Five Point Energy Minimization: A Summary

Richard Evan Schwartz

July 1, 2018

Abstract

This paper is a condensation of my monograph [**S0**], which contains a complete proof that there is a constant ש ≈ 15.0488 such that the triangular bi-pyramid is the minimizer, amongst all 5 point configurations on the sphere, with respect to the power law potential $R_s(r) = \text{sign}(s)/r^s$, if and only if $s \in (-2, 0) \cup (0,$ ש$]$. In this paper we explain the main ideas and give proofs for some of the key lemmas.

1 Introduction

Let S^2 denote the unit sphere in $\boldsymbol{R}^3$ and let $X = \{\widehat{p}_1, ..., \widehat{p}_n\}$ be a finite list of distinct points on S^2. Given some function $f : (0, 2] \to \boldsymbol{R}$ we can compute the total f-potential

$$\mathcal{E}_f(X) = \sum_{i<j} f(\|\widehat{p}_i - \widehat{p}_j\|). \tag{1}$$

For fixed f and n, one can ask which configuration(s) X minimize $\mathcal{E}_f(X)$. This problem generally goes under the heading of *Thomson's Problem*, and some form of it originates in [**Th**].

For this problem, the energy functional $f = R_s$, where

$$R_s(r) = \text{sign}(s) r^{-s}, \tag{2}$$

is a natural one to consider. When $s > 0$, this is called the *Riesz potential*. When $s < 0$ this is called the *Fejes-Toth potential*. The case $s = 1$ is specially called the *Coulomb potential* or the *electrostatic potential*.

There is a large literature on the energy minimization problem. See [**Fö**] and [**C**] for some early local results. See [**MKS**] for a definitive numerical study on the minimizers of the Riesz potential for $n \leq 16$. See [**SK**] and [**RSZ**] for results with an emphasis on the case when n is large. See [**BBCGKS**] for a survey of results about the higher dimensional case. For $n = 4, 6, 12$ it is known that the set of vertices of the corresponding platonic solid is the global minimizer with respect to any Riesz potential. See [**KY**], [**A**], [**Y**]. See also [**CK**] for a vast generalization.

The case $n = 5$ has been notoriously intractable. There is a general feeling that for a wide range of energy choices including the Riesz potentials, the global minimizer is either the triangular bi-pyramid [1] (TBP) or else some pyramid with square base – an FP. Here is a resume of results.

- The paper [**HS**] has a rigorous computer-assisted proof that the TBP is the unique minimizer for R_{-1}.

- My paper [**S1**] has a rigorous computer-assisted proof that the TBP is the unique minimizer for R_1 and R_2.

- The paper [**DLT**] gives a traditional proof that the TBP is the unique minimizer for the logarithmic potential.

- In [**BHS**, Theorem 7] it is shown that, as $s \to \infty$, any sequence of 5-point minimizers w.r.t. R_s must converge (up to rotations) to the FP made from 5 vertices of the regular octahedron. So, the TBP is not a minimizer w.r.t R_s when s is very large.

- In 1977, T. W. Melnyk, O. Knop, and W. R. Smith, [**MKS**] conjectured the existence of the phase transition constant, around $s = 15.04808$, at which point the TBP ceases to be the minimizer w.r.t. R_s.

- In [**T**], A. Tumanov gives a traditional proof of the following result. Define
$$G_k(r) = (4 - r^2)^k, \qquad k = 1, 2, 3, ... \tag{3}$$
Let $f = a_1 G_1 + a_2 G_2$ with $a_1, a_2 > 0$. The TBP is the unique global minimizer with respect to f. Moreover, a critical point of f must be the TBP In particular, the TBP is a minimizer for G_2.

[1]This is the configuration isometric to the one consisting of the north pole, the south pole, and three points placed on the equator in an equilateral triangle.

In my unpublished monograph [**S0**], I prove the following result.

Theorem 1.1 (Main) *Let* $s_0 = 15 + 25/512$*. There exists a computable number*

$$\text{ש} = 15.0480773927797... < s_0$$

with the following properties:

- *For* $s \in (-2,0) \cup (0,\text{ש})$ *the TBP is the unique minimizer w.r.t.* R_s*.*
- *For* $s = \text{ש}$*, the TBP and some FP are the two minimizers w.r.t.* R_s*.*
- *For each* $s \in (\text{ש}, s_0]$*, some FP is the unique minimizer w.r.t* R_s*.*
- *For* $s \geq s_0$ *the TBP is not the global minimizer w.r.t.* R_s*.*

The purpose of this paper is to give a condensed account of the work in [**S0**], concentrating on the overall strategy and providing proofs of, or at least insight into, the key technical lemmas. §2 of this paper gives an outline of the proof and then subsequence chapters fill in the outline.

I would like to thank Ed Saff for suggesting that I write this condensation.

2 Outline of the Proof

2.1 Interpolation

We generally consider two configurations the same if they are isometric, and as we go on we will normalize our configurations in particular ways. Given the function f and a 5-point configuration X we often write $f(X)$ in place of $\mathcal{E}_f(X)$ for ease of notation. Let R_s be as in the introduction.

Let $I \subset \boldsymbol{R}$ denote an interval, which we think of an interval of power law exponents. We say that a triple $(\Gamma_2, \Gamma_3, \Gamma_4)$ of potentials is *forcing* on the interval I if the following implication holds for any 5-point configuration X:

$$\Gamma_i(X) > \Gamma_i(\boldsymbol{T}) \quad i = 2,3,4 \qquad \Longrightarrow \qquad R_s(X) > R_s(\boldsymbol{T}) \quad \forall s \in I. \tag{4}$$

Recall that $G_k(r) = (4 - r^2)^k$. In the paper [**T**], A. Tumanov observes that (G_2, G_3, G_5) is forcing on $(-2,0)$ and $(0,2]$. He does not supply a proof for his observation, but we will take up a related idea. Consider the following functions:

$$\begin{array}{rcl} G_5^\flat & = & G_5 - 25G_1 \\ G_{10}^\# & = & G_{10} + 13G_5 + 68G_2 \\ G_{10}^{\#\#} & = & G_{10} + 28G_5 + 102G_2 \end{array} \tag{5}$$

In the §3 we will sketch the ideas behind a proof of the following result.

Lemma 2.1 (Forcing) *The following is true.*

1. (G_2, G_3, G_5) *is forcing on* $(-2, 0)$.
2. (G_2, G_4, G_6) *is forcing on* $(0, 6]$.
3. $(G_2, G_5, G_{10}^{\#\#})$ *is forcing on* $[6, 13]$.
4. $(G_2, G_5^\flat, G_{10}^\#)$ *is forcing on* $[13, 15.05]$.

We don't bother proving that (G_2, G_3, G_5) is forcing on $(0, 2]$ because we don't need this fact. Results like the Forcing Lemma are used in many papers on energy minimization – e.g. [**CK**]. See [**BDHSS**] for very general ideas like this. I found these triples through computer experimentation, taking Tumanov's observation as a starting point. One other guide to the experimets is that the TBP is not the minimizer for G_k when $k = 7, 8, 9...$, so one might look at more complicated combinations of these functions.

2.2 Divide and Conquer

As we explain in §4, we use a divide and conquer algorithm, together with a certain energy estimate, to give a rigorous computer-assisted proof of

Theorem 2.2 (Big) *The TBP is the unique global minimizer w.r.t.*

$$G_3, G_4, G_5^\flat, G_6, G_{10}^{\#\#}.$$

We prove the energy estimate in §5.

We note that G_5 is a positive combination of G_1 and $G_5^\flat$, so G_5 also satisfies the conclusion of the Big Theorem. The Big Theorem and the Forcing Lemma combine to prove the Main Theorem for all exponents $s \in (-2, 0) \cup (0, 13]$. To finish the proof, we just have to deal with the exponent interval $(13, 15 + 25/512]$. The TBP is not the global minimizer for $G_{10}^\#$, but we can still squeeze information out of this function.

Stereographic projection is the map

$$\Sigma(x,y,z) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right). \tag{6}$$

The domain of this map is $S^2 - \{(0,0,1)\}$. Here $(0,0,1)$ is the north pole. We rotate our configurations $\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3, \widehat{p}_4$ so that $\widehat{p}_4 = (0,0,1)$. We then have the correspondence

$$\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3, (0,0,1) \in S^2 \iff p_0, p_1, p_2, p_3 \in \boldsymbol{R}^2, \qquad p_k = \Sigma(\widehat{p}_k). \tag{7}$$

We write $p_k = (p_{k1}, p_{k2})$. We always normalize so that p_0 lies in the positive x-axis and $\|p_0\| \geq \|p_k\|$ for $k = 1,2,3$. Let **SMALL** denote those 5-point configurations which are represented by 4-tuples $\{p_0, p_1, p_2, p_3\}$ such that

1. $\|p_0\| \geq \|p_k\|$ for $k = 1,2,3$.
2. $512p_0 \in [433, 498] \times [0, 0]$.
3. $512p_1 \in [-16, 16] \times [-464, -349]$.
4. $512p_2 \in [-498, -400] \times [0, 24]$.
5. $512p_3 \in [-16, 16] \times [349, 364]$.

This domain is pretty tight. We tried hard to get as far away from the TBP configuration as possible. Figure 2.1 shows a picture of the sets corresponding to the definition of **SMALL**. The grey circle is the unit circle. Note that $\boldsymbol{T} \not\in$ **SMALL**. The 4 black dots and the 4 white dots are the two nearby normalized TBP configurations. As we discuss in §4, we prove the following companion to the Big Theorem.

Theorem 2.3 (Small) *Let X be some 5-point configuration. Suppose that we have $G^{\#}_{10}(X) \leq G^{\#}_{10}(\boldsymbol{T})$. Then either $X = \boldsymbol{T}$ or $X \in$* **SMALL**.

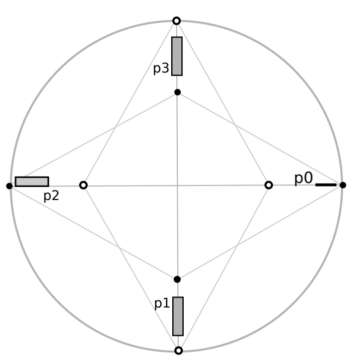

Figure 2.1: The sets defining **SMALL**.

2.3 Symmetrization

Let **K4** denote the 2-dimensional set of configurations whose stereographic projections are rhombi with points in the coordinate axes. (Here **K4** stands for "Klein 4 symmetry".) The TBP and all FPs have normalizations which lie in **K4**.

Let

$$\mathbf{SMALL4} = \mathbf{K4} \cap \mathbf{SMALL}. \tag{8}$$

This is the set of configurations $X = (p_0, p_1, p_2, p_3)$ such that

$$-p_2 = p_0 = (x, 0), \quad -p_1 = p_3 = (0, y), \quad x \geq y, \qquad x, y \in \left[\frac{348}{512}, \frac{495}{512}\right].$$

The point $(x, y) = (1, 1/\sqrt{3})$, outside our slice, represents the TBP.

We consider the following map from **SMALL** to **K4**. Starting with X, we let (p_1', p_2', p_3', p_4') be the configuration which is obtained by rotating X about the origin so that $p_{02}' = p_{22}'$ and $p_{21}' < p_{01}'$. We then define

$$-p_2^* = p_0^* = \left(\frac{p_{01}' - p_{21}'}{2}, 0\right), \qquad -p_3^* = p_1^* = \left(0, \frac{p_{12}' - p_{32}'}{2}\right) \tag{9}$$

The points $(p_0^*, p_1^*, p_2^*, p_3^*)$ define the symmetrized configuration X^*. In §6 we give some details about the following crucial result.

Lemma 2.4 (Symmetrization) *Let $s \in [12, 15+25/512]$ and suppose that $X \in$ **SMALL**. Then $R_s(X^*) \leq R_s(X)$ with equality iff $X = X^*$.*

One strange thing is that the map $X \to X^*$ is not clearly related to spherical geometry. Rather, it is a linear projection with respect to the stereographic coordinates we impose on the moduli space. I found the map $X \to X^*$ experimentally, after trying many alternatives. Combining the Symmetrization Lemma and the Small Theorem, we get the following result:

Corollary 2.5 *Let $s \in [13, 15 + 25/512]$. Suppose that $R_s(X) \leq R_s(\boldsymbol{T})$ and $X \neq \boldsymbol{T}$. Then $X \in$ **SMALL4**.*

This corollary practically finishes the proofs of the Main Theorem. It leaves us with the exploration of a 2-dimensional rectangle in the configuration space. We refer the reader to [**S0**] for a discussion of the endgame.

3 Interpolation

3.1 The General Approach

Let Γ_0 be the constant function and let $\Gamma_1 = G_1$. Let $\boldsymbol{T}$ be the TBP. We have $\Gamma_0(X) = \Gamma_0(\boldsymbol{T})$ and $\Gamma_1(X) \geq \Gamma_1(\boldsymbol{T})$ for all 5-point configurations X. Indeed, as is well known, the minimizers for Γ_1 are precisely those configurations whose center of mass is the origin. This derives from the formula

$$E_{\Gamma_1}(X) = \sum_{i<j} \|X_i + X_j\|^2 = C + \Big\| \sum_i X_i \Big\|^2,$$

where the constant C does not depend on X.

The distances involved in $\boldsymbol{T}$ are $\sqrt{2}, \sqrt{3}, \sqrt{4}$. (Writing $\sqrt{4}$ for 2 makes the equations look nicer.) Let R be some function defined on $(0, 2]$. Suppose that $(\Gamma_2, \Gamma_3, \Gamma_4)$ is one of the triples from the Forcing Lemma. Suppose we can find a combination

$$\Gamma = a_0\Gamma_0 + ... + a_4\Gamma_4, \qquad a_1, a_2, a_3, a_4 > 0. \tag{10}$$

such that

$$\Gamma(x) \leq R(x), \qquad \Gamma(\sqrt{m}) = f(\sqrt{m}), \qquad m = 2, 3, 4. \tag{11}$$

Suppose also that $\Gamma_j(X) \geq \Gamma_j(\boldsymbol{T})$ for $j = 2, 3, 4$, with strict inequality for at least one index. Then

$$R(X) \geq \Gamma(X) = \sum a_i\Gamma_i(X) > \sum a_i\Gamma_i(\boldsymbol{T}) = \Gamma(\boldsymbol{T}) = R(\boldsymbol{T}). \tag{12}$$

If the above conditions are satisfied we call our triple *good* for R.

Here is how we find the coefficients $\{a_i\}$. We impose the 5 conditions

- $\Gamma(x) = R(x)$ for $x = \sqrt{2}, \sqrt{3}, \sqrt{4}$.
- $\Gamma'(x) = R'(x)$ for $x = \sqrt{2}, \sqrt{3}$.

Here $R' = dR/dx$ and $\Gamma' = d\Gamma/dx$. These 5 conditions give us 5 linear equations in 5 unknowns. In the cases described below, the associated matrix is invertible and there is a unique solution. The proof of the Forcing Lemma, in each case, involves solving the matrix equation, checking positivity of the coefficients, and checking the under-approximation property.

In this chapter we will present solutions to the matrix equation and some discussion of the work in [**S0**, §3-6], which establishes all the claims about our triples. The interested reader can check our claims numerically in a symbolic manipulator such as Mathematica [**W**]. [**S0**, §3] has numerical plots as well.

3.2 The Matrix Solutions

For the triple is (G_2, G_3, G_5). The solution is

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ \delta \end{bmatrix} = \frac{1}{144} \begin{bmatrix} 0 & 0 & -144 & 0 & 0 & 0 \\ -312 & -96 & 408 & 24 & 80 & 0 \\ 684 & -288 & -396 & -54 & -144 & 0 \\ -402 & 264 & 138 & 33 & 68 & 0 \\ 30 & -24 & -6 & -3 & -4 & 0 \\ 2496 & 768 & -3264 & -192 & -640 & -144 \end{bmatrix} \begin{bmatrix} 2^{-s/2} \\ 3^{-s/2} \\ 4^{-s/2} \\ s2^{-s/2} \\ s3^{-s/2}, \\ s4^{-s/2} \end{bmatrix}.$$

It turns out that the triple is good for R_s when $s \in (-2, 0]$.

For the triple G_2, G_4, G_6 the solution is given by

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ \delta \end{bmatrix} = \frac{1}{792} \begin{bmatrix} 0 & 0 & 792 & 0 & 0 & 0 \\ 792 & 1152 & -1944 & -54 & -288 & 0 \\ -1254 & -96 & 1350 & 87 & 376 & 0 \\ 528 & -312 & -216 & -39 & -98 & 0 \\ -66 & 48 & 18 & 6 & 10 & 0 \\ -6336 & -9216 & 15552 & 432 & 2304 & 792 \end{bmatrix} \begin{bmatrix} 2^{-s/2} \\ 3^{-s/2} \\ 4^{-s/2} \\ s2^{-s/2} \\ s3^{-s/2}, \\ s4^{-s/2} \end{bmatrix}.$$

For the triple $G_2, G_5, G_{10}^{\#\#}$ we have

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \widehat{a}_4 \\ \delta \end{bmatrix} = \frac{1}{268536} \begin{bmatrix} 0 & 0 & 268536 & 0 & 0 & 0 \\ 88440 & 503040 & -591480 & -4254 & -65728 & 0 \\ -77586 & -249648 & 327234 & 2361 & 65896 & 0 \\ 41808 & -19440 & -22368 & -2430 & -9076 & 0 \\ -402 & 264 & 138 & 33 & 68 & 0 \\ -707520 & -4024320 & 4731840 & 34032 & 525824 & 268536 \end{bmatrix} \begin{bmatrix} 2^{-s/2} \\ 3^{-s/2} \\ 4^{-s/2} \\ s2^{-s/2} \\ s3^{-s/2}, \\ s4^{-s/2} \end{bmatrix}.$$

For the triple $G_2, G_5^\flat, G_{10}^{\#}$ we have

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \widehat{a}_4 \\ \delta \end{bmatrix} = \frac{1}{268536} \begin{bmatrix} 0 & 0 & 268536 & 0 & 0 & 0 \\ 947112 & 131520 & -1078632 & -50694 & -259072 & 0 \\ -91254 & -240672 & 331926 & 3483 & 68208 & 0 \\ 35778 & -15480 & -20298 & -1935 & -8056 & 0 \\ -402 & 264 & 138 & 33 & 68 & 0 \\ 174268608 & 24199680 & -198468288 & -9327696 & -47669248 & 268536 \end{bmatrix} \begin{bmatrix} 2^{-s/2} \\ 3^{-s/2} \\ 4^{-s/2} \\ s2^{-s/2} \\ s3^{-s/2}, \\ s4^{-s/2} \end{bmatrix}.$$

In all cases, we are also keeping track of

$$\delta = 2\Gamma'(2) - 2R'(2). \tag{13}$$

This quantity is also positive on the relevant intervals. We will use the positivity of δ in our under-approximation proof below.

Even though the expressions involved above are not polynomials in general, we establish the various claims by showing that certain polynomials are positive on certain domains. There are a variety of computational positivity certificates for polynomials; I will discuss one that I thought of myself and have used on several occasions in other work. We first discuss the general approach, and then we turn the technique towards the verification of the claims in this section.

3.3 Positive Dominance

Here we discuss some positivity certificates for polynomials in one and several variables. See [**S0**, Chapter 5] and also [**S2**], [**S3**] for the proofs of the lemmas in this section.

Given a multi-index $I = (i_1, ..., i_k) \in (\boldsymbol{N} \cup \{0\})^k$ we let

$$x^I = x_1^{i_1} ... x_k^{i_k}. \tag{14}$$

Any polynomial $F \in \boldsymbol{R}[x_1, ..., x_k]$ can be written succinctly as

$$F = \sum a_I X^I, \qquad a_I \in \boldsymbol{R}. \tag{15}$$

If $I' = (i'_1, ..., i'_k)$ we write $I' \leq I$ if $i'_j \leq i_j$ for all $j = 1, ..., k$. We call F *weak positive dominant* (WPD) if

$$A_I := \sum_{I' \leq I} a_{I'} \geq 0 \qquad \forall I, \tag{16}$$

and if the sum of all the coefficients is positive. We call f *positive dominant* (PD) if we have strict inequality in Equation 16 for all indices.

Lemma 3.1 (Positivity Criterion) *If P is PD, then $P > 0$ on $[0, 1]^k$. If P is WPD then $P > 0$ on $(0, 1]^k$.*

For the application we have in mind, Lemma 3.1 is not that useful, because the polynomials of interest to us are not WPD or PD in general. However, Lemma 3.1 feeds into a powerful divide-and-conquer algorithm. We define the maps

$$A_{j,1}(x_1, ..., x_k) = (x_1, ..., x_{j-1} \frac{x_j + 0}{2}, x_{i+1}, ..., x_k),$$

$$A_{j,2}(x_1, ..., x_k) = (x_1, ..., x_{j-1} \frac{x_j + 1}{2}, x_{j+1}, ..., x_k), \tag{17}$$

We define the jth *subdivision* of P to be the set

$$\{P_{j1}, P_{j2}\} = \{P \circ A_{j,1}, P \circ A_{j,2}\}. \tag{18}$$

Lemma 3.2 *$P > 0$ on $[0, 1]^k$ if and only if $P_{j1} > 0$ and $P_{j2} > 0$ on $[0, 1]^k$.*

Say that a *marker* is a non-negative integer vector in $\boldsymbol{R}^k$. Say that the *youngest entry* in the the marker is the first minimum entry going from left to right. The *successor* of a marker is the marker obtained by adding one to the youngest entry. For instance, the successor of $(2,2,1,1,1)$ is $(2,2,2,1,1)$. Let μ_+ denote the successor of μ.

We say that a *marked polynomial* is a pair (P,μ), where P is a polynomial and μ is a marker. Let j be the position of the youngest entry of μ. We define the *subdivision* of (P,μ) to be the pair

$$\{(P_{j1},\mu_+),(P_{j2},\mu_-)\}. \tag{19}$$

Geometrically, we are cutting the domain in half along the longest side, and using a particular rule to break ties when they occur. Now we have assembled the ingredients needed to explain the Positive Dominance Algorithm.

1. Start with a list LIST of marked polynomials. Initially, LIST consists only of the marked polynomial $(P,(0,...,0))$.

2. Let (Q,μ) be the last element of LIST. We delete (Q,μ) from LIST and test whether Q is positive dominant.

3. Suppose Q is positive dominant. We go back to Step 2 if LIST is not empty. Otherwise, we halt.

4. Suppose Q is not positive dominant. we append to LIST the two marked polynomials in the subdivision of (Q,μ) and then go to Step 2.

If the algorithm halts, it constitutes a proof that $P>0$ on $[0,1]^k$ or $(0,1]^k$. For the case of strict positivity, the algorithm halts if and only if $P>0$ on $[0,1]^k$. There is also a parallel version whose halting constitutes a proof that $\max(P_1,...,P_m)>0$, where $P_1,...,P_m$ is a finite list of polynomials. In the parallel version we pass a block if one of the functions is PD (or WPD) on it.

Remark: Below we will be somewhat imprecise about whether we use the WPD version of the algorithm of the PD version. These fine details are taken care of in [**S0**, §4-6].

3.4 Positivity of the Coefficients

The coefficients from the previous section all have the form

$$A(s) = \lambda_2 2^{-s/2} + \lambda_3 3^{-s/2} + \lambda_4 4^{-s/2} + \mu_2 s 2^{-s/2} + \mu_3 s 3^{-s/2} + \mu_4 s 4^{-s/2} \tag{20}$$

for $\lambda_i, \mu_i \in Q$ for all $i = 2, 3, 4$. We wish to show that such an expression is positive on some interval $I \subset (-2, 16]$. We break the interval $(-2, 16]$ into into the intervals $(-2, -1]$, $[-1, 0]$, $[0, 1]$, etc., and consider the problem separately on each such interval.

Let I be one of these unit intervals. Let k be the even endpoint of I. For each $m \in \{2, 3, 4\}$ we have Taylor's Theorem with Remainder:

$$m^{-s/2} = \sum_{j=0}^{11} \frac{(-1)^j \log(m)^j}{m^k 2^j j!}(s - 2k)^j + \frac{E_s}{12!}(s - 2k)^{12}. \tag{21}$$

Here E_s is the 12th derivative of $m^{-s/2}$ evaluated at some point in the interval. (Going up to 12 derivatives is a somewhat arbitrary but convenient cutoff.) We replace the powers of $\log(m)$ by nearby rational approximations in order to get close rational under-approximations and rational over-approximations to $m^{-s/2}$ on I. We then feed these approximations back into Equation 20 to get a rational polynomial $\underline{A}(s) \leq A(s)$ on I. We then take the relevant sub-interval $J \subset I$ and use the Positive Dominance Algorithm to show that $A(s) > 0$ on J. Normally we have $J = I$ but in a Case 4 of the Forcing Lemma we sometimes have $J = [15, 15 + \frac{25}{512}]$ and $I = [15, 16]$.

3.5 Under Approximation

Let Γ be as in Equation 10. Let Ω be the exponent interval of interest to us. For instance, in Case 1, $\Omega = (0, 6]$. We want to show that $\Gamma(r) \leq R_s(r)$ for all $r \in (0, 2]$ and all $s \in \Omega$. Define

$$H(r, s) = 1 - \frac{\Gamma(r)}{R_s(r)} = 1 - r^s \Gamma, \tag{22}$$

We just have to show that $H \geq 0$ on $(0, 2) \times \Omega$. In Cases 1 and 2 the following lemma, has a traditional proof that exploits the low degree of the polynomials involved. The proof we sketch here is an industrial strength proof that works (one case at a time) for all the cases.

Lemma 3.3 *For each fixed value $s \in \Omega$, the function $\partial_r H(r,s)$ has 4 roots in $(0,2)$, and these roots are all simple.*

Proof: (Sketch.) The positive roots of $\partial_r H(r,s)$ are the same as the positive roots of the polynomial $\psi(r,s) = s\Gamma(r) + r\Gamma'(r)$. To show that ψ has simple roots, we just have to show that $\psi(r,s)$ and $\partial_r\psi(r,s)$ do not simultaneously vanish for $(r,s) \in (0,2] \times \Omega$. Considered as polynomials in r, the coefficients are functions of s having the same form as in Equation 20.

We work separately with each unit integer interval I that intersects Ω. Using the same technique as in the previous section we produce rational 2-variable polynomials $\underline{u}, \overline{u}, \underline{v}, \overline{v}$ such that

$$\underline{u}(r,s) \leq \psi(r,s) \leq \overline{u}(r,s), \qquad \underline{v}(r,s) \leq \partial_r\psi(r,s) \leq \overline{v}(r,s)$$

for all $(r,s) \in (0,2] \times I$. We then apply the parallel version of the Positive Dominance Algorithm to the set of functions $\{\underline{u}, -\overline{u}, \underline{v}, -\overline{v}\}$. This time we simply work with I. There is no need to restrict to a sub-interval $J \subset I$ even in Case 4. The algorithm halts, and this constitutes a proof that $\psi(r,s)$ and $\partial_r\psi(r,s)$ do not both vanish at the same point $(r,s) \in (0,2] \times I$.

Now we turn to the number of roots. We first observe that $\partial_r H(r,s) < 0$ for $r > 0$ sufficiently small, and also $\partial_r H(2,s) > 0$. The second fact comes from the fact that the quantity δ in Equation 13 is positive for all $s \in \Omega$. We will abbreviate these two conditions as the *asymptotic conditions*.

The asymptotic conditions imply that $\partial_r H(r,s)$ has an even number of simple roots for $r \in (0,2)$. The asymptotic conditions also imply the following phenomenon: As s changes, the number of roots of $\partial_r H(r,s)$ can change only if 2 non-real and conjugate roots converge to the interval $(0,2)$. But this would give $\partial_r H(r,s)$ a double root for some s. In other words, the asymptotic conditions imply that the number of simple roots is independent of s. We then check in each case, for a single value of s, that this number is 4. ♠

Lemma 3.4 $\partial_r^2 H(r,s) > 0$ *for $r = \sqrt{2}$ and $r = \sqrt{3}$.*

Proof: We check this in each case for a single value of s. If the condition ever failed for some other value of s, then there would be a value of s such that $\partial_r^2 H(r,s) = 0$. But then $\partial_r H(r,s)$ would not have all simple roots. ♠

Lemma 3.4 combines with the asymptotic conditions and Lemma 3.3 to show that $H(r,s) \geq 0$ for all $(r,s) \in (0,2] \times I$.

4 Divide and Conquer

4.1 Normalized Configurations

Recall that $\widehat{p}_0, ..., \widehat{p}_4$ is our configuration and $p_k = \Sigma(\widehat{p}_k)$ for $k = 0, 1, 2, 3$. As usual $\widehat{p}_4 = (0, 0, 1)$. Setting $p_k = (p_{k1}, p_{k2})$, we normalize so that

- p_0 lies in the positive x-axis. That is, $p_{01} > 0$ and $p_{02} = 0$.
- $\|p_0\| \geq \max(\|p_1\|, \|p_2\|, \|p_3\|)$.
- $p_{12} \leq p_{22} \leq p_{32}$ and $0 \leq p_{22}$.

If, in addition, we are working with a monotone decreasing energy function, we also require that $p_{12} \leq 0$. If this condition fails then our normalization implies that all the points lie in the same hemisphere. But then one can decrease the energy by reflecting one of the points across the hemisphere boundary. (This argument can fail for the one energy function we consider which is not monotone, namely $G_5^\natural$.) We call such configurations *normalized*.

Lemma 4.1 *Let Γ be any of $G_3, G_4, G_5^\natural, G_5, G_6, G_{10}^\#, G_{10}^{\#\#}$. Any normalized minimizer w.r.t. Γ has $p_0 \in [0, 4]$ and $\|p_k\| \leq 3/2$ for $k = 1, 2, 3$.*

Proof: The TBP has 6 bonds of length $\sqrt{2}$, and 3 bonds of length $\sqrt{3}$, and one bond of length 2. Hence,

$$G_k(\boldsymbol{T}) = 3(2^{k+1} + 1), \qquad k = 1, 2, 3, ... \tag{23}$$

Suppose that Γ is one of the energies above, but not G_3 or $G_5^\natural$. If $\|p_0\| \geq 4$ then the distance from $\widehat{p}_0$ to $(0, 0, 1)$ is at most $d = \sqrt{4/17}$. We check this by computing $\Sigma^{-1}(4, 0)$, which would be the farthest point from $(0, 0, 1)$ with the given constraints. We check by direct calculation that $\Gamma(d) > \Gamma(\boldsymbol{T})$ in all cases. This single bond contributes too much to the energy all by itself. If the second condition fails then we have $\|p_k\| > 3/2$ for some k and also $\|p_0\| > 3/2$. The distance from $\widehat{p}_k$ and $\widehat{p}_0$ from $\widehat{p}_4$ is at most $d' = 4/\sqrt{13}$. In all cases we compute that $2\Gamma(d') > \Gamma(\boldsymbol{T})$. For example $G_4(\boldsymbol{T}) = 99$ whereas $2\Gamma(d') \approx 117.616$.

This deals with all cases except G_3 and $G_5^\natural$. See [**S0**] for the more elaborate argument that deals with these two cases. ♠

4.2 The TBP Configurations

The TBP has two kinds of points, the two at the poles and the three at the equator. When ∞ is a polar point, the points p_0, p_1, p_2, p_3 are, after suitable permutation,

$$(1,0), \qquad (-1/2, -\sqrt{3}/2), \qquad (0,0), \qquad (-1/2, +\sqrt{3}/2). \tag{24}$$

We call this the *polar configuration*. When ∞ is an equatorial point, the points p_0, p_1, p_2, p_3 are, after suitable permutation,

$$(1,0), \qquad (0, -1/\sqrt{3}), \qquad (-1,0), \qquad (0, 1/\sqrt{3}). \tag{25}$$

We call this the *equatorial configuration*. We can visualize the two configurations together in relation to the regular 6-sided star. The black points are part of the polar configuration and the white points are part of the equatorial configuration. The grey point belongs to both configurations. The points represented by little squares are polar and the points represented by little disks are equatorial.

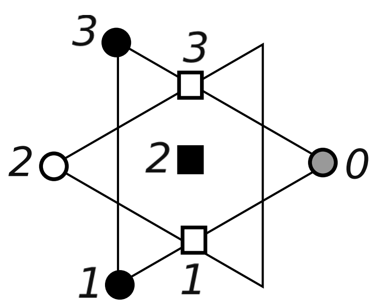

Figure 4.1: Polar and equatorial versions of the TBP.

For each $k = 0, ..., 4$ we introduce the quantity

$$\delta_k = \min_i \widehat{p}_i \cdot \widehat{p}_k. \tag{26}$$

We say that a normalized configuration is *totally normalized* if

$$\delta_4 \geq \delta_k, \qquad k = 0,1,2,3. \tag{27}$$

This condition is saying that the points in the configuration are bunched up around $\widehat{p}_4 = (0,0,1)$ as much as possible. The polar TBP has $\delta_4 = -1$ and 3 values of k for which $\delta_k = 0$. Hence, it is not totally normalized. The equatorial TBP is totally normalized.

4.3 Dyadic Blocks

For the moment we find it convenient to only require that $p_k \in [-2,2]^2$ for $k = 1,2,3$. Later on, we will enforce the stronger condition given by Lemma 4.1. Define

$$\square = [0,4] \times [-2,2]^2 \times [-2,2]^2 \times [-2,2]^2. \tag{28}$$

Any minimizer of any of the energies we consider is isometric to one which is represented by a point in this cube. This cube is our universe.

In 1 dimension, the *dyadic subdivision* of a line segment is the union of the two segments obtained by cutting it in half. In 2 dimensions, the *dyadic subdivision* of a square is the union of the 4 quarters that result in cutting the the square in half along both directions. We say that a *dyadic segment* is any segment obtained from $[0,4]$ by applying dyadic subdivision recursively. We say that a *dyadic square* is any square obtained from $[-2,2]^2$ by applying dyadic subdivision recursively. We count $[0,4]$ as a dyadic segment and $[-2,2]^2$ as a dyadic square.

Hat and Hull Notation: We let $\langle X \rangle$ denote the convex hull of any Euclidean subset. Thus, we think of a dyadic square Q as the set of its 4 vertices and we think of $\langle Q \rangle$ as the solid square having Q as its vertex set. For any $S \subset \boldsymbol{R}^2$ we let $\widehat{S} = \Sigma^{-1}(S)$, where Σ is stereographic projection. In particular,

- $\widehat{Q}$ is a set of 4 co-circular points on S^2.
- $\langle \widehat{Q} \rangle$ is a convex quadrilateral whose vertices are $\widehat{Q}$.
- $\widehat{\langle Q \rangle}$ is a "spherical patch" on S^2, bounded by 4 circular arcs.

Good Squares: A dyadic square is *good* if it is contained in $[-3/2, 3/2]^2$ and has side length at most $1/2$. Note that a good dyadic square cannot cross the coordinate axes. The only dyadic square which crosses the coordinate axes is $[-2,2]^2$, and this square is not good. Our computer program only does spherical geometry calculations on good squares.

Dyadic Blocks: We define a *dyadic block* to be a 4-tuple (Q_0, Q_1, Q_2, Q_3), where Q_0 is a dyadic segment and Q_i is a dyadic square for $j = 1,2,3$. We say that a block is *good* if each of its 3 component squares is good. By Lemma

4.1, any energy minimizer for G_k is contained in a good block. Our algorithm quickly chops up the blocks in $\square$ so that only good ones are considered.

The product

$$\langle B\rangle = \langle Q_0\rangle \times \langle Q_1\rangle \times \langle Q_2\rangle \times \langle Q_3\rangle \tag{29}$$

is a rectangular solid in the configuration space $\square$. On the other hand, the product

$$B = Q_0 \times Q_1 \times Q_2 \times Q_3 \tag{30}$$

is the collection of 128 vertices of $\langle B\rangle$. We call these the *vertex configurations* of the block.

Definition: We say that a configuration p_0, p_1, p_2, p_3 is *in* the block B if $p_i \in \langle Q_i\rangle$ for $i = 0, 1, 2, 3$. In other words, the point in $\square$ representing our configuration is contained in $\langle B\rangle$. Sometimes we will say that this configuration is *associated to* the block.

Sudvidision of Blocks: There are 4 obvious subdivision operations we can perform on a block.

- The operation S_0 divides B into the two blocks (Q_{00}, Q_1, Q_2, Q_3) and (Q_{01}, Q_1, Q_2, Q_3). Here (Q_{00}, Q_{01}) is the dyadic subdivision of Q_0.
- the operation S_1 divides B into the 4 blocks (Q_0, Q_{1ab}, Q_2, Q_3), where $(Q_{100}, Q_{101}, Q_{110}, Q_{111})$ is the dyadic subdivision of Q_1.

The operations S_2 and S_3 are similar to S_1. These subdivision operations will feed into a subdivision algorithm akin to the one discussed in §3.3.

4.4 Spherical Geometry Estimates

We introduce some geometric quantities in this section.

- Let $d(Q)$ be the diameter $\langle \widehat{Q}\rangle$.
- Let $d_1(Q)$be the length of the longest edge of $\langle \widehat{Q}\rangle$.
- Let $D_Q \subset \boldsymbol{R}^2$ denote the disk containing Q in its boundary and $d_2(Q)$ be the diameter of $\widehat{D}_Q$.

When Q is a dyadic segment, we define $\delta(Q) = \chi(2, d_2)$. When Q is a good dyadic square, We define

$$\delta(Q) = \max\Big(\chi(1, d_1), \chi(2, d_2)\Big), \qquad \chi(D, d) = \frac{d^2}{4D} + \frac{d^4}{2D^3} \tag{31}$$

We call $\delta(Q)$ the *Hull approximation constant* of Q.

Lemma 4.2 (Hull Approximation) *Let Q be a dyadic segment or a good dyadic square. Every point of the spherical patch $\widehat{\langle Q\rangle}$ is within $\delta(Q)$ of a point of the convex quadrilateral $\langle \widehat{Q}\rangle$.*

Proof: (Sketch.) The basic idea behind many of our estimates is that the set $\widehat{\langle Q\rangle}$, which is a spherical patch on S^2, is quite close to the convex hull of the vertices of this set, namely $\langle \widehat{Q}\rangle$. See [**S0**, §8] for more details. ♠

Let Q be a dyadic segment or a good dyadic square. Let δ be the hull approximation constant of Q. Let $\{q_i\}$ be the points of Q. We make all the same definitions for a second dyadic square Q'. We define

$$(Q \cdot Q')_{\max} = \max_{i,j}(\widehat{q}_i \cdot \widehat{q}'_j) + \delta + \delta' + \delta\delta'. \tag{32}$$

$$(Q \cdot \{\infty\})_{\max} = \max_i \ \widehat{q}_i \cdot (0, 0, 1) \tag{33}$$

We define the same quantities with respect to *min*, except that in the first equation we subtract $\delta + \delta' + \delta\delta'$. Let $\Omega(Q)$ denote the union of line segments that connect points in $\widehat{\langle Q\rangle}$ to a nearest point in $\langle \widehat{Q}\rangle$. Note that $\Omega(Q)$ contains $\widehat{\langle Q\rangle}$. We define $\Omega(Q')$ similarly.

Lemma 4.3 (Dot Product Bounds) *For all $V, V' \in \Omega(Q) \times \Omega(Q')$,*

$$(Q \cdot Q')_{\min} \leq V \cdot V' \leq (Q \cdot Q')_{\max}$$

Proof: (Sketch.) The dot product is bilinear, and so the restriction of the dot product to the convex polyhedral set $\langle \widehat{Q}\rangle \times \langle \widehat{Q}'\rangle$ takes on its extrema at vertices. The rest of the proof comes down to several applications of the Cauchy-Schwarz Inequality, the Hull Approximation Lemma, and the triangle inequality. See [**S0**, §8] for more details. ♠

Let $B = (Q_0, Q_1, Q_2, Q_3)$ be a good block. As usual $Q_4 = \{\infty\}$. For each index j we define

$$(B,k)_{\min} = \min_{j \neq k}(Q_j \cdot Q_k)_{\min}, \qquad (B,k)_{\max} = \min_{j \neq k}(Q_j \cdot Q_k)_{\max}. \tag{34}$$

We say that B is *disordered* if there is some $k \in \{0,1,2,3\}$ such that

$$(B,4)_{\max} < (B,k)_{\min} \tag{35}$$

The following result is an immediate consequence of Lemma 4.3 and the definition given in Equation 26.

Lemma 4.4 *If B is disordered, then B contains no totally normalized configurations.*

4.5 Irrelevant Blocks

Call a block *irrelevant* if no configuration in the interior of the block is totally normalized. Call a block *relevant* if it is not irrelevant. Every relevant configuration in the boundary of an irrelevant block is also in the boundary of a relevant block. So, to prove the Big and Small Theorems, we can ignore the irrelevant blocks. Let $\overline{Q}_{jk}$ and $\underline{Q}_{jk}$ denote the maximum and minimum kth coordinate of a point in Q_j. Call B *certifiably irrelevant* if B satisfies at least one of the following conditions.

1. $\min(|\underline{Q}_{k1}|, |\overline{Q}_{k1}|) \geq \overline{Q}_{01}$ for some $k = 1,2,3$.
2. $\min(|\underline{Q}_{k2}|, |\overline{Q}_{k2}|) \geq \overline{Q}_{01}$ for some $k = 1,2,3$.
3. $\underline{Q}_{12} \geq 0$, provided we have a monotone decreasing energy.
4. $\overline{Q}_{22} \leq 0$.
5. $\overline{Q}_{32} \leq \underline{Q}_{22}$.
6. $\overline{Q}_{22} \leq \underline{Q}_{12}$.
7. B is disordered.

The reason for our terminology is that the above conditions give a computational test for showing that B is irrelevant.

Lemma 4.5 *If B is good and certifiably irrelevant, then B is irrelevant.*

Proof: Conditions 1 and 2 each imply that there is some index $k \in \{1,2,3\}$ such that all points in the interior of B_k are farther from the origin than all points in B_0. Condition 3 implies that all points in the interior of B_1 lie above the x-axis. This violates our normalization when the energy function is monotone decreasing. Condition 4 implies that all points in the interior of B_2 lie below the x-axis. Condition 5 implies that all points in the interior of B_3 lie below all points in the interior of B_2. Condition 6 implies that all points in the interior of B_2 lie below all points in the interior of B_1. Lemma 4.4 takes care of Condition 7. ♠

4.6 The Energy Theorem

We think of the energy potential $G = G_k$ as being a function on $(\boldsymbol{R}^2 \times \infty)^2$, via the identification $p \leftrightarrow \widehat{p}$. We take $k \geq 1$ to be an integer.

Let $\mathcal{Q}$ denote the set of dyadic squares $[-2,2]^2$ together with the dyadic segments in $[0,4]$, together with $\{\infty\}$. When $Q = \{\infty\}$ the corresponding hull approximation constant and the hull diameter are 0.

Now we are going to define a function $\epsilon : \mathcal{Q} \times \mathcal{Q} \to [0,\infty)$. First of all, for notational convenience we set $\epsilon(Q,Q) = 0$ for all Q, When $Q, Q' \in \mathcal{Q}$ are unequal, we define

$$\epsilon(Q,Q') = \frac{1}{2}k(k-1)T^{k-2}d^2 + 2kT^{k-1}\delta \tag{36}$$

Here

- d is the diameter of $\widehat{Q}$.
- $\delta = \delta(Q)$ is the hull approximation constant for Q.
- $T = T(Q,Q') = 2 + 2(Q \cdot Q')_{\max}$.

This is a rational function in the coordinates of Q and Q'. The quantities d^2 and δ are essentially quadratic in the side-lengths of Q and Q'. Note that we have $\epsilon(\{\infty\},Q') = 0$ but $\epsilon(Q,\{\infty\})$ is nonzero when $Q \neq \{\infty\}$.

Let $B = (Q_0, Q_1, Q_2, Q_3)$. For notational convenience we set $Q_4 = \{\infty\}$. We define

$$\mathbf{ERR}_k(B) = \sum_{i=0}^{3}\sum_{j=0}^{4} \epsilon(Q_i, Q_j). \tag{37}$$

This is a rational function of the vertices of B.

Theorem 4.6 (Energy)

$$\min_{v\in\langle B\rangle} \mathcal{E}_k(v) > \min_{v\in B} \mathcal{E}_k(v) - \mathbf{ERR}_k(B).$$

Remark: As the form of Equation 37 suggests, a similar result holds for N point configurations. Our argument in §5 readily covers this more general case.

Theorem 4.6 suffices to deal with G_3, G_4, G_6, but we need a more general result to deal with $G_5^\flat$, $G_{10}^\#$ and $G_{10}^{\#\#}$. Suppose we have some energy of the form

$$F = \sum_{k=1}^{N} a_k G_k \tag{38}$$

where $a_1, ..., a_N$ is some sequence of numbers, not necessarily positive.

Suppressing the dependence on F, we define

$$\epsilon(Q_i, Q_j) = \sum |a_k| \ \epsilon_k(Q_i, Q_j), \tag{39}$$

where $\epsilon_k(Q_i, Q_j)$ is the above quantity computed with respect to G_k. We then define **ERR** exactly as in Equation 37. With this definition, we have

Corollary 4.7 *Suppose that F is as in Equation 38. Suppose that B is a block such that*

$$\min_{v\in B} \mathcal{E}_F(v) - \mathbf{ERR}_F(B) > \mathcal{E}_F(\mathrm{TBP}). \tag{40}$$

Then all configurations in B have higher energy than the TBP.

We can write

$$\mathbf{ERR}(B) = \sum_{i=0}^{3} \mathbf{ERR}_i(B), \qquad \mathbf{ERR}_i(B) = \sum_{j=0}^{4} \epsilon(Q_i, Q_j). \tag{41}$$

We define the *subdivision recommendation* to be the index $i \in \{0, 1, 2, 3\}$ for which $\mathbf{ERR}_i(B)$ is maximal. A tie never arises in practice, but in the event of a tie we would pick the smaller index.

4.7 Estimates on the Hessian

The Energy Theorem is too crude to deal with configurations very near the TBP. Since we are considering totally normalized configurations we only have to worry about a neighborhood of the equatorial TBP. Let $B_0 \subset \square$ denote the cube of side-length 2^{-17} centered at the configuration representating the equatorial TBP. We consider our energy functions on the configuration space $\square$. We set

$$\Gamma_k = G_k \circ \Sigma^{-1}. \tag{42}$$

Here Σ is stereographic projection. The fact that Σ^{-1} is a rational function means that all the functions of interest to us are rational functions on $\square$. In [**S0**, §11] we prove the following result.

Lemma 4.8 (Hessian) *For each* $\Gamma = \Gamma_3, \Gamma_3, \Gamma_4, \Gamma_5, \Gamma_5^\flat, \Gamma_{10}^\sharp, \Gamma_{10}^{\sharp\sharp}$, *the Hessian of* Γ *is positive definite at all points of* B_0.

The basic idea is to get a global upper bounds on the 8th partial derivatives of Γ which hold on all of B_0 using crude combinatorial information about the function. The bounds we get are on the order of 2^{90}. We then essentially use Taylor's Theorem with Remainder, evaluating at the point representing the equatorial TBP, to bound the lower derivatives. There is nothing sacred about the cutoff at the 8 derivative; this is just what worked.

4.8 The Main Algorithm

We fix some energy F that appears in either the Big Theorem of the Small Theorem. Our divide and conquer algorithm works schematically like the one described in §3.3, except that we don't use the Positivity Criterion from Lemma 3.1 and we determine the subdivision order differently. Rather, we perform the following test to each block $B = (Q_0, Q_1, Q_2, Q_3)$.

1. If B lies entirely in the region **SMALL** and we are working with the function $G_{10}^\sharp$ we pass B. This is the extra step for the Small Theorem.

2. If some component square Q_i of B has side length more than $1/2$ we fail B and recommend that B be subdivided along the first such index. This step guarantees that we only pass good blocks.

3. If B is certifiably irrelevant, we pass B.

4. If we compute that $Q_i \not\subset [-3/2, 3/2]^2$ for some $i = 1, 2, 3$, we pass B. Given Step 2, we know that Q_i is disjoint from $(-3/2, 3/2)^2$.
5. If $B \subset B_0$, the small cube from the Hessian Lemma, we pass B.
6. If B satisfies Corollary 4.7, we pass B. Otherwise, we fail B and subdivide the block according to the subdivision recommendation.

In [**S0**, §10] we explain how we implement our calculations using interval arithmetic so as to avoid round-off error. We also detail the results of the calculations. In short, everything runs to completion successfully and the whole calculation takes about 12 hours on a modern MacBook Pro. This gives a rigorous computer-assisted proof of the Big Theorem and the Small Theorem.

5 Proof of the Energy Theorem

5.1 A Polynomial Inequality

In this chapter we prove Theorem 4.6. For the (dis)interested reader, we remark that the discussion of symmetrization in the next chapter does not depend on the work here. Theorem 4.6 builds on the case $M = 4$ of the following inequality.

Lemma 5.1 *Let $M \geq 2$ and $k = 1, 2, 3...$ Suppose*

- $0 \leq x_1 \leq ... \leq x_M$.
- $\sum_{i=1}^M \lambda_i = 1$ *and* $\lambda_i \geq 0$ *for all* i.

Then

$$0 \leq \sum_{i=1}^M \lambda_i x_i^k - \Big(\sum_{I=1}^M \lambda_i x_i\Big)^k \leq \frac{1}{8} k(k-1) x_M^{k-2}\ (x_M - x_1)^2. \qquad (43)$$

The lower bound is a trivial consequence of convexity, and both bounds are trivial when $k = 1$. So, we take $k = 2, 3, 4, ...$ and prove the upper bound. I discovered Lemma 5.1 experimentally. My motivation was to get an expression that varied quadratically regardless of the exponent k.

Suppose first that $M \geq 3$. We have one degree of freedom when we keep $\sum \lambda_i x_i$ constant and try to vary $\{\lambda_j\}$ so as to maximize the left hand side of the inequality. The right hand side does not change when we do this, and the left hand side varies linearly. Hence, the left hand size is maximized when $\lambda_i = 0$ for some i. But then any counterexample to the lemma for $M \geq 3$ gives rise to a counter example for $M - 1$. Hence, it suffices to prove the inequality when $M = 2$.

In the case $M = 2$, we set $a = \lambda_1$. Both sides of the inequality in Lemma 5.1 are homogeneous of degree k, so it suffices to consider the case when $x_2 = 1$. We set $x = x_1$. The inequality is then $f(x) \leq g(x)$, where

$$f(x) = (ax^k + 1 - a) - (ax + 1 - a)^k; \qquad g(x) = \frac{1}{8}k(k-1)(1-x)^2. \quad (44)$$

This is supposed to hold for all $a, x \in [0,1]$.

We [2] will show that Equation 44 holds for all $a, x \in [0,1]$ and all $k \geq 2$. We think of f as a function of x, with a held fixed. Since $f(1) = g(1) = 1$, it suffices to prove that $f'(x) \geq g'(x)$ on $[0,1]$. Define

$$\phi(x) = akx^{k-1}, \qquad b = (1-a)(1-x). \quad (45)$$

We have

$$-f'(x) = \phi(x+b) - \phi(x). \quad (46)$$

Both x and $x + b$ lie in $[0,1]$. So, by the mean value theorem there is some $y \in [0,1]$ so that

$$\frac{\phi(x+b) - \phi(x)}{b} = \phi'(y) = ak(k-1)y^{k-2}. \quad (47)$$

Hence

$$-f'(x) = b\phi'(y) = a(1-a)k(k-1)(1-x)y^{k-2} \quad (48)$$

But $a(1-a) \in [0, 1/4]$ and $y^{k-2} \in [0,1]$. Hence

$$-f'(x) \leq \frac{1}{4}k(k-1)(1-x) = -g'(x). \quad (49)$$

Hence $f'(x) \geq g'(x)$ for all $x \in [0,1]$.

[2] I am grateful to C. McMullen, who supplied this argument after I told him about the inequality

5.2 The Local Energy Lemma

Let $Q = \{q_1, q_2, q_3, q_4\}$ be the vertex set of $Q \in \mathcal{Q}$. We allow for the degenerate case that Q is a line segment or $\{\infty\}$. In this case we just list the vertices multiple times, for notational convenience.

Note that every point in the convex quadrilateral $\langle \widehat{Q} \rangle$ is a convex average of the vertices. For each $z \in \langle Q \rangle$, there is a some point $z^* \in \langle \widehat{Q} \rangle$ which is as close as possible to $\widehat{z} \in \widehat{\langle Q \rangle}$. There are constants $\lambda_i(z)$ such that

$$z^* = \sum_{i=1}^{4} \lambda_i(z)\ \widehat{q}_j, \qquad \sum_{i=1}^{4} \lambda_i(z) = 1. \tag{50}$$

We think of the 4 functions $\{\lambda_i\}$ as a partition of unity on $\langle Q \rangle$. The choices above might not be unique, but we make such choices once and for all for each Q. We call the assignment $Q \to \{\lambda_i\}$ the *stereographic weighting system*.

Lemma 5.2 (Local Energy) *Let ϵ be the function defined in the Theorem 4.6. Let Q, Q' be distinct members of $\mathcal{Q}$. Given any $z \in Q$ and $z' \in Q'$,*

$$\left| G(z, z') - \sum_{i=1}^{4} \lambda_i(z) G(q_i, z') \right| \leq \epsilon(Q, Q'). \tag{51}$$

Proof: For notational convenience, we set $w = z'$. Let

$$X = (2 + 2z^* \cdot \widehat{w})^k. \tag{52}$$

The Local Energy Lemma follows from the triangle inequality and the following two inequalities

$$\Big| \sum_{i=1}^{4} \lambda_i G(q_i, w) - X \Big| \leq \frac{1}{2} k(k-1) T^{k-2} d^2 \tag{53}$$

$$|X - G(z, w)| \leq 2kT^{k-1}\delta. \tag{54}$$

We will establish these inequalities in turn.

Let q_1, q_2, q_3, q_4 be the vertices of Q. Let $\lambda_i = \lambda_i(z)$. We set

$$x_i = 4 - \|\widehat{q}_i - \widehat{w}\|^2 = 2 + 2\widehat{q}_i \cdot \widehat{w}, \qquad i = 1, 2, 3, 4. \tag{55}$$

Note that $x_i \geq 0$ for all i. We order so that $x_1 \leq x_2 \leq x_3 \leq x_4$. We have

$$\sum_{i=1}^{4} \lambda_i(z) G(q_i, w) = \sum_{i=1}^{4} \lambda_i x_i^k, \tag{56}$$

$$X = (2 + 2\widehat{z}^* \cdot \widehat{w})^k = \Big(\sum_{i=1}^{4} \lambda_i (2 + \widehat{q}_i \cdot \widehat{w})\Big)^k = \Big(\sum_{i=1} \lambda_i x_i\Big)^k. \tag{57}$$

By Equation 56, Equation 57, and the case $M = 4$ of Lemma 5.1,

$$\Big|\sum_{i=1}^{4} \lambda_i G(q_i, w) - X\Big| = \Big|\sum_{i=1}^{4} \lambda_i x_i^k - \Big(\sum_{i=1}^{4} \lambda_i x_i\Big)^k\Big| \leq \frac{1}{8} k(k-1) x_4^{k-2} (x_4 - x_1)^2. \tag{58}$$

By Lemma 4.3, we have

$$x_4 = 2 + 2(\widehat{q}_4 \cdot \widehat{w}) \leq 2 + 2(Q \cdot Q')_{\max} = T. \tag{59}$$

Since d is the diameter of $\langle \widehat{Q} \rangle$ and $\widehat{w}$ is a unit vector,

$$x_4 - x_1 = 2\widehat{w} \cdot (\widehat{q}_4 - \widehat{q}_1) \leq 2\|\widehat{w}\| \|\widehat{q}_4 - \widehat{q}_1\| = 2\|\widehat{q}_4 - \widehat{q}_1\| \leq 2d. \tag{60}$$

Plugging Equations 59 and 60 into Equation 58, we get Equation 53.

Now we establish Equation 54. Let γ denote the unit speed line segment connecting $\widehat{z}$ to z^*. Note that the length L of γ is at most δ, by the Hull Approximation Lemma. Define

$$f(t) = \Big(2 + 2\widehat{w} \cdot \gamma(t)\Big)^k. \tag{61}$$

We have $f(0) = X$. Since $\widehat{w}$ and $\gamma(1) = \widehat{z}$ are unit vectors, $f(L) = G(z, w)$. Hence

$$X - G(z, w) = f(0) - f(L), \qquad L \leq \delta. \tag{62}$$

By the Chain Rule,

$$f'(t) = (2\widehat{w} \cdot \gamma'(t)) \times k\Big(2 + 2\widehat{w} \cdot \gamma(t)\Big)^{k-1}. \tag{63}$$

Note that $|2\widehat{w} \cdot \gamma'(t)| \leq 2$ because both of these vectors are unit vectors. γ parametrizes one of the connectors from Lemma 4.3, so

$$|f'(t)| \leq 2k\Big(2 + 2\widehat{w} \cdot \gamma(t)\Big)^{k-1} \leq 2k\Big(2 + 2(Q \cdot Q')_{\max}\Big)^{k-1} = 2kT^{k-1}. \tag{64}$$

Equation 54 now follows from Equation 62, Equation 64, and integration. ♠

5.3 From Local to Global

Let ϵ be the function from the Energy Theorem. Let $B = (Q_0, ..., Q_N)$ be a list of $N+1$ elements of $\mathcal{Q}$. We care about the case $N = 4$ and $Q_4 = \{\infty\}$, but the added generality makes things clearer. Let $q_{i,1}, q_{i,2}, q_{i,3}, q_{i,4}$ be the vertices of Q_i. The vertices of $\langle B \rangle$ are indexed by a multi-index

$$I = (i_0, ..., i_n) \in \{1, 2, 3, 4\}^{N+1}.$$

Given such a multi-index, which amounts to a choice of vertex of $\langle B \rangle$, we define the energy of the corresponding vertex configuration:

$$\mathcal{E}(I) = \mathcal{E}(q_{0,i_0}, ..., q_{N,i_N}) \tag{65}$$

We will prove the following sharper result.

Theorem 5.3 *Let $z_0, ..., z_N \in \langle B \rangle$. Then*

$$\left| \mathcal{E}(z_0, ..., z_N) - \sum_I \lambda_{i_0}(z_0) ... \lambda_{i_N}(z_N) \mathcal{E}(I) \right| \leq \sum_{i=0}^N \sum_{j=0}^N \epsilon(Q_i, Q_j). \tag{66}$$

The sum is taken over all multi-indices.

Lemma 5.4 *Theorem 5.3 implies Theorem 4.6.*

Proof: Notice that

$$\sum_I \lambda_{i_0}(z_0) ... \lambda_{i_N}(z_N) = \prod_{j=0}^N \Big(\sum_{a=1}^4 \lambda_a(z_j) \Big) = 1. \tag{67}$$

Therefore

$$\min_{v \in B} \mathcal{E}(v) \leq \sum_I \lambda_{i_0}(z_0) ... \lambda_{i_N}(z_N) \mathcal{E}(I) \leq \max_{v \in B} \mathcal{E}(v), \tag{68}$$

because the sum in the middle is the convex average of vertex energies.

We will deal with the min case of Theorem 4.6. The max case has the same treatment. Choose some $(z_1, ..., z_N) \in B$ which minimizes $\mathcal{E}$. We have

$$0 \leq \min_{v \in B} \mathcal{E}(v) - \min_{v \in \langle B \rangle} \mathcal{E}(v) = \min_{v \in B} \mathcal{E}(v) - \mathcal{E}(z_0, ..., z_N) \leq$$

$$\sum_I \lambda_{i_0}(z_0) ... \lambda_{i_N}(z_N) \mathcal{E}(I) - \mathcal{E}(z_0, ..., z_N) \leq \sum_{i=0}^N \sum_{j=0}^N \epsilon(Q_i, Q_j). \tag{69}$$

The last expression is **ERR** when $N = 4$ and $Q_4 = \infty$. ♠

We now prove Theorem 5.3.

5.3.1 A Warmup Case

Consider the case when $N = 1$. Setting $\epsilon_{ij} = \epsilon(Q_i, Q_j)$, the Local Energy Lemma gives us

$$G(z_0, z_1) \geq \sum_{\alpha=1}^{4} \lambda_\alpha(z_0) G(q_{0\alpha}, z_1) - \epsilon_{01}. \tag{70}$$

$$G(q_{0\alpha}, z_1) \geq \sum_{\beta=1}^{4} \lambda_\beta(z_1) G(q_{1\beta}(z_1), q_{0\alpha}) - \epsilon_{10}. \tag{71}$$

Plugging the second equation into the first and using $\sum \lambda_\alpha(z_0) = 1$, we have

$$G(z_0, z_1) \geq \Big(\sum_{\alpha,\beta} \lambda_\alpha(z_0)\lambda_\beta(z_1) G(q_{0\alpha}, q_{1\beta})\Big) - (\epsilon_{01} + \epsilon_{10}). \tag{72}$$

Similarly,

$$G(z_0, z_1) \leq \Big(\sum_{\alpha,\beta} \lambda_\alpha(z_0)\lambda_\beta(z_1) G(q_{0\alpha}, q_{1\beta})\Big) + (\epsilon_{01} + \epsilon_{10}). \tag{73}$$

Equations 72 and 73 are equivalent to Equation 66 when $N = 1$.

5.3.2 The General Case

Now assume that $N \geq 2$. We rewrite Equation 72 as follows:

$$G(z_0, z_1) \geq \sum_{A} \lambda_{A_0}(z_0)\lambda_{A_1}(z_1)\ G(q_{0A_0}, q_{1A_1}) - (\epsilon_{01} + \epsilon_{10}). \tag{74}$$

The sum is taken over multi-indices A of length 2.

We also observe that

$$\sum_{I'} \lambda_{i_2}(z_2) \ldots \lambda_{i_N}(z_N) = 1. \tag{75}$$

The sum is taken over all multi-indices $I' = (i_2, \ldots, i_N)$. Therefore, if A is held fixed, we have

$$\lambda_{A_0}(z_0)\lambda_{A_1}(z_1) = \sum_{I}{}' \lambda_{I_0}(z_0) \ldots \lambda_{I_N}(z_N). \tag{76}$$

The sum is taken over all multi-indices of length $N+1$ which have $I_0 = A_0$ and $I_1 = A_1$. Combining these equations, we have

$$G(z_0, z_1) \geq \sum_I \lambda_{I_0}(z_0)...\lambda_{I_N}(z_N) G(q_{0I_0}, q_{1I_1}) - (\epsilon_{01} + \epsilon_{10}). \tag{77}$$

The same argument works for other pairs of indices, giving

$$G(z_i, z_j) \geq \sum_I \lambda_{I_0}(z_0)...\lambda_{I_N}(z_N) G(q_{iI_i}, q_{jI_j}) - (\epsilon_{ij} + \epsilon_{ji}). \tag{78}$$

Now we interchange the order of summation and observe that

$$\sum_{i<j} \Big(\sum_I \lambda_{I_0}(z_0)...\lambda_{I_N}(z_N)\ G(q_{iI_i}, q_{jI_j}) \Big) =$$

$$\sum_I \sum_{i<j} \lambda_{I_0}(z_0)...\lambda_{I_N}(z_N)\ G(q_{iI_i}, q_{jI_j}) =$$

$$\sum_I \lambda_{I_0}(z_0)...\lambda_{I_N}(z_N) \Big(\sum_{i<j} G(q_{iI_i}, q_{jI_j}) \Big) =$$

$$\sum_I \lambda_{I_0}(z_0)...\lambda_{I_N}(z_N)\ \mathcal{E}(I). \tag{79}$$

When we sum Equation 78 over all $i < j$, we get

$$\mathcal{E}(z_0, ..., z_N) \geq \sum_I \lambda_{i_0}(z_0)...\lambda_{i_N}(z_N) \mathcal{E}(I) - \sum_{i=0}^{N} \sum_{j=0}^{N} \epsilon(Q_i, Q_j). \tag{80}$$

Similary,

$$\mathcal{E}(z_0, ..., z_N) \leq \sum_I \lambda_{i_0}(z_0)...\lambda_{i_N}(z_N) \mathcal{E}(I) + \sum_{i=0}^{N} \sum_{j=0}^{N} \epsilon(Q_i, Q_j). \tag{81}$$

These two equations together are equivalent to Theorem 5.3. This completes the proof.

6 Symmetrization

6.1 A Three Step Process

In this chapter we sketch the ideas behind our Symmetrization Lemma. Our symmetrization is a retraction from **SMALL** to **K4**. Here we describe it as a 3 step process. It is convenient to set $\Omega =$ **SMALL**. We start with the configuration X having points $(p_1, p_2, p_3, p_4) \in \Omega$.

1. (Rotating) We let (p_1', p_2', p_3', p_4') be the configuration which is obtained by rotating X about the origin so that p_0' and p_2' lie on the same horizontal line, with p_0' lying on the right. This slight rotation does not change the energy of the configuration, but it does slightly change the domain. While $X \in \Omega$, the new configuration X' lies in a slightly modified domain Ω' which we describe in [**S0**, §13].

2. (Horizontal Sliding) Given a configuration $X' = (p_0', p_1', p_2', p_3') \in \Omega'$, there is a unique configuration $X'' = (p_0'', p_1'', p_2'', p_3'')$, invariant under under reflection in the y-axis, such that p_j' and p_j'' lie on the same horizontal line for $j = 0, 1, 2, 3$ and $\|p_0'' - p_2''\| = \|p_0' - p_2'\|$. There is a slightly different domain Ω'' which contains X''. Again, we describe Ω'' precisely in [**S0**, §13]. The domain Ω'' is 4-dimensional.

3. (Vertical Sliding) Given a configuration $X'' = (p_0'', p_1'', p_2'', p_3'') \in \Omega''$ there is a unique configuration $X^* = (p_0^*, p_1^*, p_2^*, p_3^*) \in$ **K4** such that p_j'' and p_j^* lie on the same vertical line for $j = 0, 1, 2, 3$.

To prove the Symmetrization Lemma, it suffices to prove the following two results.

Lemma 6.1 *$R_s(X'') \leq R_s(X')$ for all $X' \in \Omega'$ and $s \geq 2$, with equality iff $X' = X''$.*

Lemma 6.2 *$R_s(X^*) \leq R_s(X'')$ for all $X'' \in \Omega'$ and $s \in [12, 15 + 1/2]$, with equality iff $X'' = X^*$.*

Before we launch into the proofs, we mention that we tested these inequalities extensively on random inputs before attempting a proof. The inequalities seem to be quite robust, and work for a much wider domain than just the domains Ω' and Ω'' derived from **SMALL**. However, our proof strategy really requires the small domains to work. I don't know a robust proof.

6.2 Horizontal Symmetrization Proof Strategy

There are 10 bonds (i.e. distances between pairs of points) in a 5 point configuration $\widehat{p}_0, ..., \widehat{p}_4$. Ultimately, when we perform the symmetrization operations we want to see that a sum of 10 terms decreases. We find it convenient to break this sum into 4 pieces and study the pieces separately. We define

$$A_{1,s} = R_s(\widehat{p}_1, \widehat{p}_0) + R_s(\widehat{p}_1, \widehat{p}_2). \tag{82}$$

$$A_{2,s} = R_s(\widehat{p}_3, \widehat{p}_0) + R_s(\widehat{p}_3, \widehat{p}_2). \tag{83}$$

$$B_{13,s} = R_s(\widehat{p}_1, \widehat{p}_3) + R_s(\widehat{p}_1, \widehat{p}_4) + R_s(\widehat{p}_3, \widehat{p}_4). \tag{84}$$

$$B_{02,s} = R_s(\widehat{p}_0, \widehat{p}_2) + R_s(\widehat{p}_0, \widehat{p}_4) + R_s(\widehat{p}_2, \widehat{p}_4). \tag{85}$$

To prove Lemma 6.1 we establish the following stronger results. These results are meant to hold for all $s \geq 2$ and all $X' \in \Omega'$.

1. $A_{1,s}(X'') \leq A_{1,s}(X')$.
2. $A_{3,s}(X'') \leq A_{3,s}(X')$.
3. $B_{02,s}(X'') \leq B_{02,s}(X')$ with equality iff $X' = X''$.
4. $B_{13,s}(X'') \leq B_{13,s}(X')$ with equality iff $X' = X''$.

Inequalities 3 and 4 must be proved for each exponent s, but they are pretty easy. Each inequality involves a pair of points in the plane. Our proof is more or less geometric. See [**S0**, §14].

Inequalities 1 and 2 seem much harder, and I don't know a geometric proof.. However, we first note the following standard result.

Lemma 6.3 (Monotonicity) *Let $\lambda_1, ..., \lambda_n > 0$. If $\beta/\alpha > 1$. Then*

$$\sum_{i=1}^{N} \lambda_i^{\beta} - N \geq \frac{\beta}{\alpha}\Big(\sum_{i=1}^{N} \lambda_i^{\alpha} - N\Big).$$

See [**S0**, §12.1] for a proof. It follows from the Monotonicity Lemma that the truth of Inequality 1 at $s = 2$ implies the truth of Inequality 1 for $s > 2$. The same goes for Inequality 2. We deal with the case $s = 2$ with a direct calculation involving the positivity certificate from Lemma 3.1.

Consider Inequality 1. Since the x-coordinates of p_0 and p_2 are the same, we just have a 5-dimensional space of possibilities. We find 4 maps f_1, f_2, f_3, f_4 from the unit cube $[0,1]^5$ into the configuration space of triples of points such that the union $\bigcup f_j([0,1]^5)$ covers the projection of the domain Ω' onto the first 6 coordinates. That is, every triple (p_0, p_1, p_2) that arises in a configuration of Ω' lies in the images of one of the cubes.

We then define

$$\phi_j = (A_s^{\text{bottom}}(X') - A_s^{\text{bottom}}(X'')) \circ f_j. \tag{86}$$

The function ϕ_j is a rational function on the unit cube for each $j = 1,2,3,4$. The numerator and denominators of ϕ_j are enormous. They have several thousand terms and degree about 40. Nonetheless, we show that both the numerator and denominator satisfy the Positivity Criterion in Lemma 3.1! This establishes Inequality 1. See [**S0**, §13,14] for more details.

We mention that the result of Lemma 6.1, which we have fairly well explained, already shows that any minimizer for R_s with $s \in [13, 15 + \frac{25}{512}]$ has to have 3 points on an equator and the remaining two points symmetrically placed with respect to reflection in this equator. It seems that there should be many endgames from here.

6.3 Vertical Symmetrization in Brief

It seems that Inequalities 1-4 also work in the case of Lemma 6.2, except that for Inequalities 1 and 2 we need to take s fairly large. I found failures for exponents as high as $s = 9$. In principle, we could prove Lemma 6.2 using a similar strategy, except that we would pick, say $s = 12$, for Inequality 1. The problem is that the high exponent leads to enormous polynomials. When trying to understand the failure of Inequality 1 for small exponents, I noticed the following phemonenon: When Inequalities 1 or 2 fail, Inequalities 3 and 4 hold by a wide margin. The proof of Lemma 6.2, given in [**S0**, §13-15], makes calculations for Inequalities 1 and 2 at the exponent $s = 2$ similar to the ones described for Lemma 6.2, and then uses L_p estimates to exploit the phenomenon just mentioned. The proof is considerably more delicate.

7 References

[**A**] A. N. Andreev, *An extremal property of the icosahedron* East J Approx **2** (1996) no. 4 pp. 459-462

[**BBCGKS**] Brandon Ballinger, Grigoriy Blekherman, Henry Cohn, Noah Giansiracusa, Elizabeth Kelly, Achill Schurmann,
Experimental Study of Energy-Minimizing Point Configurations on Spheres, arXiv: math/0611451v3, 7 Oct 2008

[**BDHSS**] P. G. Boyvalenkov, P. D. Dragnev, D. P. Hardin, E. B. Saff, M. M. Stoyanova, *Universal Lower Bounds and Potential Energy of Spherical Codes*, Constructive Approximation 2016 (to appear)

[**C**] Harvey Cohn, *Stability Configurations of Electrons on a Sphere*, Mathematical Tables and Other Aids to Computation, Vol 10, No 55, July 1956, pp 117-120.

[**CK**] Henry Cohn and Abhinav Kumar, *Universally Optimal Distributions of Points on Spheres*, J.A.M.S. **20** (2007) 99-147

[**CCD**] online website:
http://www-wales.ch.cam.ac.uk/∼ wales/CCD/Thomson/table.html

[**DLT**] P. D. Dragnev, D. A. Legg, and D. W. Townsend, *Discrete Logarithmic Energy on the Sphere*, Pacific Journal of Mathematics, Volume 207, Number 2 (2002) pp 345–357

[**HS**], Xiaorong Hou and Junwei Shao, *Spherical Distribution of 5 Points with Maximal Distance Sum*, arXiv:0906.0937v1 [cs.DM] 4 Jun 2009

[**I**] IEEE Standard for Binary Floating-Point Arithmetic (IEEE Std 754-1985) Institute of Electrical and Electronics Engineers, July 26, 1985

[**RSZ**] E. A. Rakhmanoff, E. B. Saff, and Y. M. Zhou, *Electrons on the Sphere*,
Computational Methods and Function Theory, R. M. Ali, St. Ruscheweyh, and E. B. Saff, Eds. (1995) pp 111-127

[**S1**] R. E. Schwartz, *The* 5 *Electron Case of Thomson's Problem*, Journal of Experimental Math, 2013.

[**S2**] R. E. Schwartz, *The projective Heat Map*, A.M.S. Mathematical Surveys and Monographs **219**, 2017

[**SK**] E. B. Saff and A. B. J. Kuijlaars, *Distributing many points on a Sphere*, Math. Intelligencer, Volume 19, Number 1, December 1997 pp 5-11

[**Th**] J. J. Thomson, *On the Structure of the Atom: an Investigation of the Stability of the Periods of Oscillation of a number of Corpuscles arranged at equal intervals around the Circumference of a Circle with Application of the results to the Theory of Atomic Structure*. Philosophical magazine, Series 6, Volume 7, Number 39, pp 237-265, March 1904.

[**T**] A. Tumanov, *Minimal Bi-Quadratic energy of* 5 *particles on* 2*-sphere*, Indiana Univ. Math Journal, **62** (2013) pp 1717-1731.

[**W**] S. Wolfram, *The Mathematica Book*, 4th ed. Wolfram Media/Cambridge University Press, Champaign/Cambridge (1999)

[**Y**], V. A. Yudin, *Minimum potential energy of a point system of charges* (Russian) Diskret. Mat. **4** (1992), 115-121, translation in Discrete Math Appl. **3** (1993) 75-81