Five Point Energy Minimization 3: Local Analysis

Richard Evan Schwartz

April 12, 2024

Abstract

This is Paper 3 of series of 7 self-contained papers which together prove the Melnyk-Knopf-Smith phase transition conjecture for 5-point energy minimization. (Paper 0 has the main argument.) This paper deals with a local analysis of configurations near the triangular bi-pyramid.

1 Introduction

1.1 Context

During the past decade I have written several versions of a proof that rigorously verifies the phase-transition for 5 point energy minimization first observed in [**MKS**], in 1977, by T. W. Melnyk, O, Knop, and W. R. Smith. See [**S0**] for the latest version. This work implies and extends my solution [**S1**] of Thomson's 1904 5-electron problem [**Th**]. Unfortunately, after a number of attempts I have not been able to publish my work on this. Even though I have taken great pains to make the proof modular and checkable, the monograph still gives the impression of being too difficult to referee.

Now I am taking a new approach. I have broken down the proof into a series of 7 independent papers, each of which may be checked without any reference to the others. The longest of the papers is 20 pages. The drawback of this approach is twofold. First, there will necessarily be some redundancy in these papers. Second, none of the papers has a blockbuster result in itself. To help offset the second drawback, I will state the main result in full in each paper, and I will try to explain how the small result proved in each paper relates to the overall goal.

1.2 The Phase Transition Result

Let S^2 be the unit sphere in $\boldsymbol{R}^3$. Given a configuration $\{p_i\} \subset S^2$ of N distinct points and a function $F : (0,2] \to \boldsymbol{R}$, define

$$\mathcal{E}_F(P) = \sum_{1 \le i < j \le N} F(\|p_i - p_j\|). \tag{1}$$

This quantity is commonly called the *F-potential* or the *F-energy* of P. A configuration P is a *minimizer* for F if $\mathcal{E}_F(P) \le \mathcal{E}_F(P')$ for all other N-point configurations P'.

We are interested in the *Riesz potentials*:

$$R_s(d) = d^{-s}, \qquad s > 0. \tag{2}$$

R_s is also called a *power law potential*, and R_1 is specially called the *Coulomb potential* or the *electrostatic potential*. The question of finding the N-point minimizers for R_1 is commonly called *Thomson's problem*.

We consider the case $N = 5$. The *Triangular Bi-Pyramid* (TBP) is the 5 point configuration having one point at the north pole, one point at the south pole, and 3 points arranged in an equilateral triangle on the equator. A *Four Pyramid* (FP) is a 5-point configuration having one point at the north pole and 4 points arranged in a square equidistant from the north pole.

Define

$$15_+ = 15 + \frac{25}{512}. \tag{3}$$

Theorem 1.1 (Phase Transition) *There exists $\mathfrak{w} \in (15, 15_+)$ such that:*

1. *For $s \in (0, \mathfrak{w})$ the TBP is the unique minimizer for R_s.*

2. *For $s = \mathfrak{w}$ the TBP and some FP are the two minimizers for R_s.*

3. *For each $s \in (\mathfrak{w}, 15_+)$ some FP is the unique minimizer for R_s.*

The proof has many moving parts. The largest part involves eliminating most of the configurations using a divide-and-conquer algorithm. Since the algorithm is based on a finite calculation, it would get irreparable bogged down if it had to eliminate configurations arbitrarily near the TBP. In this paper we do some local analysis which automatically eliminates all the configurations in a definite, explicit neighborhood of the TBP.

1.3 The Result of This Paper

We first give some background information.

Stereographic Projection: Let $S^2 \subset \boldsymbol{R}^3$ be the unit 2-sphere. *Stereographic projection* is the map $\Sigma : S^2 \to \boldsymbol{R}^2 \cup \infty$ given by the following formula.

$$\Sigma(x,y,z) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right). \tag{4}$$

Here is the inverse map:

$$\Sigma^{-1}(x,y) = \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, 1 - \frac{2}{1+x^2+y^2}\right). \tag{5}$$

Σ^{-1} maps circles in $\boldsymbol{R}^2$ to circles in S^2 and $\Sigma^{-1}(\infty) = (0,0,1)$.

Avatars: Stereographic projection gives us a correspondence between 5-point configurations on S^2 having $(0,0,1)$ as the last point and planar configurations:

$$\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3, (0,0,1) \in S^2 \iff p_0, p_1, p_2, p_3 \in \boldsymbol{R}^2, \qquad \widehat{p}_k = \Sigma^{-1}(p_k). \tag{6}$$

We call the planar configuration the *avatar* of the corresponding configuration in S^2. By a slight abuse of notation we write $\mathcal{E}_F(p_0,p_1,p_2,p_3)$ when we mean the F-potential of the corresponding 5-point configuration. One of the avatars representing the TBP is

$$p_0 = -p_2 = (1,0), \qquad p_1 = -p_3 = (0,-u), \qquad u = \sqrt{3}/3.$$

The Special Potentials: Rather than work with the Riesz potentials, we work with potentials that have a more polynomial flavor.

$$G_k(r) = (4 - r^2)^k. \tag{7}$$

Also define

$$G_5^\flat = G_5 - 25G_1, \quad G_{10}^{\sharp\sharp} = G_{10} + 28G_5 + 102G_2, \quad G_{10}^\sharp = G_{10} + 13G_5 + 68G_2$$

These potentials are related to the Riesz potentials in a way that we discuss below.

The Definite Neighborhood: We specially treat avatars very near the TBP. When we string out the points of ξ_0, we get $(1, \cdot, 0, -u, -1, 0, 0, u)$ where $u = \sqrt{3}/3$. The $(\cdot)$ indicates that we do not record $p_{02} = 0$. We let Ω_0 denote the cube of side-length 2^{-17} centered at ξ_0. For all our choices of F, the function $\mathcal{E}_F$ is a smooth function on $\boldsymbol{R}^7$. We check first of all that the gradient of $\mathcal{E}_F$ vanishes at ξ_0. This probably follows from symmetry, but to be sure we make a direct calculation in all cases.

Recall that the *Hessian* of a function is its matrix of second partial derivatives. Here is the main result of this paper.

Theorem 1.2 (Local Convexity) *For each $F = G_4, G_6, G_5^\flat, G_{10}^\sharp$, the Hessian of $\mathcal{E}_F$ is positive definite at every point of Ω_0.*

Corollary 1.3 *Let F be any of $G_4, G_5^\flat, G_5, G_6, G_{10}^\sharp, G_{10}^{\sharp\sharp}$. Then ξ_0 is the unique minimizer for $\mathcal{E}_F$ inside Ω_0.*

Proof: Let F be any of the functions from the Local Convexity Theorem. Let $\xi \in \Omega_0$ be other than ξ_0. The Local Convexity Theorem combines with the vanishing gradient to show that the restriction of $\mathcal{E}_F$ to the line segment γ joining ξ_0 to ξ is convex and has 0 derivative at ξ_0. Hence $\mathcal{E}_F(\xi) > \mathcal{E}_F(\xi_0)$.

It remains to deal with $F = G_5$ and $F = G_{10}^{\sharp\sharp}$. As is well known, ξ_0 is a minimizer for G_1. Since ξ_0 is the unique minimizer for $G_5^\flat$ in Ω_0, we see that ξ_0 is also the unique minimizer for $G_5 = G_5^\flat + 25G_1$ in Ω_0.

By the main result in [**T**], ξ_0 is the unique global minimizer for G_2. With this in mind, we see that the same kind of argument we just gave for G_5 also works for $G_{10}^{\sharp\sharp} = G_{10}^\sharp + 15G_5 + 34G_2$. ♠

Our result here combines with the main result in Paper 4, the Interpolation Theorem, to show that the TBP is the unique minimizer, amongst configurations having avatars in Ω_0, for all R_s with $s \in (0, 15_+]$. Since we prefer to use the potentials from Corollary 1.3, we do not directly use this result for any purpose. We mention it just to show how it ties in with the Phase Transition Theorem.

The proofs in this paper are computer-assisted. All calculations are all done using exact arithmetic in Mathematica. The reader can download and inspect the files I wrote for this. The basic idea of the proof is to get *a priori* bounds on high order partial derivatives of the functions involved and then combine these bounds with explicit evaluations of lower order partial derivatives at the point representing the TBP. Equation 20 is the magic equation for the high order partial derivatives.

2 Proof of the Local Convexity Theorem

2.1 Reduction to Simpler Statements

We consider F to be any of the 4 functions

$$G_4, \quad G_6, \quad G_5^\flat = G_5 - 25G_1, \quad 2^{-5}G_{10}^\sharp = 2^{-5}(G_{10} + 13G_5 + 68G_2).$$

Scaling the last function by 2^{-5} makes our estimates more uniform.

Recall that Ω_0 is the cube of side length 2^{-17} centered at the point

$$\xi_0 = \left(1, 0, \frac{-1}{\sqrt{3}}, -1, 0, 0, \frac{1}{\sqrt{3}}\right) \in \boldsymbol{R}^7 \tag{8}$$

In general, the point $(x_1, ..., x_7)$ represents the avatar

$$p_0 = (x_1, 0),\ p_1 = (x_2, x_3),\ p_2 = (x_4, x_5),\ p_3 = (x_6, x_7). \tag{9}$$

The quantity $\mathcal{E}_F(x_1, ..., x_7)$ is the F-potential of the 5-point configuration associated to the avatar under inverse stereographic projection Σ^{-1}.

$$\mathcal{E}_F(x_1, ..., x_7) = \sum_{i<j} F(\|\widehat{p}_i - \widehat{p}_j\|), \qquad \widehat{p} = \Sigma^{-1}(p). \tag{10}$$

Equation 5 gives the formula for Σ^{-1}.

Let $H\mathcal{E}_F$ be the Hessian of $\mathcal{E}_F$. The Local Convexity Theorem says $H\mathcal{E}_F$ is positive definite in Ω_0. Let $\partial_J \mathcal{E}_F$ be the (iterated) partial derivative of $\mathcal{E}_F$ with respect to a multi-index $J = (j_1, ..., j_7)$. Let $|J| = j_1 + ... + j_7$. Let

$$M_N = \sup_{|J|=N} M_J, \qquad M_J = \sup_{\xi \in \Omega_0} |\partial_J \mathcal{E}_F(\xi)|, \tag{11}$$

Let $\lambda(M)$ be the smallest eigenvalue of a real symmetric matrix M. The Local Convexity Theorem is an immediate consequence of the following two lemmas.

Lemma 2.1 *If $M_3(\mathcal{E}_F) < 2^{12}\lambda(H\mathcal{E}_F(\xi_0))$ then $\lambda(H\mathcal{E}_F(\xi)) > 0$ for all points $\xi \in \Omega_0$.*

Lemma 2.2 *$M_3(\mathcal{E}_F) < 2^{12}\lambda(H\mathcal{E}_F(\xi_0)))$ in all cases.*

2.2 Proof of Lemma L1

Let

$$H_0 = H\mathcal{E}_F(\xi_0), \qquad H = H\mathcal{E}_F(\xi), \qquad \Delta = H - H_0. \tag{12}$$

For any real symmetric matrix X define the L_2 matrix norm:

$$\|X\|_2 = \sqrt{\sum_{ij} X_{ij}^2} = \sup_{\|v\|=1} \|Xv\|. \tag{13}$$

Given a unit vector $v \in \boldsymbol{R}^7$ we have $H_0 v \cdot v \geq \lambda$. Hence

$$Hv \cdot v = (H_0 v + \Delta v) \cdot v \geq H_0 v \cdot v - |\Delta v \cdot v| \geq \lambda - \|\Delta v\| \geq \lambda - \|\Delta\|_2 > 0.$$

So, to prove Lemma L1 we just need to establish the implication

$$M_3 < 2^{12}\lambda(H_0) \quad \Longrightarrow \quad \|\Delta\|_2 < \lambda(H_0).$$

Let $t \to \gamma(t)$ be the *unit speed parametrized* line segment connecting p_0 to p in Ω_0. Note that γ has length $L \leq \sqrt{7} \times 2^{-18}$. We write $\gamma = (\gamma_1, ..., \gamma_7)$. Let H_t denote the Hessian of $\mathcal{E}_F$ evaluated at $\gamma(t)$. Let D_t denote the directional derivative along γ.

Now $\|D_t(H_t)\|_2$ is the speed of the path $t \to H_t$ in $\boldsymbol{R}^{49}$, and $\|\Delta\|_2$ is the Euclidean distance between the endpoints of this path. Therefore

$$\|\Delta\|_2 \leq \int_0^L \|D_t(H_t)\|_2 \; dt. \tag{14}$$

Let $(H_t)_{ij}$ denote the ijth entry of H_t. From the definition of directional derivatives, and from the Cauchy-Schwarz inequality, we have

$$(D_t H_t)_{ij}^2 = \left(\sum_{k=1}^{7} \frac{d\gamma_k}{dt} \frac{\partial H_{ij}}{\partial k} \right)^2 \leq 7M_3^2. \qquad \|D_t(H_t)\|_2 \leq 7^{3/2} M_3. \tag{15}$$

The second inequality follows from summing the first one over all 7^2 pairs (i, j) and taking the square root. Equation 14 now gives

$$\|\Delta\|_2 \leq L \times 7^{3/2} M_3 = 49 \times 2^{-18} M_3 < 2^{-12} M_3 < \lambda(H_0). \tag{16}$$

This completes the proof.

2.3 Proof of Lemma 2.2

Let F be any of our functions. Let $H_0 = H\mathcal{E}_F(\xi_0)$.

Lemma 2.3 $\lambda(H_0) > 39$.

Proof: Let χ be the characteristic polynomial of H_0. This turns out to be a rational polynomial. We check in Mathematica that the signs of the coefficients of $\chi(t+39)$ alternate. Hence $\chi(t+39)$ has no negative roots. The file we use is `LemmaL21.m`. ♠

Recalling that $\xi_0 \in \boldsymbol{R}^7$ is the point representing the TBP, we define

$$\mu_N(\mathcal{E}_F) = \sup_{|I|=N} |\partial_I \mathcal{E}_F(\xi_0)|. \tag{17}$$

Lemma 2.4 *For any of our functions we have the bound*

$$\mu_3 < 45893, \qquad \frac{(7\times 2^{-18})^j}{j!}\mu_{j+3} < 38, \qquad j=1,2,3. \tag{18}$$

Proof: We compute this in Mathematica. The file we use is `LemmaL22.m`. ♠

Lemma 2.5 *For any of our functions we have the bound*

$$\frac{(7\times 2^{-18})^4}{4!}M_7 < 2354.$$

Proof: We give this proof in the next section. ♠

Lemma 2.6 *We have*

$$M_3 \le \mu_3 + \sum_{j=1}^{3} \frac{(7\times 2^{-18})^j}{j!}\mu_{j+3} + \frac{(7\times 2^{-18})^4}{4!}M_7 \tag{19}$$

Proof: Choose any multi-index J with $|J|=3$. Let γ be the line segment connecting ξ_0 to any $\xi \in \Omega$. We parametrize γ by unit speed and furthermore set $\gamma(0)=\xi_0$. Let

$$f(t) = \partial_J \mathcal{E}_F \circ \gamma(t).$$

The bound for $|M_J|$ follows from Taylor's Theorem with remainder once we notice that

$$0 \leq t \leq \sqrt{7} \times 2^{-18}, \qquad \left|\frac{\partial^n f(0)}{\partial t^n}\right| \leq (\sqrt{7})^n \mu_n \qquad \left|\frac{\partial^n f}{\partial t^n}\right| \leq (\sqrt{7})^n M_n.$$

Since this works for all J with $|J| = 3$ we get the same bound for M_3. ♠

The lemmas above and Equation 18 imply

$$M_3 < 45893 + 3 \times 38 + 2354 \leq 65536 = 2^{16} \leq 2^{12}\lambda(H_0).$$

This completes the proof of Lemma 2.2.

2.4 Proof of Lemma 2.5

Now we come to the interesting part of the proof, the one place where we need to go beyond specific evaluations of our functions. When $r, s \geq 0$ and $r + s \leq 2d$ we have

$$\sup_{(x,y)\in \boldsymbol{R}^2} \frac{x^r y^s}{(1 + x^2 + y^2)^d} \leq (1/2)^{\min(r,s)}. \tag{20}$$

One can prove Equation 20 by factoring the expression into pieces with quadratic denominators. Here is a more general version. Say that a function $\phi : \boldsymbol{R}^4 \to \boldsymbol{R}$ is *nice* if it has the form

$$\sum_i \frac{C_i a^{\alpha_i} b^{\beta_i} c^{\gamma_i} d^{\delta_i}}{(1 + a^2 + b^2)^{u_i}(1 + c^2 + d^2)^{v_i}}, \quad \alpha_i, \beta_i, \gamma_i, \delta_i \geq 0, \quad \alpha_i + \beta_i \leq 2u_i, \qquad \gamma_i + \delta_i \leq 2v_i.$$

It follows from Equation 20 that

$$\sup_{\boldsymbol{R}^4} |\phi| \leq \langle\phi\rangle, \qquad \langle\phi\rangle = \sum_i |C_i|(1/2)^{\min(\alpha_i,\beta_i)+\min(\gamma_i,\delta_i)}. \tag{21}$$

Equation 21 is useful to us because it allows us to bound certain kinds of functions without having to evaluate then anywhere. We also note that if ϕ is nice, then so is any iterated partial derivative of ϕ. Indeed, the nice functions form a ring that is invariant under partial differentiation. This fact makes it easy to identify nice functions.

For any $\phi : \boldsymbol{R}^n \to \boldsymbol{R}$ we define

$$\overline{M}_7(\psi) = \sup_{|J|=7} \overline{M}_J(\psi), \qquad \overline{M}_J(\psi) = \sup_{\xi\in \boldsymbol{R}^n} |\partial_J(\phi)|. \tag{22}$$

We obviously have

$$M_7(\mathcal{E}_F) \leq \overline{M}_7(\mathcal{E}_F). \tag{23}$$

Recall that $\widehat{p} = \Sigma^{-1}(p)$, the inverse stereographic image of p. Define

$$f(a,b) = 4 - \|\widehat{(a,b)} - (0,0,1)\|^2 = \frac{4(a^2+b^2)}{1+a^2+b^2}. \tag{24}$$

$$g(a,b,c,d) = 4 - \|\widehat{(a,b)} - \widehat{(c,d)}\|^2 = \frac{4(1+2ac+2bd+(a^2+b^2)(c^2+d^2))}{(1+a^2+b^2)(1+c^2+d^2)}. \tag{25}$$

Notice that g is nice. Hence g^k is nice and $\partial_I g^k$ is nice for any multi-index. That means we can apply Equation 21 to $\partial_I g^k$.

$\mathcal{E}_{G_k}$ is a 10-term expression involving 4 instances of f^k and 6 of g^k. However, each variable appears in at most 4 terms. So, as soon as we take a partial derivative, at least 6 of the terms vanish. Moreover, $\partial_I f$ is a limiting case of $\partial_I g$ for any multi-index I. From these considerations, we see that

$$\overline{M}_7(\mathcal{E}_{G_k}) \leq 4 \times \overline{M}_7(g^k). \tag{26}$$

The function $\partial_I(g^k)$ is nice in the sense of Equation 21. Therefore

$$4 \times \overline{M}_7(g^k) \leq 4 \times \max_{|I|=7} \langle \partial_I g^k \rangle. \tag{27}$$

Using this estimate, and the Mathematica file `LemmaL23.m`, we get

$$\max_{k \in \{1,2,3,4,5,6\}} \frac{(7 \times 2^{-18})^4}{4!} \times 4 \times \overline{M}_7(g^k) \leq \frac{1}{1000}.$$

$$2^{-5} \times \frac{(7 \times 2^{-18})^4}{4!} \times 4 \times \overline{M}_7(g^{10}) \leq 2353. \tag{28}$$

The bounds in Lemma 2.5 follow directly from Equations 26 - 28 and from the definitions of our functions.

3 References

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

[**MKS**], T. W. Melnyk, O. Knop, W.R. Smith, *Extremal arrangements of point and and unit charges on the sphere: equilibrium configurations revisited*, Canadian Journal of Chemistry 55.10 (1977) pp 1745-1761

[**S0**] R. E. Schwartz, *Divide and Conquer: A Distributed Approach to* 5-*Point Energy Minimization*, Research Monograph (preprint, 2023)

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

[**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)

See Paper 0 for an extended bibliography.