On the Equilibrium Distributions of Leja Sequences on Curves

Mateo Attanasio, Praneeth Kolichala, Chandler Watson
Advisor: Rodrigo Angelo

August 2020

Abstract

A classic problem in physics asks what positions of N electrons on the circle constitute equilibria, ai.e. positions of minimal potential energy. In this physical case, the potential energy is (proportional to) a sum of inverse squares of distances. We consider a similar problem with potential as the inverse of the product of distances, i.e. what positions of N points on the circle minimize

$$E = -\sum_{i \neq j} \log |z_i - z_j|.$$

Such sequences are known as *Leja sequences*. It is known that an equilibrium is achieved exactly when the N points lie at the vertices of a regular N-gon, and as such the density of points approaches the uniform distribution as $N \to \infty$. The problem becomes more interesting when one greedily chooses points to minimize the energy, fixing all previous points. Such sequences are known as *Leja sequences*. We generalize this notion to closed bounded sets $K \subset \mathbb{C}$, and demonstrate analogous results to existing literature, exhibiting an application to Leja sequences on the interval.

1 Introduction

Consider a sequence of points $(x_n)_{n=1}^{\infty}$ in $[0, 1]$. We say such a sequence is *equidistributed* if for any $[a, b] \subset [0, 1]$ the limit as $N \to \infty$ of the fraction of the first N points lying in $[a, b]$ is $b - a$, i.e. each interval has its proportional share of points. An example equidistributed sequence is $x_n = \{n\sqrt{2}\}$, where $\{\cdot\}$ denotes taking the fractional part.

Quantitatively, to see how quickly a sequence becomes "well-distributed," we might consider the *discrepancy* of the first N points:

$$D_N = \sup_{[a,b] \subset [0,1]} \left| \frac{\#\{1 \leq n \leq N : x_n \in [a, b]\}}{N} - (b - a) \right|,$$

Intuitively, this is how badly $(x_n)_{n=1}^{\infty}$ fails at being equidistributed after N points. The *star discrepancy* D_N^* is defined in the same way with $a = 0$ fixed for the supremum. The bound $D_N^* \leq D_N \leq 2D_N^*$ tells us that their behaviours are always of the same order of magnitude. There is particular interest in low-discrepancy sequences, as they are well-distributed not only in the limit, but also well-distributed even if you stop "placing points" at any point in time.

We note that having the discrepancy going to zero precisely tells us that

$$\frac{1}{N}\sum_{i=1}^{N}\chi_{[a,b]}(x_i) \to \int_0^1 \chi_{[a,b]}(x)dx = (b-a)$$

In fact, since we may approximate Riemann-integrable functions by sums of characteristic functions, this is equivalent to

$$\frac{1}{N}\sum_{i=1}^{N} f(x_i) \to \int_0^1 f(x)dx$$

Weyl's famous equidistribution theorem tells us that it suffices to check this condition on the functions $f_n(x) = e^{2\pi i n x}$, since these form a basis for $L^2(\mathbb{T})$, where $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ is the circle.

It is a result of Schmidt in [3] that for any sequence $(x_n)_{n=1}^{\infty}$ in $[0,1]$, there is a fixed C such that for infinitely many N,

$$D_N^* \geq C\frac{\log N}{N},$$

implying the best we can hope for is a sequence with $D_N^* = O(\log N/N)$. Two optimal constructions are known which achieve this. The first is the Kronecker sequence, which is defined as $x_n = \{\gamma n\}$, where γ is any irrational. The second is the van der Corput sequence, which has x_n defined in binary as $0.a_1a_2\cdots a_n$ where $a_1a_2a_3\cdots$ is the reversed binary expansion of n. Such sequences are well-studied; see e.g. [4] for a complete treatment.

We can generalize discrepancy and star discrepancy to $[0,1]^d$ by considering a supremum over the fraction of points lying in rectangles $[a_1,b_1]\times\cdots\times[a_n,b_n]$ minus their areas. Remarkably, for $d \geq 2$, *no tight bound is known for* D_N^*. Roth, in [1], showed that for any C,

$$D_N^* \geq C\,\frac{(\log N)^{d/2}}{N}$$

for infinitely many N, but current conjectures from empirical data expect much tighter bounds, i.e. d instead of $d/2$ in the exponent (see [6]).

Until recently, the only low-discrepancy sequences known for $d \geq 2$ were simple generalizations of Kronecker and Van der Corput sequences, and it is unclear whether these are optimal. This suggests exploring new constructions is of use.

A recent construction is this: in [6], Steinerberger demonstrated that for a special class of functions $f : [0,1] \to \mathbb{R}$, we can take an "energy" on n points $x_1, \cdots, x_n$ as

$$E = \sum_{\substack{i,j=1 \\ i \neq j}} f(x_i - x_j).$$

Steinerberger showed that starting with a single point x_1 and iteratively greedily choosing x_n to minimize E tends to yield fairly regular sequences.

In particular, one intuitive f to consider is this: sending points in $[0,1]$ to their images on the circle $\mathbb{T}$, we might consider the distance between points when $\mathbb{T}$ is identified with the unit circle. Then, taking f as measuring the negative log distance between two such points, i.e.

$$f = -\log\left(2 \sin \pi |x|\right),$$

E then becomes the negative log of product of distances. This construction satisfies favorable bounds on D_N^* for all d.

Recently, in [8] Steinerberger rephrased his energy-based construction for $d = 1$ in terms of *Leja sequences*: for some choice of points $z_1, \cdots, z_m$ on the complex unit circle, take $p_m(z) = \prod_{k=1}^{m}(z - z_k)$, and iteratively define

$$p_{N+1}(z) = p_N(z)(z - z^*) \qquad \text{where } z^* = \arg\max_{|z|=1} |p_N(z)|.$$

The choices of z^* form the desired sequence: by taking their argument and scaling to $[0,1]$, we recover the original greedy energy-based sequence.

In this paper, we consider a new notion of *generalized Leja sequences* on connected closed subsets $K \subset \mathbb{C}$. In particular, we consider the sequence formed by placing a single point in K and then greedily placing additional points in K to minimize the product of pairwise distances between points. The following section provides some details about generalized Leja sequences, and introduces our approach via optimal distributions. In the third section, we demonstrate some results regarding special cases (namely, K as the closed disk and K as the interval) and provide a framework for generalizing to arbitrary K, and lastly pose an open question whose resolution would allow for a much deeper understanding of the topic. Finally, in the appendix, we give empirical results for the case of the interval.

2 Preliminaries

2.1 Generalized Leja sequences

We first begin with the definition of our object of study:

Definition 2.1. *A generalized Leja sequence $(z_n)_{i=1}^{\infty}$ on a closed connected set $K \subset \mathbb{C}$ is the sequence iteratively constructed by choosing $z_1 \in \partial K$ and then* *afterward greedily choosing z_{n+1} to maximize the product of distances between points, i.e. we take*

$$z_{n+1} = \arg\max_{z \in \partial K} \prod_{i=1}^{n} |z - z_i|.$$

One important detail central to our exploration is this:

Proposition 2.2. *A generalized Leja sequence on the closed connected set $K \subset \mathbb{C}$ lies entirely within ∂K.*

Proof. By induction. $z_1 \in \partial K$ by definition. If we have $z_1, \cdots, z_n \in \partial K$, then since $\prod_{i=1}^{n} z - z_i$ is holomorphic in z, the maximum modulus principle implies

$$\left| \prod_{i=1}^{n} z - z_i \right| = \prod_{i=1}^{n} |z - z_i|$$

is maximized in $\partial K \subset K$, meaning $z_{n+1} \in \partial K$. $\square$

In particular, if we take $K \subset \mathbb{C}$ homeomorphic to the disc, we merely only need understand generalized Leja sequences on choices of ∂K homeomorphic to the circle.

2.2 Relaxing on the circle

One observation we make is that in general, the first n points in a generalized Leja sequence will *not* minimize the product of pairwise distances. Indeed, as one might expect,

Proposition 2.3. *For $z_1, \cdots, z_n$ with $|z_k| = 1$*

$$\prod_{i,j=1}^{N} |z_i - z_j|$$

is minimized when a regular n-gon.

Proof. Recall that the square Vandermonde matrix

$$V = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{bmatrix}$$

has determinant

$$\det(V) = \prod_{1 \le j < k \le n} (\alpha_k - \alpha_j).$$

This can be seen e.g. by row reduction of V. Moreover, Hadamard's inequality states that for a complex matrix N with elements $|N_{jk}| \le B$

$$|\det(N)| \le B^n n^{n/2}.$$

Taking these together, we note that when $|\alpha_k| = 1$, we have

$$\prod_{1 \le j < k \le n} |\alpha_j - \alpha_k| \le n^{n/2}$$

and moreover this upper bound is achieved when $\alpha_k = e^{2\pi i k/n}$, as these are precisely the roots of unity, i.e. the roots of $x^n - 1 = (x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n)$.

Thus, the product of the distances from each root to a fixed root α_k is

$$\left| \frac{x^n - 1}{x - \alpha_k} \right|$$

evaluated at $x = \alpha_k$. By symmetry, we may take $\alpha_k = 1$, since distance is preserved under rotation, to get

$$\left| \prod_{1 \le k < n} (1 - e^{2\pi i k/n}) \right| = \lim_{x \to 1} \left| \frac{x^n - 1}{x - 1} \right| = 1 + x + x^2 + \dots + x^{n-1} = n$$

Thus, the product is

$$\prod_{1 \le j, k \le n} |\alpha_j - \alpha_k| = \left| \prod_{1 \le k \le n} (1 - e^{2\pi i k/n}) \right|^n = n^n$$

If we count each pair once instead of twice, we get that the product is $n^{n/2}$, so a regular polygon does indeed achieve this maximum. □

It is clear that because Leja sequences are greedily constructed, there is no way to have them form a n-gon at each step. Note, however, that even though Leja sequences aren't optimal in this sense, it is useful to consider the above "non-greedy" problem–indeed, for any n, the first 2^n points from a Leja sequence from a single point will lie on a 2^n-gon.

2.3 Optimal distributions

One major feature of this is that as $N \to \infty$, the points spread evenly about the circle: in fact, they approach a uniform distribution.

We can make one more relaxation to our problem which will be useful: considering the problem in the continuum limit. Note that maximizing the product of distances is the same as minimizing the negative log sum of distances, so it is natural to consider

$$E_{\text{cont}} = -\int_{|z|=1} \int_{|z'|=1} f(z) f(z') \log|z - z'| \, dz \, dz'.$$

Indeed, the above expression weights pairs of points according to their density, and replacing f with an indicator function on a discrete set, we recover the previous problem. We seek a L^2 function $f(z)$ with unit integral for $|z| = 1$ such that E_{cont} is minimized. First, we recall a crucial definition:

Definition 2.4. *A positive-definite function $h : \mathbb{T} \to \mathbb{R}$ is a L^2 function for which the k^{th} Fourier coefficient*

$$\hat{h}(k) = \int_0^{2\pi} h(\theta) e^{-2\pi i k\theta} \, d\theta$$

is non-negative for all $k \in \mathbb{Z}$.

As a side note, most literature defines this property differently (in terms of a matrix of h applied to differences of points, hence the name), but our definition is known to be equivalent. Additionally,

Lemma 2.5. *If a function h is positive-definite, the quadratic form*

$$\int_0^1 \int_0^1 g(\theta) g(\theta') h(\theta - \theta') \, d\theta \, d\theta'$$

for an L^2 function g with unit integral is minimized by g constant.

Proof. We note

$$\begin{aligned}
&\int_0^1 \int_0^1 g(x) g(x') \left(\sum_{k \in \mathbb{Z}} \hat{h}(k) e^{2\pi i k(x - x')} \right) dx \, dx' \\
&= \sum_{k \in \mathbb{Z}} \hat{h}(k) \int_0^1 \int_0^1 g(x) g(x') e^{2\pi i k(x - x')} \, dx \, dx' \\
&= \sum_{k \in \mathbb{Z}} \hat{h}(k) \left| \int_0^1 g(x) e^{2\pi i k x} \, dx \right|^2 \\
&= \sum_{k \in \mathbb{Z}} \hat{h}(k) \left| \hat{g}(-k) \right|^2 \\
&\geq \hat{h}(0)
\end{aligned}$$

where we can swap the sum and integral in the first step due to Fubini's theorem applied on a counting measure, given that the expression converges. The inequality is tight iff g is constant. ☐

Lastly, we replicate an argument from Steinerberger in [7]. First, C.-S. Lin in [5] gives the following result:

Proposition 2.6 (C.-S. Lin)**.** *For any $x \notin \pi\mathbb{Z}$,*

$$\log |\sin x| = -\log 2 - \sum_{k=1}^{\infty} \frac{\cos(2kx)}{k}.$$

Using this, Steinerberger thus observes that

$$-\log |2 \sin x| = \sum_{k=1}^{\infty} \frac{\cos(2kx)}{k} = \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{1}{2k} e^{2\pi i k x}$$

meaning all Fourier coefficients of $h(x) = -\log|2\sin(\pi x)|$ are non-negative and thus $h(x)$ is positive-definite.

Proposition 2.7. *Over all f with unit integral, f constant minimizes E_{cont}.*

Proof. We parametrize our integrals with $z = e^{2\pi i x}$ and $z' = e^{2\pi i x'}$, which yields

$$\begin{aligned}
|z - z'| &= |e^{2\pi i x} - e^{2\pi i x'}| \\
&= |1 - e^{2\pi i (x - x')}| \\
&= \sqrt{(1 - \cos(2\pi(x - x'))^2 + (\sin(2\pi(x - x')))^2} \\
&= \sqrt{2 - 2\cos(2\pi(x - x'))} \\
&= 2|\sin(\pi(x - x'))|
\end{aligned}$$

and thus

$$E_{\text{cont}} = -\int_0^1 \int_0^1 g(x)g(x') \log|2\sin(\pi(x - x'))| \, dx \, dx'.$$

where $g(x) = f(e^{2\pi i x})$. Since $h(x) = -\log|2\sin(\pi x)|$ is positive-definite, the result holds. $\square$

Thus, we have relaxed Leja sequences on the circle to non-greedy choices of n points. In general, since the points of a generalized Leja sequence always lie on the boundary of a given K, we can nicely define

Definition 2.8. *An optimal distribution on K is an L^2 function $g : \partial K \to \mathbb{R}_{\geq 0}$ which minimizes*

$$E_{K,cont} = -\int_{\partial K} \int_{\partial K} g(x)g(x') \log|x - x'| \, dx \, dx'.$$

The above shows that the lone optimal distribution for the disk is $f = 1$. While this problem has been significantly relaxed from the original question of generalized Leja sequences, optimal distributions appear to strongly inform the behavior of generalized Leja distributions, and as we will see transform nicely under conformal equivalences.

2.4 Conformal equivalences

Recall that a conformal equivalence is a bijection which preserves angles between tangent vectors of curves: in particular, conformal equivalences in $\mathbb{C}$ are biholomorphisms. We motivate the idea that conformal equivalences can help us determine optimal distributions on various choices of K.

In [2], Goluzin describes the concept of the *transfinite diameter* of a closed bounded infinite set K: effectively the limit of a function of the product of distances (as in section 2.2) between a minimizing set of points in K.

He then considers the limit γ of the constant term in the Green's function of a set of domains approaching the component of the complement of K containing ∞, and denotes $e^{-\gamma}$ the *capacity* of K. After, he demonstrates that the transfinite diameter and capacity are equal.

Next, when K is connected, he demonstrates that the Green's function on K is such that the capacity equals the *conformal radius*, defined as the radius R for which the complement of K maps conformally via some F onto $|z| > R$, fixing infinity with $F'(\infty) = 1$.

Definitionally, the conformal radius is thus invariant under conformal equivalences on the exterior of K to the exterior of another closed bounded simply connected K', implying via the previous that K and K' share a transfinite diameter.

The transfinite diameter encodes information regarding the limiting behavior on the product of distances between points as we add arbitrarially many points, and thus it would seem that conformal equivalences on the exteriors of sets K as chosen preserve information about equilbria. In particular, from empirical data as well as this motivation, we believe that conformal equivalences on the exteriors of such K preserve optimal distributions:

Question 2.9. *Does a conformal equivalence between the exteriors of two simply connected closed bounded sets $K, K' \subset \mathbb{C}$ send optimal distributions of K to optimal distributions of K'?*

We strongly believe the answer to this question is yes, or at least yes for similar conditions on K, but do not as yet have a proof of this statement. Special cases we have considered are given in the results section, and empirical results for the interval are given in the appendix.

If this statement holds, then we obtain optimal distributions of a large set of possible sets K as follows: if the exterior of K is given by applying $1/z$ to the interior of K with the interior nonempty, then using a conformal equivalence $\operatorname{int} K \to \operatorname{int} D$ given by the Riemann mapping theorem and composing on both sides $1/z$ yields a conformal equivalence from the exterior of K to the exterior of D as desired. Moreover, any such K will then have a unique optimal distribution: the image of the constant function on D under this conformal equivalence.

As a last detail, we note that knowing the optimal distribution on a large number of Ks is useful, but to truly understand generalized Leja sequences on such a K in the mode of Steinerberger, it is crucial to have an analogue of Fourier analysis. We return to this point in two subsections, stopping briefly to understand Steinerberger's methodology first.

2.5 Positive definite functionals

Steinerberger in [6] takes an arbitrary positive-definite real-valued function f with average value 0, (e.g. $-\log|2\sin x|$), and shows that a greedy sequence on a torus $\mathbb{T}^d$ which at each point minimises f results in a well-distributed sequence, giving explicit bounds for the discrepancy. We will outline the results in one dimension for simplicity.

Steinerberger's argument consists of two key steps. The first is that using Fourier analysis to rewrite f as a Fourier series with Fourier coefficients a_k, the total energy E_{total} after adding N points becomes:

$$E_{total} = \sum_{k\in\mathbb{Z}} a_k N^2 \left|\mathrm{Weyl}_{N-1}(k)\right|^2 \qquad \text{where } \mathrm{Weyl}_{N-1}(k) = \frac{1}{N}\sum_{j=1}^{N-1} e^{2\pi i k x_j}$$

Recalling Weyl's equidistribution theorem, we see that if $\mathrm{Weyl}_N(k) \to 0$, then our sequence is well-distributed. Furthermore, by the Erdős–Turán inequality, better bounds on these sums give us better distributed sequences. Therefore it suffices to control this energy to see how well distributed our sequence is.

The second key step is to get a bound on E_{total}. Steinerberger does this by bounding the minimum increase in energy at each step by the average increase in energy. This gives us a bound on E_{total} of the form $a_0N^2 + CN$. We subtract off the a_0N^2 on both sides and since all the a_k are positive by assumption, drop all but one term from the sum. This gives us bounds on the Weyl sums of $O(1/\sqrt{N})$, though better results are believed to be achievable. We believe that the total energy (above the base a_0N^2) is of the form $O(\log N)$, though we were unable to prove this.

In higher dimensions, the process is exactly analogous, except that instead of considering $\mathrm{Weyl}_N(k)$, we instead consider the equivalent average for the basis $f_{\mathbf{k}}(\mathbf{x}) = e^{2\pi i\langle \mathbf{k},\mathbf{x}\rangle}$ where $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{Z}^d$. Then as before, we require that the coefficients of f in this basis be positive, write the energy as a sum of "Weyl sums," and compare both sides.

2.6 Generalizing Steinerberger

The second step can be very easily generalised to a more general bounded set K on the plane. The trick of bounding the minimum energy increase by the average energy increase can be easily repeated.

It is the first step which is harder to do. We wish to find an orthonormal basis e_i of $L^2(\partial K)$. Then given such a basis, we wish to relate the standard $\mathbb{R}^2$ energy to an analogue of the Weyl sums on K: namely for each i, considering $\frac{1}{N}\sum_{j=1}^{N} e_i(x_j)$. Then just as in Weyl's equidistribution theorem, the x_i will be well-distributed when these converge to $a_i = \int e_i$.

3 Results

To begin, consider any nice curve ∂K with interior K such that the conditions of the previous section hold, i.e. we have a conformal equivalence M sending the complement of K to the complement of the disk D with ∂K mapping to ∂D. Additionally, take μ_D as the arc length measure on ∂D and let $\mu_K = M_*(\mu_D)$, the pushforward measure of μ_D by M.

3.1 Rotation and Energy Operators

Let R_θ be a map which "rotates" the space while preserving K. In particular, take

$$R_\theta z = M(e^{i\theta} \cdot M^{-1}(z)).$$

In words, R_θ takes z from ∂K to the unit circle, rotates by θ, and moves back to ∂K. Notice that when K is a disk and $M = Id$, then R_θ corresponds to ordinary rotation by θ. Moreover, we will also apply R_θ to functions. If $f : \partial K \to \mathbb{C}$ is a function, then we define $R_\theta f(z) = f(R_\theta z)$. Notice that since R_θ acts on the inputs to a function, it is clearly linear over functions: for any functions f and g and $a \in \mathbb{C}$ and $R_\theta(f+g) = R_\theta f + R_\theta g$ and $R_\theta(af) = aR_\theta(f)$.

Now, consider the following linear map T which, given a mass distribution f on ∂K (say of "electrons"), computes the potential field:

$$Tf(z) = -\int_{\partial K} f(z') \log|z - z'| \, dz'.$$

In particular, note that the integral is with respect to dz', *not* $d\mu_K$. We can now consider R_θ and T as linear operators over $L^2(\partial K, \mu_K)$ i.e. square-integrable functions on ∂K with inner product $\langle f, g\rangle = \int_{\partial K} f(t)\overline{g(t)} \, d\mu_K$.
We see that R_θ certainly takes square-integrable functions to square-integrable functions, for $\|f\|_{L^2(\partial K)} = \|R_\theta f\|_{L^2(\partial K)}$ by a change of variable.

3.2 Analysis of the disk

Notice that when K is a disk, we can take M as the identity map, meaning R_θ corresponds precisely to ordinary rotations and $\mu_K = \mu_D$. Suppose we wish to find the eigenfunctions of R_θ. Note that in principle, different values of θ may have different eigenfunctions, but by a continuity argument this cannot happen: any value of θ can be built up to arbitrary precision by repeated application of, say, R_1, since $n \pmod{2\pi}$ is dense in the circle for $n \in \mathbb{Z}$.

In this case, our eigenfunctions are exactly functions such that $f(e^{i\theta}z) = \lambda f(z)$. In particular, functions $f_n(z) = z^n$ for $n \in \mathbb{Z}$ work, since $\left(e^{i\theta}z\right)^n = e^{in\theta}z^n = \lambda z^n$ where $\lambda = e^{in\theta}$. This is the Fourier basis, since if we identify points on the circle with an angle $\alpha \in [0, 2\pi)$, then f_n indeed maps the point at α to $e^{in\alpha}$.

Moreover, when K is the disk, we see that T and R_θ easily commute as linear operators. In particular, since R_θ is just a rotation, we have that TR_θ is

$$\begin{aligned}
TR_\theta f(z) &= Tf(R_\theta z) \\
&= -\int_0^{2\pi} f(e^{i(\theta+\alpha)}) \log|z - e^{i\alpha}| \, d\alpha \\
&= -\int_0^{2\pi} f(e^{i\beta}) \log|z - e^{i(\beta-\theta)}| \, d\beta \\
&= -\int_0^{2\pi} f(e^{i\beta}) \log\left(|e^{-i\theta}| \, |e^{i\theta}z - e^{i\beta}|\right) \, d\beta
\end{aligned}$$

$$= -\int_0^{2\pi} f(e^{i\beta}) \log |R_\theta z - e^{i\beta}| \, d\beta$$
$$= R_\theta T f(z)$$

In the middle, we made the substitution $\beta = \theta + \alpha$. This confirms what is physically obvious: rotating a ring of electrons simply rotates its potential field. Note, however, that in our case f (the distribution of charge) can have a complex output, so in this case, the analogy with electrons sometimes fails.

It is a well-known fact that linear operators which commute share an eigenbasis. In fact, in this case, we have in particular that the eigenfunctions mentioned above are the common eigenbasis.

In general, we conjecture that some property like the following holds: if for some $w : K \to \mathbb{C}$ we have $d\mu_K = w(z)\, dz$, i.e. for any measurable $E \subset \partial K$

$$\mu_K(E) = \int_E w(z)\, dz$$

then the linear operators Tw and R_θ share an eigenbasis. Here, Twf is interpreted as $T(wf)$, where wf is pointwise multiplication. In other words,

$$Twf(z) = -\int_{\partial K} f(z')w(z') \log |z - z'| \, dz' = -\int_{\partial K} f(z') \log |z - z'| \, d\mu_K.$$

We have just seen this holds for K a disk, and will see in the next section that it holds for K an interval, but the statement appears to fail for K in general. We hope that some reformulation of this result (possibly for a better choice of operator than R_θ) will be possible.

Note that to obtain a $w(x)$ as above, it suffices (by change of variables for pushforward measures) to take $w(x)$ as the image of the constant map on the circle. Thus, our earlier hypothesis regarding conformal equivalences sending optimal distributions to optimal distributions suggests that $w(x)$ should be an optimal distribution–this is trivially the case for the circle, and as we will see holds for the interval as well.

3.3 Analysis of the interval

One conformal equivalence between the circle and the interval is the Joukowsky transform $M(z) = \frac{1}{2}\left(z + \frac{1}{z}\right)$. There are some issues since this is a “degenerate” map in the sense that every point on the interval has two points on the circle which map to it, but these issues will be resolved.

Note that our map sends $z = e^{i\theta}$ to $\cos\theta$. In particular, this implies that if we take $E = (a, b) \subset [-1, 1]$, then

$$\mu_K(E) = 2(\cos^{-1}(a) - \cos^{-1}(b)) = 2\int_a^b \left(-\frac{1}{\sqrt{1 - z^2}}\right) dz$$

suggesting

$$w(x) = -\frac{2}{\sqrt{1 - x^2}}.$$

where x is used in place of z as K lies on the real line. We note that the factor of 2 is simply an artifact of every open interval considered mapping to an interval on the upper half of the circle and the lower half: considering $w(z)$ as the limiting case of the ellipse as in the next subsection, we will take

$$w(x) = -\frac{1}{\sqrt{1-x^2}}$$

but note that the factor of 2 does not affect the eigenfunctions of Tw. For the sake of illustration, we take for granted the case when E is a more general measurable set. Note that by a similar argument to that in Section 2.3, along with the fact that the Chebyshev polynomials form an orthonormal basis with respect to the inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)} \frac{1}{\sqrt{1-x^2}}\, dx$$

we can show this is indeed the optimal distribution on the interval.

We claim that the Chebyshev polynomials are a common eigenbasis for the operators R_θ and Tw. It is well known that they are orthogonal with respect to our measure μ_K, which suggests that they may satisfy other nice properties as well. First, we will verify that R_θ preserves the Chebyshev polynomials. In fact, there is a minor issue – due to the degeneracy, there are two values that each point could rotate to. We fix this by letting

$$R'_\theta f(z) = \frac{1}{2}\left(f(M(e^{i\theta}z_1)) + f(M(e^{i\theta}z_2))\right)$$

where z_1 and z_2 are the two preimages of z in M. Note that $z_1 = z_2$ for $z = \pm 1$. Here, we could also use the sum of both terms instead of average; the intuition here is that the two points both contribute their "energy." We wish to show that

$$R'_\theta f_n = \lambda f_n$$

where in this case we happen to have $\lambda = \cos(n\theta)$, where f_n are the Chebyshev polynomials:

$$f_n(x) = \cos\left(n\cos^{-1}(x)\right).$$

This ultimately amounts to simple algebra:

$$\begin{aligned} R'_\theta f_n(x) &= \frac{1}{2}\left(\cos(n\cos^{-1}(x) + n\theta) + \cos(-n\cos^{-1}(x) + n\theta)\right) \\ &= \cos(n\theta)\cos(n\cos^{-1}(x)) \\ &= \lambda f_n(x) \end{aligned}$$

The first equation follows from the definition of R'_θ. Then, the second line follows from using the cosine sum identity and simplifying.

Next, we will show that the Chebyshev polynomials are also eigenfunctions of Tw. We have

$$Twf_n(x) = -\int_0^1 f_n(x')w(x')\log|x-x'|\,dx' = -\int_0^1 \frac{\cos(n\cos^{-1}(x'))\log|x-x'|}{\sqrt{1-x'^2}}\,dx'$$

Making the substitution $x' = \cos\theta$, we get

$$Twf_n(x) = -\int_0^\pi \cos(n\theta)\log|x - \cos\theta|\, d\theta$$

From a StackExchange post [9], we see that this integral is equal to:

$$Twf_n(x) = \frac{\pi}{n} f_n(x)$$

for $n > 0$. Moreover, for $n = 0$, $f_n(x)$ is constant, and it is not hard to see that the integral evaluates to the constant $\pi \log 2 \approx 2.178$ for any $-1 \leq x \leq 1$.

3.4 Application to Leja Sequences

Once we have a Fourier-like decomposition of functions in a general space K, we can decompose any function $g \in L^2(\partial K, \mu_K)$. If we take the eigenfunctions of Tw as our basis, which we call f_k with eigenvalues λ_k then assuming that w is not too bad, then we can decompose g as

$$g = \sum_k c_k w f_k.$$

Here g will represent a distribution of mass around the circle, either continuous or discrete. In the latter case we can approximate by a continuous distribution by considering sums of good kernels. By an identical argument to the one presented by Steinerberger and the case of the disk, we can use this "Fourier-like" series to analyze greedy sequences on ∂K analogous to those of Steinerberger in [6].

We will now see how these eigenfunctions of Tw let us relate our energy to an analogue of the Weyl sums using this Fourier decomposition.

We recall the definitions of E and T to be

$$E = \iint_{\partial K} -\log|z - z'| g(z) g(z') dz dz' \qquad Tg = \int_{\partial K} -\log|z - z'| g(z) dz.$$

This almost looks like $E = \langle Tg, g \rangle$, except that for the true inner product in $L^2(\partial K)$, we must integrate against $d\mu$ and not dz. Therefore we see that the energy is in fact $\langle Tg, g/w \rangle$. Now applying our decomposition of g into f_k, we get that

$$E = \sum_{k \geq 0} \lambda_k |c_k|^2$$

We will now assume the first eigenfunction of Tw is the constant function 1. This is reasonable, and is at least true on the interval and the circle.

Thus, c_0 is simply

$$\langle g/w, f_0 \rangle = \int_{\partial K} \frac{g(x)}{w(x)}\, d\mu_K = \int_{\partial K} g(x)\, dx$$

which is the total amount of mass.

Performing rearrangements identical to Steinerberger, we extract the $k = 0$ term of the energy to get:

$$E = \sum_{k \geq 0} \lambda_k |c_k|^2 = \lambda_0 N^2 + \sum_{k \geq 1} \lambda_k N^2 \left| \frac{c_k}{N} \right|^2$$

In a sense, c_k/N are our new "Weyl sums" which cannot be too large if E is small. In fact, when we let g be a discrete mass distribution, this is precisely what they are.

This immediately tells us that the lowest-energy configuration is that in which g/w decomposes only as f_0 with no other components. In other words, this occurs when $g = w$ or is some other multiple of w.

We now proceed to the second of Steinerberger's key steps: for a Leja sequence, we may again use the earlier argument to bound the energy by its average value. In particular, there is some constant C such that for each N:

$$E \leq \lambda_0 N^2 + CN$$

Combining this with our expansion of E together with the fact that each λ_k is at least 0 so each term in the sum is less than the entire sum, we can bound c_k/N as follows:

$$\left| \frac{c_k}{N} \right| \leq C \sqrt{\frac{1}{\lambda_k N}}$$

This yields a notion of regularity; indeed, we see the analogous Weyl sums tend to 0. It still remains open for us to rigorously connect the c_k with a notion of discrepancy with respect to the equilibrium density $w(x)$, but very roughly, if the c_k for $k > 0$ are not too large, then $f \approx w f_0 = w$, i.e. f becomes close to our optimal distribution.

3.5 Outlook

In this paper, we demonstrated a framework for connecting results regarding optimal distributions on the disk to those regarding optimal distributions on a large class of sets $K \subset \mathbb{C}$, and posed a conjecture regarding how such connections allow for useful analogues to Fourier decomposition. Moreover, we briefly illustrated how such a tool may be used for research on Leja sequences, with an eye toward proving regularity results for generalized Leja sequences, yielding a construction for low-discrepancy sequences on such sets $K \subset \mathbb{C}$. Further results could hopefully use the machinery presented here to finish the generalization of bounds given in works such as [8] to our case.

Acknowledgements

This project was made possible by the kind oversight of Rodrigo Angelo, who spent countless hours helping us navigate the literature, find suitable approaches, and ultimately create the above. His patient guidance and unique insights into the problem were the cornerstone of this research. We also would like to extend gratitude to Pawel Grzegrzolka for organising the SURIM research program and going to great lengths to make this opportunity a wonderful research experience.

References

[1] K. F. Roth. "On irregularities of distribution". In: *Mathematika* 1.2 (1954), pp. 73–79. DOI: 10.1112/S0025579300000541.

[2] Goluzin Gennadii Mikhailovich. *Geometric Theory of Functions of a Complex Variable*. American Mathematical Society, 1969, pp. 309–314. ISBN: 0821815768.

[3] Wolfgang M. Schmidt and T. N. Shorey. *Lectures on irregularities of distribution*. Tata Institute of Fundamental Research, 1977.

[4] Michael Drmota and Robert F. Tichy. *Sequences, Discrepancies and Applications, Lecture Notes in Mathematics, 1651*. 1997.

[5] C.-S. Lin. "On Log-trigonometric Functions". In: *Crux Mathematicorum* 29 (7 2003), pp. 460–463.

[6] Louis Brown and Stefan Steinerberger. "Positive-definite Functions, Exponential Sums and the Greedy Algorithm: a curious Phenomenon". In: *arXiv e-prints*, arXiv:1908.11228 (Aug. 2019), arXiv:1908.11228. arXiv: 1908.11228 [math.CA].

[7] Stefan Steinerberger. "A Nonlocal Functional Promoting Low-Discrepancy Point Sets". In: *arXiv e-prints*, arXiv:1902.00441 (Feb. 2019), arXiv:1902.00441. arXiv: 1902.00441 [math.OC].

[8] Stefan Steinerberger. "Polynomials with Zeros on the Unit Circle: Regularity of Leja Sequences". In: *arXiv e-prints*, arXiv:2006.10708 (June 2020), arXiv:2006.10708. arXiv: 2006.10708 [math.CA].

[9] Nguyen Quan (https://math.stackexchange.com/users/794566/nguy%e1%bb%85n-qu%c3%a2n). *Show that this integral is a multiple of a Chebyshev polynomial*. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/3802781 (version: 2020-08-27). eprint: https://math.stackexchange.com/q/3802781. URL: https://math.stackexchange.com/q/3802781.

Appendix: Empirical Results for the Interval

One of the major motivators for the question posed in Section 2.4 was the following empirical result. Assuming that the question is answered in the affirmative, we find as in Section 3.3 that

$$w(x) = -\frac{1}{\sqrt{1-x^2}}.$$

Discretizing the interval $[-1,1]$ into points spaced 0.0001 apart, we take a single point at 0 and then iteratively add the discretized point to our list which maximizes the product of distances between points. 1,000 points are chosen in this manner, and their locations are binned into a normalized histogram with bin width 0.05. $w(x)$ is then normalized and plotted on top, yielding this plot:

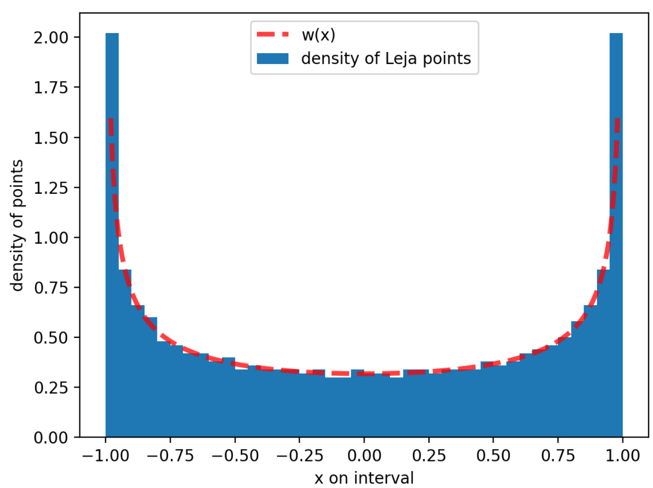

Figure 1: A plot of the density of points in the Leja sequence on the interval as compared to $w(x)$.

The closeness of the normalized version of $w(x)$ to the distribution of points from the Leja sequence illustrates the close connection between Leja sequences, non-greedy sets of points which minimize energy, and optimal distributions as suggested in Sections 2.1-2.4. Below are copied the scripts used to generate the above plot:

`dots.py`: generates point data

```
import csv
import numpy as np
from scipy import optimize
from tqdm import tqdm

NUM_POINTS = 1000
STEP = 0.0001

grid = np.reshape(np.arange(-1, 1 + STEP, STEP), (-1, 1))
points = [0.0]
for _ in tqdm(range(NUM_POINTS)):
    ps = np.reshape(points, (1, -1))
    dists = np.abs(ps - grid)
    prods = np.prod(dists, axis=1)
    max_ind = np.argmax(prods)
    points.append(grid[max_ind, 0])

with open('dump.csv', 'w') as f:
    writer = csv.writer(f)
    writer.writerow(points)
```

`dotplot.py`: plots generated point data

```
import csv
import matplotlib.pyplot as plt
import numpy as np

BIN_SIZE = 0.05
STEP = 0.0001

with open('dump.csv') as f:
    reader = csv.reader(f)
    points = [float(val) for val in next(reader)]

bins = np.arange(-1, 1+BIN_SIZE, BIN_SIZE)
plt.hist(points, bins=bins, density=True, alpha=1.0)

x = np.arange(-0.98 + STEP, 0.98, STEP)
y = 1.0 / (np.pi * np.sqrt(1 - x**2))
plt.xlabel('x on interval')
plt.ylabel('density of points')
plt.plot(x, y, linestyle='dashed', color='red', linewidth=3, alpha
    =0.75)
plt.legend(['w(x)', 'density of Leja points'])

plt.show()
```