans. Amer. Math. Soc. **351**, 4561–4584 (1999)

[10] Damelin, S.B., Maymeskul, V.: On point energies, separation radius and mesh norm for s-extremal configurations on compact sets in $\mathbb{R}^n$. J. Complexity **21**, 845–863 (2005)

[11] Damelin, S.B., Dragnev, P.D., Kuijlaars, A.B.J.: The support of the equilibrium measure for a class of external fields on a finite interval. Pac. J. Math.**199**, 303–321 (2001)

[12] Damelin, S.B., Levesley, J., Ragozin, D.L., Sun, X.: Energies, group invariant kernels and numerical integration on compact manifolds. J. Complexity **25**, 152–162 (2005)

[13] Deift, P.: Orthogonal polynomials and random matrices: a Riemann–Hilbert approach. Courant Lecture Notes in Mathematics, Courant Institute, New York (1999).

[14] Du, Q., Faber, V., Gunzburger, M.: Centroidal Voronoi tessellations: applications and algorithms. SIAM Rev. **41**(4), 637–676 (1999)

[15] Fasshauer, G.E.: Meshfree Approximation Methods with MATLAB. Interdisciplinary Mathematical Sciences, vol. 6. World Scientific Publishing Co., Singapore (2007)

[16] Grabner, P.J., Tichy, R.F.: Spherical designs, discrepancy and numerical integration. Math. Comp. **60**, 327–336 (1993)

[17] Hardin, D.P., Saff, E.B.: Discretizing manifolds via minimum energy points. Notices Amer. Math. Soc. **51**, 1186–1194 (2004)

[18] Hickernell, F.J.: A generalized discrepancy and quadrature error bound. Math. Comp. **67**, 299–322 (1998)

[19] Hickernell, F.J.: Goodness-of-fit statistics, discrepancies and robust designs. Statist. Probab. Lett. **44**, 73–78 (1999)

[20] Hickernell, F.J.: What affects the accuracy of quasi-Monte Carlo quadrature?, In: Niederreiter, H., Spanier, J. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 1998, pp. 16–55. Springer, New York (2000)

[21] Hickernell, F.J., Sloan, I.H., Wasilkowski, G.W.: On tractability of weighted integration over bounded and unbounded regions in $\mathbb{R}^s$. Math. Comp. **73**, 1885–1901 (2004)

[22] Landkof, N.S.: Foundations of Modern Potential Theory. Springer-Verlag, Berlin (1972)

[23] Larcher, G.: Digital point sets: Analysis and applications. In: Hellekalek, P., Larcher, G. (eds.), Random and Quasi-Random Point Sets. Lecture Notes in Statistics, vol. 138, pp. 167–222. Springer, New York (1998)

[24] Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on the sphere I. Comm. Pure Appl. Math. **39**, 148–186 (1986).

[25] Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on the sphere II. Comm. Pure Appl. Math. **40**, 401–420 (1987)

[26] Mhaskar, H.N.: Weighted polynomials, radial basis functions, and potentials on locally compact spaces. Numer. Funct. Anal. Optim. **11**, 987–1017 (1990/1991)

[27] Mhaskar, H.N.: Introduction to the Theory of Weighted Polynomial Approximation. World Scientific Publishing Co., Singapore (1996)

[28] Mhaskar, H.N.: On the tractability of multivariate integration and approximation by neural networks. J. Complexity **20**, 561–590 (2004)

[29] Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia (1992)

[30] Novak, E., Woźniakowski, H.: When are integration and discrepancy tractable? In: Foundations of Computational Mathematics. London Math. Soc. Lecture Note Ser., vol. 284. Cambridge University Press, Cambridge (2001)

[31] Owen, A.B.: Monte Carlo, quasi-Monte Carlo, and randomized quasi-Monte Carlo. In: Niederreiter, H., Spanier, J. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 1998. Springer, New York (2000)

[32] Saff, E.B., Totik, V.: Logarithmic Potentials with External Fields. Springer, New York (1997)

[33] Santner, T.J., Williams, B.J., Notz, W.I.: The Design & Analysis of Computer Experiments. Springer, New York (2003)

[34] SAS Institute, JMP 7.0 (2007).

[35] Sjögren, P.: Estimates of mass distributions from their potentials and energies. Ark. Mat. **10**, 5977–5987 (1972)

[36] Sloan, I.H., Joe, S.: Lattice Methods for Multiple Integration. Oxford University Press, Oxford (1994)

[37] Stahl, H., Totik, V.: General Orthogonal Polynomials. Cambridge University Press, Cambridge (1992)

[38] Totik, V.: Weighted Approximation with Varying Weight. Lecture Notes in Mathematics, vol. 1569. Springer, Berlin (1994)

[39] Wendland, H.: Scattered Data Approximation. Cambridge Monographs on Applied and Computational Mathematics, vol. 17. Cambridge University Press, Cambridge (2005)

[40] Weyl, H.: Über die Gleichverteilung der Zahlen mod Eins. Math. Ann. **77**, 313–352 (1916)