1

The State of the Art in Smale's 7th Problem

Carlos Beltrán[a]

Depto. de Matemáticas, Estadística y Computación
Universidad de Cantabria

1.1 A very brief historical note

Smale's 7th problem is the computational version of an old problem dating back to Thomson [30] and Tammes [29], see Whyte's early review [32] for its history, namely, the sensible distribution of points in the two–dimensional sphere. In Whyte's paper different possible definitions of "well–distributed points in the sphere" are suggested:

1. Points which maximise the product of their mutual distances (called elliptic Fekete points [1] after [14]).
2. Points which minimise the sum of the inverse of their mutual distances (Thomson's problem), and more generally which minimise some sum of potentials which depend on the mutual distances (like Riesz potentials).
3. Points which maximise the least distance between any pair.
4. Points which are the center of the optimal packing problem, that is, the problem of finding the smallest radius of a sphere such that one can place on its surface k non–overlapping circles of a given radius.

This beautiful problem is terribly challenging! A first shocking result by Leech [19] showed that even though the set of N particles on the sphere which are critical points for the problem in item (2) for *every possible* potential can be completely described, this description is not enough to solve the problem for *any particular* potential. Namely, solving problem (2) for some particular potential may be completely meaningless for solving problem (2) for another, different potential. We quote Leech:

[a] Partially supported by MTM2010-16051 (Spanish Ministry of Science and Innovation MICINN).

[1] Not to be confused with the so called Fekete points.

There is no obvious way of relating the present problems to other extremal problems such as minimising the greatest distance at which an arbitrary point can be placed from the nearest point of a configuration. In fact, since a configuration which is not balanced is out of equilibrium under almost all laws of force, it is not to be expected that any such configuration will be found to be of significance in respect to both an equilibrium problem and another extremal problem, or even under two different significant equilibrium problems.

The problem has so many ramifications that it is difficult even to mention all of them. There are dozens of papers written about each of the mentioned problems. In this paper we focus only on the explicit version proposed by Smale [27]: the problem of finding elliptic Fekete points in the two–dimensional sphere. Our reference list will also refer only to this problem, and thus many important articles dealing with the other versions are omitted, for the sake of brevity.

1.2 The problem

Given N different points $x_1, \dots, x_N \in \mathbb{R}^3$, let $X = (x_1, \dots, x_N)$ and

$$\mathcal{E}(X) = \mathcal{E}(x_1, \dots, x_N) = -\sum_{i<j} \log \|x_i - x_j\|$$

be its *logarithmic potential* or *logarithmic energy* [2]. Let

$$\begin{aligned}\mathbb{S} &= \{(a,b,c) \in \mathbb{R}^3 : a^2 + b^2 + (c-1/2)^2 = 1/4\} \\ &= \{(a,b,c) \in \mathbb{R}^3 : a^2 + b^2 + c^2 = c\}\end{aligned}$$

be the Riemann sphere, i.e. the sphere in $\mathbb{R}^3$ of radius $1/2$ centered at $(0,0,1/2)^T$, and let

$$m_N = \min_{x_1,\dots,x_N \in \mathbb{S}} \mathcal{E}(x_1, \dots, x_N)$$

be the minimum value of $\mathcal{E}$. A minimising N–tuple $X = (x_1, \dots, x_N)$ is called a set of elliptic Fekete points. Note that such a N–tuple can also be defined as a set of N points in the sphere which *maximise the product of their mutual distances.*

Smale's 7th problem [27]: *Can one find* $X = (x_1, \dots, x_N)$ *such that*

$$\mathcal{E}(X) - m_N \leq c \log N, \qquad c \text{ a universal constant.} \tag{1.1}$$

[2] Sometimes $\mathcal{E}(X)$ is denoted by $\mathcal{E}_0(X)$, $\mathcal{E}(0,X)$ or $V_N(X)$

By "can one find" Smale means "can one describe an algorithm (in the BSS model of computation[3]) which on input N produces such a N–tuple, in running time bounded by a polynomial in N?".

Remark 1.1 The original question in [27] is not for points in $\mathbb{S}$ but for points in the unit sphere $\{(a,b,c) \in \mathbb{R}^3 : a^2+b^2+c^2 = 1\}$. We prefer to use the Riemann sphere instead of the unit sphere because some of the results look more natural when stated in $\mathbb{S}$. Another powerful reason to do so is that $-\log \|x-y\|$ is positive for every $x, y \in \mathbb{S}$ while the same claim is not true for x, y in the unit sphere. This helps intuition at some moments. Of course, the problems of finding a set of elliptic Fekete points in $\mathbb{S}$ or in the unit sphere are equivalent via the transformation

$$(a,b,c) \in \mathbb{S} \mapsto (2a, 2b, 2c-1).$$

If $x_1, \ldots, x_N \in \mathbb{S}$ and we denote $\hat{x}_1, \ldots, \hat{x}_N$ their associated unit sphere points via this transform, then we have:

$$\mathcal{E}(\hat{x}_1, \ldots, \hat{x}_N) = \mathcal{E}(x_1, \ldots, x_N) - \frac{\log 2}{2} N(N-1)$$

We thus state all the results for $\mathbb{S}$, translating them from their original citations when necessary.

1.3 The value of m_N

The first problem one encounters in dealing with Smale's 7th problem is that the value of m_N is not known, even to $O(N)$. A general technique (valid for Riemannian manifolds) given by Elkies shows that

$$m_N \geq \frac{N^2}{4} - \frac{N \log N}{4} + O(N).$$

Wagner [31] used the stereographic projection and Hadamard's inequality to get another lower bound. His method was refined by Rakhmanov, Saff and Zhou [21], who also proved an upper bound for m_N using partitions of the sphere. The lower bound was subsequently improved upon

[3] For the nonexpert reader, a BSS algorithm is just an algorithm in the natural sense of the word: a sequence of instructions (arithmetic operations, comparisons and, in general, any of the usual instructions present in a computer program) that, correctly executed, gives an answer. The arithmetic operations are assumed to be exact when performed on real numbers. See [9, 8] for details.

by Dubickas and Brauchart [13], [10]. The following result summarizes the best known bounds:

Theorem 1.2 *Let C_N be defined by*

$$m_N = \frac{N^2}{4} - \frac{N \log N}{4} + C_N N.$$

Then,

$$-0.4375 \leq \liminf_{N \mapsto \infty} C_N \leq \limsup_{N \mapsto \infty} C_N \leq -0.3700708...$$

1.4 The separation distance

The separation distance of a N–tuple $X = (x_1, \ldots, x_N) \in \mathbb{S}^N$ is defined by

$$d_{\|\cdot\|,sep}(X) = \min_{i \neq j} \|x_i - x_j\|.$$

By the definition of $\mathcal{E}$, it is clear that if X is a set of elliptic Fekete points then $d_{\|\cdot\|,sep}(X)$ cannot be too small. Using tools from classical potential theory, Rakhmanov, Saff and Zhou [22, 21] first proved the lower bound $3/(10\sqrt{N})$ for the separation distance of a set of elliptic Fekete points. Their result was improved by Dubickas [13] to $7/(8\sqrt{N})$. The sharpest known bound is due to Dragnev [11]:

Theorem 1.3 *Let X be a set of elliptic Fekete points. Then,*

$$d_{\|\cdot\|,sep}(X) \geq \frac{1}{\sqrt{N-1}}.$$

Recall that given two points $x, y \in \mathbb{S}$, the Riemannian distance $d_R(x, y)$ is the length of the shortest curve in $\mathbb{S}$ joining x and y. Elementary trigonometry shows that

$$d_R(x, y) = \arcsin \|x - y\|.$$

Thus, if we define

$$d_{R,sep}(X) = \min_{i \neq j} d_R(x_i, x_j),$$

we have

$$d_{R,sep}(X) \geq \arcsin \frac{1}{\sqrt{N-1}} \geq \frac{1}{\sqrt{N-1}},$$

for X as above.

1.5 The condition number of polynomials and Bombieri–Weyl norm

According to [27], one of Smale's motivations for studying the problem of elliptic Fekete points was to find polynomials all of whose zeros are well conditioned. Shub and Smale [24] defined a certain quantity (the "condition number") and used it to measure the stability and complexity of polynomial zero finding algorithms. Given a homogeneous polynomial

$$h(z,t)=\sum_{k=0}^{N} a_k z^k t^{N-k}, \qquad a_k\in\mathbb{C}, a_N\neq 0$$

of degree $N\geq 1$, and a projective zero $\zeta\in\mathbb{P}(\mathbb{C}^2)$ of h, the condition number of h at ζ is[4]

$$\mu(h,\zeta)=N^{1/2}\|(Dh(\zeta)\mid_{\zeta^\perp})^{-1}\|\|h\|\|\zeta\|^{N-1},$$

or $+\infty$ if $Dh(\zeta)\mid_{\zeta^\perp}$ is not an invertible mapping. Here, $Dh(\zeta)\mid_{\zeta^\perp}$ is the restriction of the derivative to the orthogonal complement of ζ in $\mathbb{C}^2$, and

$$\|h\|=\left(\sum_{k=0}^{N}\binom{N}{k}^{-1}|a_k|^2\right)^{1/2}$$

is the Bombieri–Weyl norm (sometimes called the Kostlan norm) of h. If no zero of h is specified, we just take the maximum:

$$\mu(h)=\max_{\zeta\in\mathbb{P}(\mathbb{C}^2):h(\zeta)=0}\mu(h,\zeta).$$

Now, let f be a univariate polynomial

$$f(X)=\sum_{k=0}^{N} a_k X^k, \qquad a_k\in\mathbb{C}, a_N\neq 0,$$

and let $z\in\mathbb{C}$ be a zero of f. We define

$$\mu(f,z)=\mu(h,(z,1)), \qquad \mu(f)=\max_{z\in\mathbb{C}:f(z)=0}\mu(f,z),$$

where $h(X,Y)=\sum_{k=0}^{N} a_k X^k Y^{N-k}$ is the homogeneous counterpart of f. Taking $\|f\|=\|h\|$, one can write:

$$\mu(f,z)=\frac{N^{1/2}(1+|z|^2)^{\frac{N-2}{2}}}{|f'(z)|}\|f\|.$$

[4] Sometimes μ is denoted μ_{norm} or μ_{proj} but we here keep the simpler notation.

In [26] Shub and Smale proved the following relation between the condition number and elliptic Fekete points. Let $\mathfrak{Re}$ and $\mathfrak{Im}$ be, respectively, the real and complex part of a complex number.

Theorem 1.4 *Let $z_1, \dots, z_N \in \mathbb{C}$ be a set of complex numbers. For $1 \leq i \leq N$, let $x_i \in \mathbb{S}$ be the preimage of z_i under the stereographic projection, that is*

$$x_i = \left(\frac{\mathfrak{Re}(z_i)}{1+|z_i|^2}, \frac{\mathfrak{Im}(z_i)}{1+|z_i|^2}, \frac{1}{1+|z_i|^2} \right)^T \in \mathbb{S}, \qquad 1 \leq i \leq N. \tag{1.2}$$

Assume that $x_1, \dots, x_N$ are a set of elliptic Fekete points. Let $f : \mathbb{C} \to \mathbb{C}$ be a degree N polynomial such that its zeros are $z_1, \dots, z_N$. Then,

$$\mu(f) \leq \sqrt{N(N+1)}.$$

More generally, let $z_1, \dots, z_N \in \mathbb{C}$ be any collection of N distinct complex numbers, let f be a polynomial with zeros $z_1, \dots, z_N$ and let $x_1, \dots, x_N$ be given by (1.2)*. Then,*

$$\mu(f) \leq \sqrt{N(N+1)} \frac{e^{\mathcal{E}(x_1,\dots,x_N)}}{e^{m_N}}.$$

It is interesting to remark that there exists no explicit known way of describing a sequence of polynomials satisfying $\mu(f) \leq N^c$, for any fixed constant c and $N \geq 1$. Theorem 1.4 says that, if a N–tuple satisfying (1.1) can be described for any N, then such a sequence of polynomials can also be generated.

Here is a nice formula (which just follows from the definitions) relating $\mathcal{E}$ to μ and Bombieri–Weyl norm:

$$\mathcal{E}(x_1, \dots, x_N) = \frac{1}{2} \sum_{i=1}^{N} \log \mu(f, z_i) + \frac{N}{2} \log \frac{\prod_{i=1}^{N} \sqrt{1+|z_i|^2}}{\|f\|} - \frac{N}{4} \log N.$$

Note that the term

$$\frac{\prod_{i=1}^{N} \sqrt{1+|z_i|^2}}{\|f\|} \tag{1.3}$$

in the previous formula is the quotient between the product of the Bombieri–Weyl norm of the factors of f and the Bombieri–Weyl norm of f. That quantity is always greater than 1, see [3]. Experiments suggest that minimising $\mathcal{E}$ is a problem similar to minimising the sum of $\log \mu(f, z_i)$, and to maximising the quotient (1.3)[5]. We recall from [3,

[5] This may seem surprising at a first glance. It turns out that (1.3) is minimal, i.e., equal to 1, precisely when all the z_i are equal, which implies $\mathcal{E}(x_1, \dots, x_N) = \infty$.

Theorem 2.1] (see also [4]) that for two polynomials f, g of respective degrees r and s,

$$\|f \cdot g\| \geq \sqrt{\frac{r!s!}{(r+s)!}}\, \|f\| \cdot \|g\|, \tag{1.4}$$

and this bound is optimal. Maximising (1.3) would be solved by finding an analogue of (1.4) for products of N polynomials, a nice mathematical problem in its own right. As pointed out in [4], it follows from (1.4) that

$$\frac{\prod_{i=1}^{N} \sqrt{1+|z_i|^2}}{\|f\|} \leq \sqrt{N!},$$

but this inequality is far from optimal (for example, if the $z_i's$ are the Nth roots of unity, then the value of this quotient is $\sqrt{2^{N-1}} \ll \sqrt{N!}$).

1.6 The average value and random polynomials

Random polynomials[6] have been known, since [25], to be well–conditioned on average, meaning that their condition number is polynomially bounded by their degree, on average. This, combined with Theorem 1.4, suggests that spherical points associated with zeros of random polynomials should produce small values of $\mathcal{E}$. To properly state this fact, let us consider $\mathcal{E}$ as a function defined on $\mathbb{S}^N \setminus \Sigma$ where

$$\Sigma = \{(x_1, \ldots, x_N) \in \mathbb{S}^N : x_i = x_j \text{ for some } i \neq j\}.$$

Note that Σ is the set of N–tuples $(x_1, \ldots, x_N)$ of points in $\mathbb{S}$ such that the polynomial f whose zeros are associated with the x_i satisfies $\mu(f) = \infty$.

Here and throughout this paper, for a measure space X, a measurable, finite volume subset $U \subseteq X$, and a measurable function $f : U \to [0, \infty)$ we set

$$\fint_U f = \frac{\int_U f}{\int_U 1} = \frac{1}{\text{Volume}(U)} \int_U f.$$

One can easily compute the average value of $\mathcal{E}$ when $x_1, \ldots, x_N$ are chosen at random in $\mathbb{S}$, uniformly and independently with respect to the probability distribution induced by Lebesgue measure in $\mathbb{S}$:

$$\fint_{x \in \mathbb{S}^N \setminus \Sigma} \mathcal{E}(X) = \frac{N^2}{4} - \frac{N}{4}.$$

[6] in the sense of Theorem 1.5 below.

By comparing this with Theorem 1.2, we can see that random choices of points in the sphere already produce pretty low values of the minimal energy. One can ask if other (simple) probability distributions produce even lower average results. In [2] we proved a relation with random polynomials.

Theorem 1.5 *Let $f(X) = \sum_{k=0}^{N} a_k X^k$ be a random polynomial where the a_k are independent complex random variables, such that the real and imaginary parts of a_k are independent (real) Gaussian random variables centered at 0 with variance $\binom{N}{k}$. Let $z_1, \ldots, z_N$ be the complex zeros of f, and let x_i be given by* (1.2)*. Then, the expected value of $\mathcal{E}(x_1, \ldots, x_N)$ equals*

$$\frac{N^2}{4} - \frac{N \log N}{4} - \frac{N}{4}.$$

Again, by comparing this with Theorem 1.2, we conclude that spherical points coming from zeros of random polynomials are pretty well distributed, as they agree with the minimal value of $\mathcal{E}$, to order $O(N)$. This result fits into a more general (yet, less precise) kind of result related to random sections on Riemann surfaces, see [33, 34]. Note that the notion of "random polynomial" used in Theorem 1.5 is strongly related to the Bombieri–Weyl norm. It is the natural Gaussian distribution associated with the space of polynomials, considered as a normed vector space with the Bombieri–Weyl norm.

1.7 Properties of the critical points of $\mathcal{E}$

One of the first things to do when faced with an optimization problem is to study the critical points of the objective function, i.e. the points where the derivative vanishes. In our case the derivative of $\mathcal{E}$ is easy to compute. Algebraic manipulation of its expression was used in [7, 12] to get the following [7]:

Theorem 1.6 *Let $x_1, \ldots, x_N \in \mathbb{S}$ be a critical point of $\mathcal{E}$. Let*

$$\mathfrak{o} = \left(0, 0, \frac{1}{2}\right)^T$$

be the center of $\mathbb{S}$. Then,

[7] In [12] the result is stated for global minima, but the proof is indeed valid for any critical point of $\mathcal{E}$. Moreover, the third item in this theorem is here stated in greater generality than in [12], but the same proof holds.

- *The center of mass of the x_i is $\mathfrak{o}$. Namely,*

$$\sum_{i=1}^{N} \overrightarrow{\mathfrak{o}x_i} = \mathfrak{o}, \qquad \text{or equivalently} \qquad \frac{1}{N}\sum_{i=1}^{N} x_i = \mathfrak{o}.$$

- *For every $1 \leq i \leq N$, we have:*

$$\sum_{j \neq i} \frac{\overrightarrow{x_j x_i}}{\|x_j - x_i\|^2} = 2(N-1)\overrightarrow{\mathfrak{o}x_i}.$$

- *For every $x \in \mathbb{S}$, we have:*

$$\sum_{i=1}^{N} \|x - x_i\|^2 = \frac{N}{2}.$$

Other natural questions are, which types of critical points does $\mathcal{E}$ have and how many of them exist? There are some conjectures about their number (some authors conjecture that the number of local minima grows exponentially on N, see the references in Section 1.12 below) but no precise result is known. It was pointed out in [26] that there exist critical points of $\mathcal{E}$ of index N, namely N points evenly distributed on some equator of $\mathbb{S}$. It follows from (1.5) below and the maximum principle of harmonic analysis that no local maximum of $\mathcal{E}$ can exist.

1.8 Harmonic properties of $\mathcal{E}$

Let us endow $\mathbb{S}^N$ with its natural Riemannian structure, that is the product structure (or equivalently, the structure inherited from $\mathbb{R}^{3N}$). Again viewing $\mathcal{E}$ as a function $\mathcal{E} : \mathbb{S}^N \setminus \Sigma \to \mathbb{R}$, we computed in [5] its (Riemannian) Laplacian. It turns out that

$$\Delta \mathcal{E} \equiv 2N(N-1), \tag{1.5}$$

is a constant. If a function defined on an open set of $\mathbb{R}^n$ has a constant Laplacian, then the classical mean value theorem of harmonic analysis gives a formula for the mean value of the function on a ball centered at every point. In the case of $\mathbb{S}^N$, one can use the theory of harmonic manifolds to analyze the mean value of $\mathcal{E}$ in products of spherical caps, that is in sets of the form

$$\begin{aligned} B_\infty(X, \vec{\varepsilon}) &= \{(y_1, \ldots, y_N) \in \mathbb{S}^N : d_R(x_i, y_i) < \varepsilon_i, 1 \leq i \leq N\} \\ &= B(x_1, \varepsilon_1) \times \cdots \times B(x_N, \varepsilon_N) \subseteq \mathbb{S}^N, \end{aligned}$$

where $X = (x_1, \dots, x_N)$, $\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)$ and for $x \in \mathbb{S}$, $\varepsilon > 0$, $B(x, \varepsilon)$ is the open spherical cap of (Riemannian) radius equal to ε. Abusing notation, if $\varepsilon = 0$ we define $B(x, 0) = \{x\}$. The mean value of $\mathcal{E}$ in $B_\infty(X, \vec{\varepsilon})$ was studied in [5]:

Theorem 1.7 *Let $X \in \mathbb{S}^N \setminus \Sigma$ and $\vec{\varepsilon} \in [0, \pi/2)^N$ be such that $B_\infty(X, \vec{\varepsilon}) \subseteq \mathbb{S}^N \setminus \Sigma$. Then,*

$$\fint_{B_\infty(X,\vec{\varepsilon})} \mathcal{E}(Y)\, dY = \mathcal{E}(X) + C_N(\vec{\varepsilon}),$$

where

$$C_N(\vec{\varepsilon}) = (N-1) \sum_{j=1}^{N} \left(\frac{1}{2} + \frac{\log(\cos \varepsilon_j)}{\tan^2 \varepsilon_j} \right) \in \left[0, \frac{N-1}{2} \right),$$

with the convention that

$$\frac{1}{2} + \frac{\log(\cos 0)}{\tan^2 0} = 0.$$

The reader may find useful the estimate $\frac{1}{2} + \frac{\log(\cos \varepsilon)}{\tan^2 \varepsilon} \approx \frac{\varepsilon^2}{4}$ for small values of ε.

1.9 The limiting distribution

It follows from classical potential theory that optimal logarithmic energy points are uniformly distributed over $\mathbb{S}$, asymptotically as $N \mapsto \infty$, in the following sense: Let $\{X^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)})\}$ be a sequence such that $X^{(N)} \in \mathbb{S}^N$ is a set of N elliptic Fekete points in $\mathbb{S}$ for every $N \geq 2$. Then, for any continuous function $f : \mathbb{S} \to \mathbb{R}$ we have:

$$\fint_{\mathbb{S}} f = \lim_{N \mapsto \infty} \frac{1}{N} \sum_{j=1}^{N} f(x_j^{(N)}). \tag{1.6}$$

One way to analyze this qualitative result is to study the so called spherical cap discrepancy, that is for fixed $N \geq 2$:

$$D_C(X^{(N)}) = \sup_{\mathcal{C}} \left| \frac{\sharp(X^{(N)} \cap \mathcal{C})}{N} - \fint_{\mathbb{S}} \chi_{\mathcal{C}} \right|,$$

where $\chi_{\mathcal{C}}$ is the characteristic function of $\mathcal{C}$ and the supremum is taken over all possible spherical caps $\mathcal{C}$ in $\mathbb{S}$. Note that $D_C(X^{(N)})$ measures how far the counting measure is from the probability measure associated with Lebesgue measure in $\mathbb{S}$. In [10], Brauchart proved the following estimate: