Discretizing Manifolds via Minimum Energy Points

D. P. Hardin and E. B. Saff

There are a variety of needs for the discretization of a manifold—statistical sampling, quadrature rules, starting points for Newton's method, computer-aided design, interpolation schemes, finite element tessellations—to name but a few. So let us assume we are given a d-dimensional manifold A in the Euclidean space $\mathbb{R}^{d'}$ and wish to determine, say, 5,000 points that "represent A". How can we go about this if A is described by some geometric property or by some parametrization of the unit cube $U^d := [0,1]^d$ in $\mathbb{R}^d$? Naturally, we must be guided by the particular application in mind.

For a historical perspective as well as a brief motivational journey, let us look at the simple case when A is the interval $[-1,1] \subset \mathbb{R}$. One obvious choice for N points that discretize A is the set of equally spaced points

$$x_{k,N} = -1 + \frac{2k}{N-1}, \qquad N \geq 2,\ k = 0, \ldots, N-1.$$

These points also enjoy the property of solving the "best-packing" problem on $[-1,1]$; in general, a set of distinct points $\omega_N^* = \{x_1^*, \ldots, x_N^*\} \subset A$ solves the *N-point best-packing problem* on a compact set A if

$$\min_{i \neq j} |x_i^* - x_j^*| = \max_{\omega_N \subset A} \min_{i \neq j} |x_i - x_j|,$$

where the maximum is taken over all N-point subsets $\omega_N = \{x_i\}_1^N$ of A. But suppose our interest is in selecting N points for quadrature or for polynomial interpolation of a smooth function $f(x)$ on $[-1,1]$. Then, as shown by Runge, the choice of equally spaced points (or, indeed, any asymptotically uniformly distributed sets of points) can be disastrous (in fact, the norm of the polynomial interpolation operator grows *geometrically* large with N). Rather, choosing N points of $[-1,1]$ that asymptotically (as $N \to \infty$) have the arcsine distribution $(1/\pi)dx/\sqrt{1-x^2}$ (such as the zeros of the classical Chebyshev polynomials $T_N(x) = \cos(N \arccos x)$ shown in Figure 1) does a much better job—one can achieve polynomial interpolation operator norm $\mathcal{O}(\log N)$.

The connection between efficient univariate polynomial interpolation (or Gaussian quadrature) and the arcsine distribution becomes clearer on observing that any monic polynomial $p_N(x) = \prod_{i=1}^N (x - x_i)$ satisfies

$$\frac{1}{N} \log \frac{1}{|p_N(x)|} = \int \log \frac{1}{|x-t|}\, d\nu_N(t), \tag{1}$$

where ν_N is the normalized counting measure

$$\nu_N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$$

with δ_x denoting the unit point mass at x. In other words, $(1/N)\log(1/|p_N(x)|)$ is a logarithmic potential

D. P. Hardin is associate professor of mathematics at Vanderbilt University. His email address is hardin@math.vanderbilt.edu.

E. B. Saff is professor of mathematics at Vanderbilt University. His email address is esaff@math.vanderbilt.edu. *The research of this author was supported in part by the U. S. National Science Foundation under grant No. DMS-0296026.*

for a discrete probability measure. Classical potential theory shows that the energy integral

$$I_0[\mu] := \iint \log \frac{1}{|x-t|}\, d\mu(x)\, d\mu(t), \tag{2}$$

where μ is any probability measure (normalized, positive, Radon measure) supported on $[-1,1]$, attains its minimum when $d\mu$ is the arcsine distribution, which is called the *equilibrium measure* or *Robin measure* for $A = [-1,1]$.

It was M. Fekete who explored the connection between polynomial interpolation and the discretized version of (2), which, for given N, consists in finding an N-point set $\omega_N^F = \{x_{k,N}^F\}_1^N \subset A$ that minimizes the logarithmic energy

$$E_0(\omega_N) := \sum_{i\neq j} \log \frac{1}{|x_i - x_j|} = 2 \sum_{1\le i<j\le N} \log \frac{1}{|x_i - x_j|} \tag{3}$$

over all N-point subsets $\omega_N = \{x_i\}_1^N$ of A. (For $A = [-1,1]$, such points are the zeros of the Jacobi polynomial $P_{N-2}^{(1,1)}(x)$ together with $x = \pm 1$.) Provided that a compact set $A \subset \mathbb{R}^2$ has positive logarithmic capacity, the sequence of normalized counting measures for the *Fekete point sets* $\{\omega_N^F\}$ converges in the weak-star topology (as $N \to \infty$) to the unique measure μ_A that minimizes the energy integral (2) over all probability measures supported on A (cf. [17]); moreover,

$$\lim_{N\to\infty} \frac{E_0(\omega_N^F)}{N^2} = I_0[\mu_A]\,.$$

(So, in particular, the Fekete point sets for $A = [-1,1]$ asymptotically have the arcsine distribution.) The condition that A have positive logarithmic capacity simply means that there is at least one probability measure μ on A for which the energy integral is finite. This condition will play a crucial role as we discuss point sets that minimize other energy functionals.

Both the equally spaced points and the Fekete points for $A = [-1,1]$ can be regarded as limiting cases of point sets that minimize the discrete Riesz energy. For a fixed parameter $s > 0$, the **Riesz s-energy** of a set $\omega_N = \{x_i\}_1^N$ of N distinct points in $\mathbb{R}^{d'}$ is given by

$$E_s(\omega_N) := \sum_{i\neq j} \frac{1}{|x_i - x_j|^s}\,, \tag{4}$$

where $|\cdot|$ denotes Euclidean distance. For $s = 0$, we use the definition in (3). Given a compact set $A \subset \mathbb{R}^{d'}$ with infinitely many points, we denote the ***N*-point minimal *s*-energy over** A by

$$\mathcal{E}_s(A,N) := \inf_{\omega_N \subset A} E_s(\omega_N). \tag{5}$$

Notice that as $s \to \infty$, with N fixed, the s-energy (5) is increasingly dominated by the term(s) involving the smallest of pairwise distances and, in this sense, leads to the best-packing problem on A. On the other hand, as $s \to 0$, it is easily verified that for each $N \ge 2$,

$$\frac{\mathcal{E}_s(A,N) - N(N-1)}{s} \to \mathcal{E}_0(A,N).$$

So natural questions that arise are:

Q1: *How are minimal s-energy configurations for A distributed for large N?*

Q2: *How does the asymptotic behavior of $\mathcal{E}_s(A,N)$ in* (5) *depend on A and s?*

For $A = [-1,1]$ and $0 \le s < 1$, explicit answers can be found in [11], where it is shown using potential theoretic arguments that optimal s-energy points have the limit distribution (as $N \to \infty$)

$$d\lambda_s = \frac{c_s}{(1-x^2)^{(1-s)/2}} dx, \quad x \in (-1,1), \tag{6}$$

where c_s is a normalizing constant. Furthermore,

$$\lim_{N\to\infty} \frac{\mathcal{E}_s([-1,1],N)}{N^2} = \frac{\sqrt{\pi}\,\Gamma(1+s/2)}{\cos(\pi s/2)\Gamma((1+s)/2)}, \quad 0 < s < 1. \tag{7}$$

The potential theoretic argument proceeds as in the case of Fekete points by showing that any limit distribution of optimal s-energy points minimizes the energy integral

$$I_s[\mu] := \iint \frac{1}{|x-y|^s} d\mu(x) d\mu(y) \tag{8}$$

over all probability measures μ supported on $[-1,1]$ and then appealing to the fact that such a measure is unique and given by (6). The limit in (7) is simply $I_s[\lambda_s]$.

But What If $s \ge 1$? In this case we have $I_s[\mu] = \infty$ for all probability measures μ on $[-1,1]$, and so the preceding argument fails. Yet a glance at the distribution (6) reveals that as s increases from 0 to 1, the equilibrium distributions λ_s transform from the arcsine to the uniform (normalized Lebesgue) distribution, which is the distribution of the best-packing points corresponding to $s = \infty$. Thus we might expect that for *every* fixed $s \ge 1$, optimal s-energy points are uniformly distributed in the

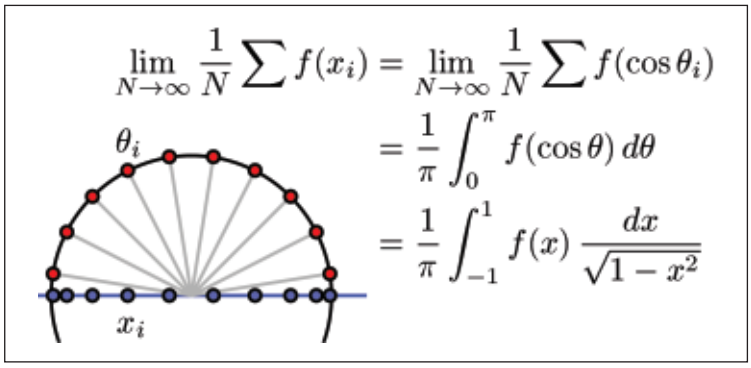


Figure 1. Chebyshev polynomial zeroes (blue dots) have arcsine limit distribution.

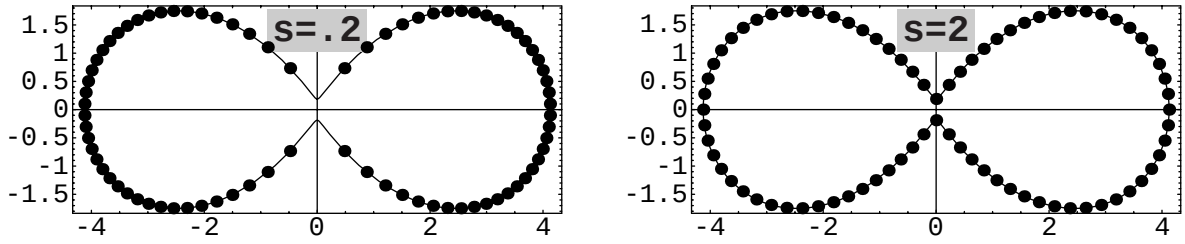


Figure 2. Near optimal energy configurations for $s = 0.2$ (left) and for $s = 2$ (right) with 100 points for the affinely scaled Cassinian oval given by (11).

limit, and this turns out to be true in a much more general context that we describe below. The predicted analog of (7) is, however, less obvious.

Minimal Energy Points on Curves. For the case when A is a rectifiable Jordan arc or curve in $\mathbb{R}^{d'}$, answers to Q1 and Q2 are given by A. Martinez-Finkelshtein et al. [12]. They show that for $s = 1$

$$\lim_{N\to\infty} \frac{\mathcal{E}_1(A,N)}{N^2 \log N} = \frac{2}{L} \tag{9}$$

and for $s > 1$

$$\lim_{N\to\infty} \frac{\mathcal{E}_s(A,N)}{N^{1+s}} = \frac{2\zeta(s)}{L^s}, \tag{10}$$

where L is the arclength of A and $\zeta(s)$ denotes the classical Riemann zeta function. Moreover, for each $s \geq 1$, the limit distribution of asymptotically optimal s-energy configurations for A is uniform with respect to arclength measure on A. The situation for $0 \leq s < 1$ on such curves is treated, as above, via potential theory; the limit distribution of asymptotically optimal points is the unique measure $\lambda_{A,s}$ that minimizes the energy integral $I_s[\mu]$ in (8) (or (2)) over all probability measures μ supported on A, and the energy $\mathcal{E}_s(A,N)$ grows like N^2 (more precisely, $\mathcal{E}_s(A,N)/N^2 \to I_s[\lambda_{A,s}]$ as $N \to \infty$). Hence as s increases from zero, the minimum energy growth switches from order N^2 to order N^{1+s}, with the transition occurring at $s = 1$ where the energy growth is of order $N^2 \log N$. This transition is signaling a change from global to local effects, with the influence of nearby neighbors becoming more and more dominant as s increases beyond 1 (indeed, at $s = \infty$, *only* the nearest neighbors are significant).

The (curious) appearance of the zeta function in (10) arises from the following observation: For any set of points $\omega_N = \{x_k\}_1^N$ that are listed in consecutive order along a Jordan arc A, we can get a lower bound for $E_s(\omega_N)$ by setting $d_{i,j}$ equal to the length of the subarc from x_i to x_j and noting that

$$E_s(\omega_N) \geq \sum_{k=1}^{N-1} \hat{E}_k, \text{ where } \hat{E}_k := \sum_{|i-j|=k} \frac{1}{d_{i,j}^s}.$$

From the convexity of x^s for $s \geq 1$, together with the harmonic-arithmetic mean inequality, we deduce that

$$\hat{E}_k = 2\sum_{j=1}^{N-k} \frac{1}{d_{j,j+k}^s} \geq 2(N-k)^{1-s} \left(\sum_{j=1}^{N-k} \frac{1}{d_{j,j+k}} \right)^s$$
$$\geq \frac{2(N-k)^{1+s}}{\left(\sum_{j=1}^{N-k} d_{j,j+k}\right)^s} \geq \frac{2(N-k)^{1+s}}{(kL)^s}.$$

Adding these lower estimates, dividing by N^{1+s}, and letting $N \to \infty$ give rise to $\sum_{k=1}^{\infty} k^{-s} = \zeta(s)$. In the case of a piecewise smooth Jordan curve without cusps, equally spaced points along A provide an asymptotically sharp upper bound. (The upper bound for general Jordan curves requires finer analysis.)

By way of illustration, consider the (affinely scaled) Cassinian oval A given parametrically by

$$(x(t), y(t)) := r(t)(2\cos t, 3\sin t) \quad (0 \leq t \leq 2\pi), \tag{11}$$

where $r(t) := \cos(2t) + \sqrt{a^4 + \cos^2(2t)}$ and $a = 0.6$. Figure 2 shows numerically computed optimal s-energy configurations for A with $N = 100$ points for $s = 0.2$ and $s = 2$, demonstrating the dependence of the limit distribution on s. The case $s = 2$ clearly indicates nearly uniformly distributed points on A as is expected for any $s \geq 1$. In contrast, the configuration for $s = 0.2$ is distributed according to the (nonuniform) equilibrium measure for the corresponding Riesz energy integral.

Moving to Higher-Dimensional Manifolds. Here a canonical choice for the manifold A is the sphere $S^d = \{x \in \mathbb{R}^{d+1} : |x| = 1\}$, for which optimal point configurations have been the subject of considerable investigation. Indeed, the case $s = \infty$ of best-packing is the famous *Tammes's problem* or *hard-spheres problem*, which has its origin in a botanist's attempt to describe patterns of pores on spherical pollen grains (optimal configurations for this problem are known explicitly only for a handful of integers N). The case $s = 0$, which is the same as maximizing the product of pairwise distances $\prod_{i\neq j} |x_i - x_j|$ over all N-point sets $\omega_N \subset S^d$,

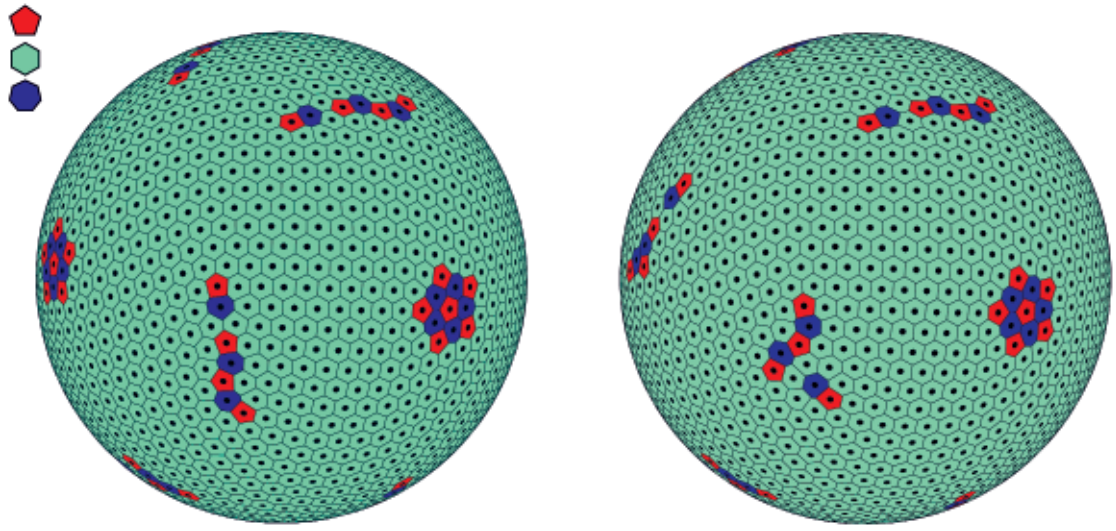


Figure 3. Near optimal s-energy configurations for $s = 1$ (left) and $s = 4$ (right) with 1,600 points on the sphere S^2.

arises, for example, in the study of computational complexity, where M. Shub and S. Smale [18] investigate good starting points for Newton's method on the sphere. Smale, in his list [19] of problems for the current century, states as Problem #7 the challenge to design a fast algorithm for generating "nearly optimal" logarithmic energy points: namely, to compute (in polynomial time with respect to N) an N-point set $\omega_N \subset S^2$ so that

$$(12)\quad E_0(\omega_N) \le \mathcal{E}_0(S^2, N) + C \log N\,, \quad N = 2, 3, \ldots,$$

for some positive constant C. While far from meeting this challenge, a variety of fast methods have been devised (see, e.g., the algorithms for "spiral points" in [14], [16] and for "equal area points" in [14], [20], the latter being downloadable from `http://math.vanderbilt.edu/~esaff/sphere_points.html` and recently extended to S^d for arbitrary d by I. Sloan, R. Womersley, and P. Leopardi).

We further note that for $s = 1$ and $A = S^2$, the minimization in (5) is the classical *Thomson problem* of electrons restricted to the sphere and interacting through the Coulomb potential (see, e.g., [7], [4]), which is relevant not only in electrostatics but also in molecular modeling (crystallography, stable carbon molecules, fullerenes) as well as in the study of certain viral structures. Extensive computations of optimal configurations appear in a number of articles spanning the physics, chemistry, and mathematics literature. A particularly convenient listing is provided by Hardin, Sloane, and Smith, whose findings are accessible via the Internet address `http://www.research.att.com/~njas/electrons/` (see also [15]).

For large N the numerical determination of minimum energy points is a difficult constrained optimization problem. Indeed, it appears (cf. [7], [15]) that the number of relative minima (ignoring rotations and reflections) grows exponentially with N (at least for certain subsequences of integers). Beyond a few hundred points, finding a global minimum of energy is always accompanied with some uncertainty. Yet ad hoc numerical methods devised by Alar Toomre and others for N in the thousands have generated configurations on the sphere that reveal rather startling features. Figure 3 provided by R. Womersley shows (near) optimal s-energy configurations for $N = 1,600$ points when $s = 1$ and $s = 4$. These illustrations display the tessellations

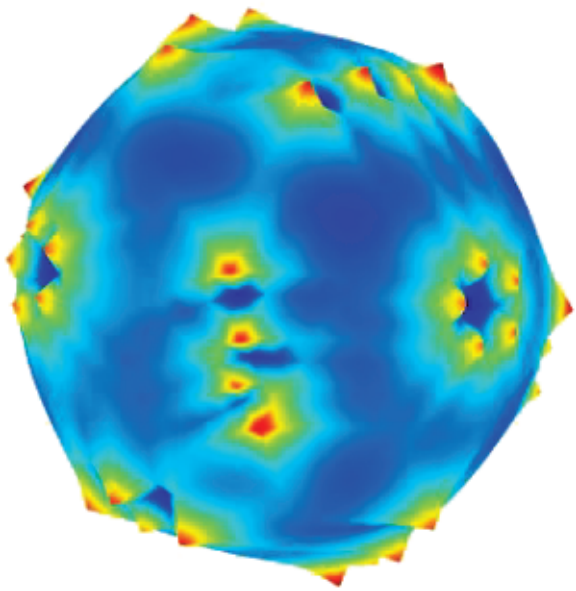


Figure 4. The point energies for the near optimal 1-energy configuration shown in Figure 3.

of the sphere created by the Voronoi cells ("school districts") corresponding to these optimal points. Notice that the vast majority of cells are nearly regular hexagons (imitating best coverings of the plane). But there also appear spherical pentagons (as in the standard soccer ball design; see red cells) as well as heptagons (see blue cells). The heptagons seem not to be present for N less than 300, but for thousands of points they do occur in significant number and are paired with pentagonal cells. Furthermore, the nonhexagonal cells (called *defects* or *disclinations*) appear to form 12 "scars" (or sometimes "buttons") roughly centered at the vertices of an inscribed icosahedron (cf. Figure 3). As illustrated in Figure 4, the *point energies* $E_s(x_i) := \sum_{j=1,\ j\neq i}^{N} |x_i - x_j|^{-s}$ are nearly equal for the sea of "hexagonal points", while the "pentagonal points" have relatively elevated energies and the "heptagonal points" have relatively lower energies.

The appearance of nonhexagonal cells is no surprise, since an Euler characteristic computation readily implies that the sphere cannot be covered by hexagons alone. But what is fascinating is that the twelve formations of these five and seven nearest neighbor points appear to be independent of the ground potential (e.g. independent of the parameter s for the Riesz potential). These observations by Bowick et al. [4] have also been confirmed by laboratory experiments in which polystyrene beads (one micron in diameter) attach themselves to a water droplet suspended in an oily mixture (cf. [2]). Focusing on these twelve scars has the considerable advantage of reducing the number of variables in the optimization problem and may in the future lead to fast generation of nearly optimal configurations for N in the thousands. Whether such scars persist for even larger orders of N is as yet unknown but may be of crucial importance for asymptotic results.

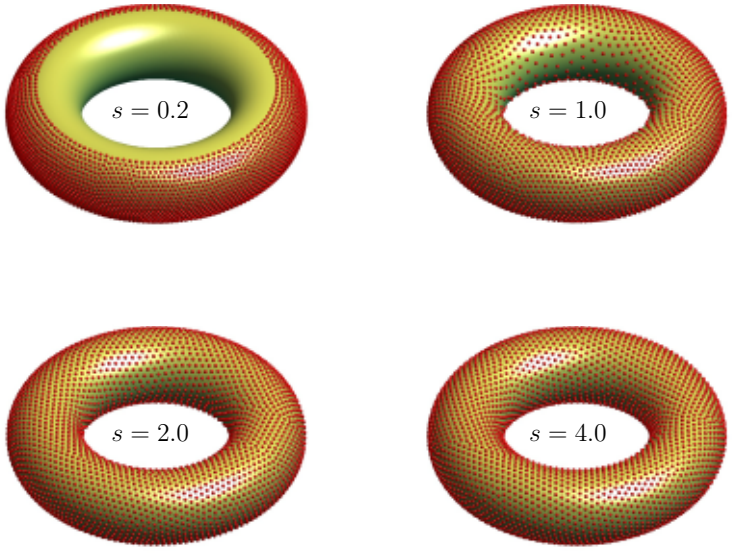


Figure 5. Near optimal s-energy configurations for $s = 0.2$, $s = 1$, $s = 2$, and $s = 4$ with 4,000 points for a torus in $\mathbb{R}^3$.

Questions Q1 and Q2 for the d-Sphere. The asymptotics for the minimal energy $\mathcal{E}_s(S^d, N)$ is (as one would suspect from the above discussion for curves) quite different for the three cases $0 \le s < d$, $s = d$, and $s > d$. Indeed, for $0 \le s < d$, the energy integral (8) or (2) for probability measures supported on S^d attains its finite minimum when μ is normalized surface area σ^d on the sphere and potential theory then gives

$$\lim_{N\to\infty} \frac{\mathcal{E}_s(S^d, N)}{N^2} = I_s(\sigma^d) = \frac{\Gamma((d+1)/2)\Gamma(d-s)}{\Gamma((d-s+1)/2)\Gamma(d-s/2)}, \quad 0 < s < d, \tag{13}$$

as well as the fact that optimal s-energy configurations are asymptotically uniformly distributed with respect to σ^d. For $s \ge d$, however, $I_s(\mu) = +\infty$ for all probability measures on S^d and different methods are needed for analysis. Using spherical harmonics and positivity results, it is shown by Kuijlaars and Saff [10] that for $s = d$

$$\lim_{N\to\infty} \frac{\mathcal{E}_d(S^d, N)}{N^2 \log N} = \frac{\Gamma((d+1)/2)}{d\sqrt{\pi}\,\Gamma(d/2)} = \frac{\mathrm{Vol}(\mathcal{B}^d)}{\mathrm{Area}(S^d)},$$

where $\mathcal{B}^d$ is the unit ball in $\mathbb{R}^d$, from which it follows that d-optimal configurations are asymptotically uniformly distributed. For $s > d$, it is not difficult to show that the order of growth of the minimal energy becomes $N^{1+s/d}$, again signaling the increasing dominance of local interactions. Yet more precise limit formulas such as the analogue of (13) as well as a rigorous proof that optimal configurations for $s > d$ are asympotically uniformly distributed (as symmetry would tend to dictate) require a completely different approach that we will describe below. Unlike the simple case $d = 1$ for a rectifiable curve where we can consider points in the systematic order described earlier, handling nearest neighbors in higher dimensions can present quite a challenge. What is fortuitous is that attempts to deal with the sphere for $s > d$ have led to a general argument that resolves questions Q1 and Q2 for a large class of d-manifolds in $\mathbb{R}^{d'}$.

A General Result. For an arbitrary compact set $A \subset \mathbb{R}^{d'}$ with Hausdorff dimension d_H, potential theory provides answers to Q1 and Q2 for $0 \le s < d_H$ via the s-energy equilibrium measure $\lambda_{A,s}$ that minimizes (8) over all probability measures supported on A (cf. [11] and [13]). For $s \ge d_H$, the following recent result applies to rectifiable manifolds A.

Recall that a mapping $\phi : B \to \mathbb{R}^{d'}$ is said to be *bi-Lipschitz on $B \subset \mathbb{R}^d$ with constant $L > 0$* if

$$(1/L)|x-y| \le |\phi(x)-\phi(y)| \le L|x-y|,\ x,y \in B.$$

We say that $A \subset \mathbb{R}^{d'}$ is a *d-rectifiable manifold* if it is a compact subset of a finite union of bi-Lipschitz images of open sets in $\mathbb{R}^d$.

Theorem 1 *([9]). Suppose $s \ge d$ and $A \subset \mathbb{R}^{d'}$ is a d-rectifiable manifold. When $s = d$ we further assume A is a subset of a d-dimensional C^1 manifold. Let $\mathcal{H}_d$ denote d-dimensional Hausdorff measure on $\mathbb{R}^{d'}$. Then for $s = d$ we have*

$$\lim_{N\to\infty} \frac{\mathcal{E}_d(A,N)}{N^2 \log N} = \frac{\mathcal{H}_d(\mathcal{B}^d)}{\mathcal{H}_d(A)}, \tag{14}$$

while for $s > d$, the limit $\lim_{N\to\infty} \mathcal{E}_s(A,N)/N^{1+s/d}$ exists and is given by

$$\lim_{N\to\infty} \frac{\mathcal{E}_s(A,N)}{N^{1+s/d}} = \frac{C_{s,d}}{\mathcal{H}_d(A)^{s/d}}, \tag{15}$$

where $C_{s,d}$ is a finite positive constant independent of A and d'.

If $\mathcal{H}_d(A) > 0$, then for each $s \ge d$, any sequence of optimal (or asymptotically optimal) s-energy configurations ω_N is uniformly distributed (as $N \to \infty$) with respect to d-dimensional Hausdorff measure restricted to A.

In particular, the theorem holds for any compact subset A of $\mathbb{R}^d$ as well as any compact subset of a smooth d-dimensional manifold. It is interesting to note that the limit (14) is simply $1/\rho^d$ when A is a ball in $\mathbb{R}^d$ with radius ρ.

The constant $C_{s,d}$ in (15) certainly depends on the normalization for Hausdorff measure. Here we choose $\mathcal{H}_d$ on $\mathbb{R}^{d'}$ normalized so that any isometric image of the unit cube $U^d := [0,1]^d$ in $\mathbb{R}^{d'}$ has $\mathcal{H}_d$-measure 1. Then, for $s > d$,

$$C_{s,d} = \lim_{N\to\infty} \frac{\mathcal{E}_s(U^d,N)}{N^{1+s/d}}. \tag{16}$$

For $d = 1$ we deduce from (10) that $C_{s,1} = 2\zeta(s)$ for $s > 1$; however, for $d \ge 2$, the determination of the constant $C_{s,d}$ for $s > d$ remains an open problem. For $d = 2$ (as seen for the sphere) the hexagonal lattice $L \subset \mathbb{R}^2$ consisting of points of the form $m(1,0) + n(1/2, \sqrt{3}/2)$ for $m, n \in \mathbb{Z}$ appears to play the central role in determining $C_{s,2}$. Assuming that most points in optimal configurations live in the "hexagonal sea" and are centers of regular hexagons with area $\approx \mathcal{H}_2(A)/N$, it is natural to conjecture that the constant $C_{s,2}$ is given by $(\sqrt{3}/2)^{s/2}\,\zeta_L(s)$, where $\zeta_L(s) := \sum_{X\in L, X\neq 0} |X|^{-s}$ is the zeta function for the lattice L. It is shown in [10] that for the sphere S^2

$$\limsup_{N\to\infty} \frac{\mathcal{E}_s(S^2,N)}{N^{1+s/2}} \le \left(\frac{\sqrt{3}}{8\pi}\right)^{s/2} \zeta_L(s),\ (s > 2), \tag{17}$$

which implies that $(\sqrt{3}/2)^{s/2}\,\zeta_L(s)$ is an upper bound for $C_{s,2}$.

Theorem 1 provides the order of growth of $\mathcal{E}_s(A,N)$ and the limit distribution of optimal configurations for a d-rectifiable manifold A only if $\mathcal{H}_d(A) > 0$. If $\mathcal{H}_d(A) = 0$, then the right-hand side of (15) is understood to be ∞, and, in this case, Theorem 1 provides only a lower bound on the order of growth of $\mathcal{E}_s(A,N)$. When $0 < \mathcal{H}_d(A) < \infty$, we observe, as with the sphere S^d, that the minimum energy experiences a transition in order of growth as s increases from values less than d to values greater than d (that is, from N^2 to $N^{1+s/d}$ with the transition value of $s = d$ giving growth of order $N^2 \log N$). Moreover, as s increases, the limit distribution of optimal (or near optimal) points becomes and remains the uniform distribution when $s \ge d$. The latter is particularly significant with regard to applications that involve integration with respect to Hausdorff (Lebesgue) measure.

For example, let A be the torus in $\mathbb{R}^3$ obtained by revolving about the z-axis the circle in the xz-plane of radius 1 centered at (3,0,0). Figure 5 shows near optimal s-energy configurations for A with $s = 0.2$, $s = 1$, $s = 2$, and $s = 4$ with $N = 4000$ points. In contrast to the case of the sphere, the equilibrium measure $\lambda_{A,s}$ for $s < 2$ is no longer uniform, and thus we find qualitatively different s-energy configurations for $s < 2$ (points are distributed more densely around the outer ring) from those for $s \ge 2$ (points are distributed uniformly on A in accordance with Theorem 1). The pictures in Figure 4, together with animations prepared by R. Womersley showing near optimal s-energy configurations for the torus A, are available at the Internet address `http://www.maths.unsw.edu.au/~rsw/Torus/`.[1] Because the Euler characteristic for a torus is 0, nonhexagonal "defects" are not required as they are for the sphere. In [5] Bowick et al. investigate configurations on the torus with long-range bond-orientational order. For such "hexatic" phase configurations they predict that pentagonal and heptagonal defects should be present in optimal configurations when N is below a critical value N_c depending on the "aspect ratio" and that there should be no defects when $N > N_c$. For the torus A with aspect ratio 3, they predict $N_c \approx 10^4$.

So how does one go about proving Theorem 1, given that our potential theoretic tools are no longer available when $s \ge d$? Here we focus our discussion on the $s > d$ case, since the $s = d$ case is more technical, relying on the previously

[1] *A natural question suggested by the configurations shown in Figure 5 (and even more so by the animations of Womersley) is whether $s^* = 1$ is the critical value of s such that for $s < s^*$ the support of the equilibrium measure $\lambda_{A,s}$ is a proper subset of A.*

mentioned result for the d-sphere. The first step in the proof is to show the existence of the limit (15) for the cube.

The Argument for the Unit Cube U^d. We begin with an optimal arrangement ω_N^* minimizing the s-energy of N points in U^d and use it to obtain an upper bound for the minimal energy of m^dN points of U^d. So let $m \in \mathbb{N}$ and set $\mathbb{Z}_m^d := \{0, 1, ..., m-1\}^d$. For $0 < \gamma < 1$ and $\mathbf{i} = (i_1, ..., i_d) \in \mathbb{Z}_m^d$, consider the m^d disjoint subcubes

$$U_{\mathbf{i}} = \frac{1}{m}(\gamma U + \mathbf{i}) = \Big[\frac{1}{m}i_1, \frac{1}{m}(i_1+\gamma)\Big] \times \cdots \times \Big[\frac{1}{m}i_d, \frac{1}{m}(i_d+\gamma)\Big].$$

Then we obtain m^dN points of U^d by scaling and translating ω_N^* to each $U_{\mathbf{i}}$ (see Figure 6), yielding

$$\omega_{m^dN} := \bigcup_{\mathbf{i}\in\mathbb{Z}_m^d} \frac{1}{m}(\gamma\omega_N^* + \mathbf{i}),$$

and we note that $\frac{1}{m}(\gamma\omega_N^* + \mathbf{i})$ is an optimal N-point configuration for $U_{\mathbf{i}}$. To estimate $E_s(\omega_{m^dN})$, we separate the energy terms arising from pairs of points in the same subcube (the terms involving interactions between distinct subcubes will turn out to be relatively negligible):

$$(18)\quad \mathcal{E}_s(U^d, m^dN) \le E_s(\omega_{m^dN}) \le \sum_{\mathbf{i}\in\mathbb{Z}_m^d} \Big\{\mathcal{E}_s(U_{\mathbf{i}}, N) + \sum_{\mathbf{j}\in\mathbb{Z}_m^d,\ \mathbf{j}\neq\mathbf{i}} N^2 \mathrm{dist}(U_{\mathbf{i}}, U_{\mathbf{j}})^{-s}\Big\}.$$

Next, from the translation invariance and scaling properties of the Riesz s-energy we can write $\mathcal{E}_s(U_{\mathbf{i}}, N) = m^s\gamma^{-s}\mathcal{E}_s(U^d, N)$, which we use along with the inequality $\mathrm{dist}(U_{\mathbf{i}}, U_{\mathbf{j}}) \ge \frac{1-\gamma}{m}|\mathbf{i}-\mathbf{j}|$ to deduce from (18) the estimate

$$(19)\quad \mathcal{E}_s(U^d, m^dN) \le m^{s+d}\{\gamma^{-s}\mathcal{E}_s(U^d, N) + K(1-\gamma)^{-s}N^2\},$$

where $K := \sum_{\mathbf{k}\in\mathbb{Z}^d, \mathbf{k}\neq\mathbf{0}} |\mathbf{k}|^{-s} < \infty$ for $s > d$.

Let L_t denote a subsequence for which

$$\overline{g}_{s,d} := \limsup_{N\to\infty} \mathcal{E}_s(U^d, N)/N^{1+s/d} = \lim_{t\to\infty} \mathcal{E}_s(U^d, L_t)/L_t^{1+s/d}$$

and choose N_* so that

$$(20)\quad N_*^{1-s/d} < (1-\gamma)^{2s} \quad \text{and} \quad \mathcal{E}_s(U^d, N_*)/N_*^{1+s/d} \le \underline{g}_{s,d} + (1-\gamma),$$

where $\underline{g}_{s,d} := \liminf_{N\to\infty} \mathcal{E}_s(U^d, N)/N^{1+s/d}$. Selecting m_t so that $(m_t-1)^dN_* < L_t \le m_t^dN_*$, we have from (19) and (20) that

$$\begin{aligned}\overline{g}_{s,d} &= \lim_{t\to\infty} \frac{\mathcal{E}_s(U^d, L_t)}{L_t^{1+s/d}} \le \limsup_{t\to\infty} \frac{\mathcal{E}_s(U^d, m_t^dN_*)}{[(m_t^*-1)^dN_*]^{1+s/d}} \\ &\le \limsup_{t\to\infty} \Big(\frac{m_t}{m_t-1}\Big)^{s+d} \Big(\gamma^{-s}\frac{\mathcal{E}_s(U^d, N_*)}{N_*^{1+s/d}} + K(1-\gamma)^{-s}N_*^{1-s/d}\Big) \\ &\le \gamma^{-s}[\underline{g}_{s,d} + (1-\gamma)] + K(1-\gamma)^s.\end{aligned}$$

Taking $\gamma \to 1$ gives $\overline{g}_{s,d} \le \underline{g}_{s,d}$, proving the existence of the limit (15) for $A = U^d$.

It remains to show that the limit $g_{s,d} := \overline{g}_{s,d} = \underline{g}_{s,d}$ is positive and finite. This follows for the upper estimate by considering the configurations $\omega_{m^d} = \frac{1}{m}Z_m^d$, $m = 2, 3, \ldots$, and for the lower

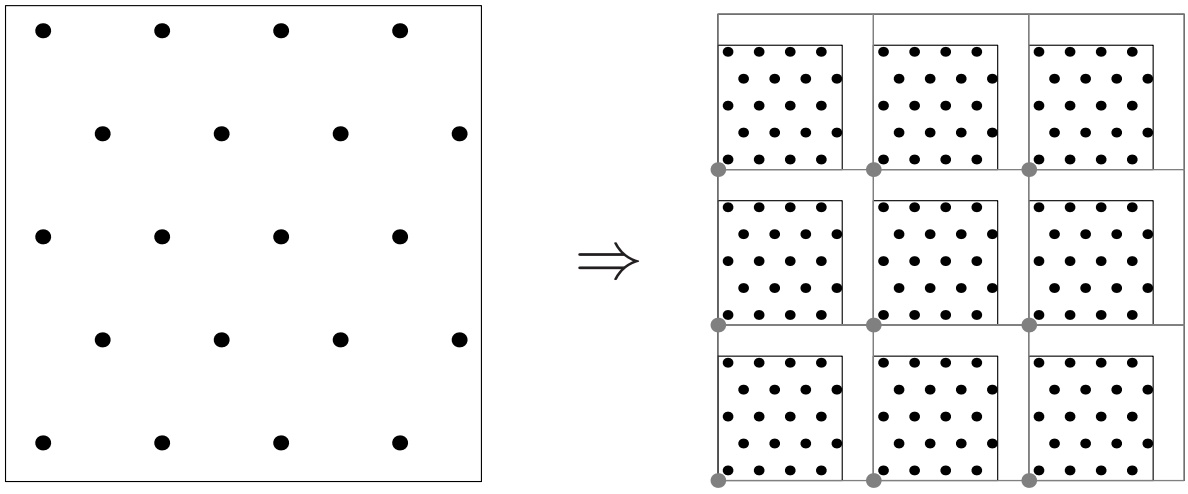


Figure 6. Scaling an optimal configuration of N points in the unit cube U^d to disjoint subcubes provides bounds for the minimal Riesz s-energy for m^dN points in terms of the minimal Riesz s-energy for N points.

estimate by using the convexity of $f(r) = r^{s/d}$ and the arithmetic-harmonic mean inequality.

From the Unit Cube to Compact Sets in $\mathbb{R}^d$. For $s > d$, we define

$$\mathcal{g}_{s,d}(A) := \lim_{N\to\infty} \mathcal{E}_s(A,N)/N^{1+s/d}$$

for any compact set A for which the limit exists. As we have seen, for $s > d$ it is the near neighbor terms that dominate the Riesz s-energy. If B and C are disjoint compact sets such that both $\mathcal{g}_{s,d}(B)$ and $\mathcal{g}_{s,d}(C)$ exist and $A = B \cup C$, then by neglecting terms in $\mathcal{E}_s(A,N)$ involving pairs of points $x \in B$ and $y \in C$ and considering optimal s-energy configurations for B and C separately, one can show that $\mathcal{g}_{s,d}(A)$ exists and satisfies

$$\mathcal{g}_{s,d}(A)^{-d/s} = \mathcal{g}_{s,d}(B)^{-d/s} + \mathcal{g}_{s,d}(C)^{-d/s}. \tag{21}$$

For an arbitrary compact set $A \subset \mathbb{R}^d$ we use finite unions of cubes with pairwise disjoint interiors to approximate A to show that $\mathcal{g}_{s,d}(A)$ exists and satisfies $\mathcal{g}_{s,d}(A) = \mathcal{g}_{s,d}(U^d)\mathcal{H}_d(A)^{-s/d}$ (the latter equation clearly holding for arbitrary cubes). The Besicovitch Covering Theorem and the Lebesgue Density Theorem play key roles in this argument. Letting $C_{s,d} = \mathcal{g}_{s,d}(U^d)$ gives (15) in the case A is a compact set in $\mathbb{R}^d$.

From Compact Sets in $\mathbb{R}^d$ to d-Rectifiable Manifolds in $\mathbb{R}^{d'}$. Clearly, if $A \subset \mathbb{R}^{d'}$ is *exactly* the isometric image of some compact set $B \subset \mathbb{R}^d$, then $\mathcal{g}_{s,d}(A)$ exists and equals

$$\mathcal{g}_{s,d}(B) = C_{s,d}\mathcal{H}_d(B)^{-s/d} = C_{s,d}\mathcal{H}_d(A)^{-s/d}.$$

Moreover, if A is the union of a finite collection of pairwise disjoint compact sets $\{K_i\}_{i=1}^L$ where each K_i is the isometric image of some compact set $B_i \subset \mathbb{R}^d$, then using (21) (which also holds when $d < d'$), one may conclude that $\mathcal{g}_{s,d}(A)$ exists and

$$\mathcal{g}_{s,d}(A) = C_{s,d}\left(\sum_{i=1}^{L}\mathcal{H}_d(K_i)\right)^{-s/d} = C_{s,d}\mathcal{H}_d(A)^{-s/d}.$$

Finally, we appeal to a basic result of Federer [8] that for every d-rectifiable manifold A, $\mathcal{H}_d$-almost all of A can be covered by a countable union of pairwise disjoint images of bi-Lipschitz mappings of compact sets in $\mathbb{R}^d$ with bi-Lipschitz constants uniformly close to 1. This completes an outline of the proof of Theorem 1.

Other Energy Functions. Although the Riesz s-energy is relevant to a variety of physical and mathematical problem areas, there are many other energy functionals that have significant application. For example, Benedetto and Fickus [3] show how to generate finite normalized tight frames by minimizing the energy

$$\sum_{i\neq j} f(|x_i - x_j|), \tag{22}$$

for N points on the sphere S^d, where $f(t) = t^4/4 - t^2$. J. Brauchart has shown that for a suitable function f depending on N, the minimum energy points on the sphere give spherical designs for cubature. A further recent example arises in the work of M. Atiyah and P. Sutcliffe [1], where the energy is minimized over $\mathbb{R}^3$ for a function of the form $\log 1/|D|$, where $D : C_N(\mathbb{R}^3) \to \mathbb{C}$ is a smooth function on the space of N distinct unordered points in $\mathbb{R}^3$. Du, Gunzburger, and Ju [6] consider point sets on a surface A obtained by minimizing the energy function

$$\sum_{i=1}^{N}\int_{x\in V_i}\rho(x)|x - x_i|^2\,dx,$$

where V_i denotes the Voronoi regions on A generated by $\{x_1,\dots,x_N\} \subset A$ and ρ is a weight function. Many of these energy functionals provide fertile ground for asymptotic analysis.

A natural generalization that is currently being investigated is that of the *weighted* Riesz s-energy problem:

$$\min_{\{x_i\}_{i=1}^N\subset A}\sum_{i\neq j}\frac{\rho(x_i,x_j)}{|x_i - x_j|^s},$$

where $\rho(x,y)$ is a given weight function on a d-dimensional manifold A. The methods described above can be adapted to show that under suitable conditions on the weight and manifold, the limit density (with respect to $\mathcal{H}_d$) for optimal s-energy configurations is a multiple of $\rho(x,x)^{-d/s}$ when $s > d$.

Acknowledgments. We thank R. Womersley for providing the images for Figures 3–5, M. Bowick for helpful discussions, and S. Borodachov for his useful comments.

Added in proof: The authors, together with S. Borodachov, have recently shown that Theorem 1 holds for a more general class of rectifiable sets.

References

[1] M. Atiyah and P. Sutcliffe, The geometry of point particles, *Royal Soc. London Proc. Ser. A Math. Phys. Eng. Sci.* **458** (2002), 1089–1115.

[2] A. R. Bausch, M. J. Bowick, A. Cacciuto, A. D. Dinsmore, M. F. Hsu, D. R. Nelson, M. G. Nikolaides, A. Travesset, and D. A. Weitz, Grain boundary scars and spherical crystallography, *Science* **299** (March 14, 2003), 1716–1718.

[3] J. Benedetto and M. Fickus, Finite normalized tight frames, *Adv. Comput. Math.* **18** (2003), 357–385.

[4] M. Bowick, D. R. Nelson, and A. Travesset, Interacting topological defects in frozen topographies, *Phys. Rev.* **B62** (2000), 8738–8751.

[5] M. Bowick, D. R. Nelson, and A. Travesset, Curvature-induced defect unbinding in toroidal geometries, *Phys. Rev. E* **69** (2004), 041102-(1–12).

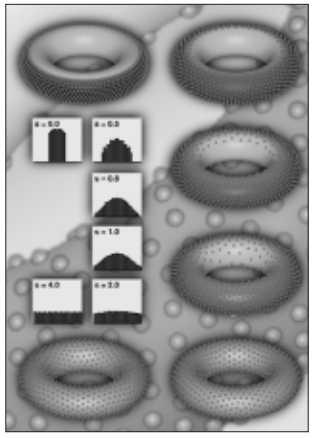
[6] Q. Du, M. D. Gunzburger, and L. Ju, Constrained centroidal Voronoi tessellations for surfaces, *SIAM J. Sci. Comput.* **24** (2003), 1488–1506.

[7] T. Erber and G. M. Hockney, Complex systems: equilibrium configurations of N equal charges on a sphere $2 \le N \le 112$, *Advances in chemical physics, Vol. XCVIII*, 495–594, Adv. Chem. Phys., XCVIII, Wiley, New York, 1997.
[8] H. Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin, 1969.
[9] D. P. Hardin and E. B. Saff, Minimal Riesz energy point configurations for rectifiable d-dimensional manifolds, to appear in *Adv. Math.*
[10] A. B. J. Kuijlaars and E. B. Saff, Asymptotics for minimal discrete energy on the sphere, *Trans. Amer. Math. Soc.* **350** (1998), 523–538.
[11] N. S. Landkof, *Foundations of Modern Potential Theory,* Springer-Verlag, Heidelberg, 1972.
[12] A. Martinez-Finkelshtein, V. V. Maymeskul, E. A. Rakhmanov, and E. B. Saff, Asymptotics for minimal discrete Riesz energy on curves in $\mathbf{R}^d$, *Canad. J. Math.* **56** (2004), 529–552.
[13] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, Cambridge University Press, Cambridge, 1995.
[14] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou, Minimal discrete energy on the sphere, *Math. Res. Lett.* **1** (1994), 647–662.
[15] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou, Electrons on the sphere, *Computational Methods and Function Theory* (CMFT '94) (R. M. Ali et al., eds.), World Scientific, 1995, pp. 293–309.
[16] E. B. Saff and A. B. J. Kuijlaars, Distributing many points on a sphere, *Math. Intelligencer* **19** (1997), 5–11.
[17] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Grundlehren series, Springer-Verlag, Heidelberg, 1997.
[18] M. Shub and S. Smale, Complexity of Bezout's theorem III: Condition number and packing, *J. Complexity* **9** (1993), 4–14.
[19] S. Smale, Mathematical problems for the next century, *Math. Intelligencer* **20** (1998), 7–15.
[20] Y. M. Zhou, *Arrangements of points on the sphere*, thesis, University of South Florida, 1995.

About the Cover

Approximating minimal Riesz

This month's cover originated with the article by Doug Hardin and Ed Saff in this issue. It shows 1,000 points on a torus distributed so as to minimize the total energy determined by an interaction through Riesz potentials $-\log r$ (labeled $s = 0$) as well as $1/r^s$ for 0.5, 0.8, 1, 2, and 4. As explained in the article, the limit distribution as the number of points becomes infinite is, remarkably, just the uniform distribution for all values of $s \ge 2$ (the dimension of the torus). This is also shown by the histograms on the cover, displaying the distribution in somewhat carefully selected radial bands. If the uniformity for $s = 2$ doesn't seem to be quite valid in the figure, that's because convergence to the equilibrium is very slow at the transitional s-value.



The data for all images, and indeed the actual images of all those blue & green bagels with red poppy seeds (as Saff refers to them), were produced by Rob Womersley of the University of New South Wales. Asked how he produced the pictures, Womersley said it had been "by a combination of local and global large-scale optimization techniques running on a Linux cluster to find close to minimum energy point sets, as well as visualization using Matlab." He added, "Numerical experiments such as these help illustrate theoretical results, but also suggest new results waiting to be proved."

—Bill Casselman
Graphics Editor
(notices-covers@ams.org)