

Eventual Square Pyramid Optimality for Five-Point Riesz Energy

Alper Ferudun*

14 October 2026

Abstract

We prove that, for every sufficiently large real exponent s , the unique five-point minimizer of chordal Riesz s -energy on the unit two-sphere, up to orthogonal congruence and relabeling, is a square pyramid. Its apex is the north pole and its base has latitude $-t$, where $(1 - t)^{s/2+1} = (2 + 2^{-s/2})t$. Starting from the limiting configuration theorem of Bondarenko, Hardin and Saff, we use angular stationarity and relative height balance to localize every minimizer within $o(1/s)$ of this moving pyramid. A uniform positive Hessian in a surrounding coordinate box proves uniqueness. The threshold is existential: the argument does not identify the transition exponent or settle the intervening range of exponents.

1 The eventual optimality theorem

For five distinct points $X = \{x_0, \dots, x_4\}$ on the unit sphere $S^2 \subset \mathbb{R}^3$, consider the unordered-pair energy

$$\mathcal{E}_s(X) = \sum_{0 \leq i < j \leq 4} |x_i - x_j|^{-s}, \quad s > 0.$$

Distances are Euclidean chordal distances, not spherical geodesic distances. For $0 < t < 1$, write $P(t)$ for the square pyramid with apex $e_3 = (0, 0, 1)$ and base points

$$\left(\sqrt{1-t^2} \cos \frac{j\pi}{2}, \sqrt{1-t^2} \sin \frac{j\pi}{2}, -t \right), \quad 0 \leq j \leq 3. \quad (1)$$

Theorem 1. *There is a finite $S > 0$ such that, for every real $s \geq S$, every global minimizer of $\mathcal{E}_s$ among five-point subsets of S^2 is congruent to $P(t_{s/2})$. Here $t_q \in (0, 1)$ is the unique solution of*

$$(1 - t_q)^{q+1} = (2 + 2^{-q})t_q. \quad (2)$$

Thus the minimizer is unique up to orthogonal congruence and relabeling.

The five-point problem appears as Problem 1.38 in the AIM list [1]. Bondarenko, Hardin and Saff identify the limit of energy minimizers as the exponent tends to infinity [2, Theorem 7]: after congruence and relabeling, the limit is one pole together with an equatorial square. Their theorem supplies qualitative convergence, not exact optimality of a square pyramid at finite exponents. Schwartz [3, Theorem 1.1] establishes the transition to square pyramids on a specified finite interval near the transition. Theorem 1 concerns the eventual large-exponent range; it does not bridge the entire remaining interval.

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: [AlperTheKing](https://github.com/AlperTheKing).

Our proof follows the localization and convexity architecture used for eight points by Kryvonos, Liehr and Taylor [4, Section 7]. Their positive contact-stress localization does not transfer directly to five points. The four apex-to-base contacts have zero coefficients in the unique stress of the limiting five-point packing. Moreover, $t_q \sim (\log q)/q$, so the candidate itself is not within $O(1/q)$ of that packing. We replace this step by relative height balance and localize at the *moving* pyramid $P(t_q)$.

2 Normalization and the symmetric candidate

Put $q = s/2$ and use normalized squared distances. The energy

$$E_q(X) = \sum_{i < j} \left(\frac{|x_i - x_j|^2}{2} \right)^{-q} = 2^q \mathcal{E}_{2q}(X) \quad (3)$$

has the same minimizers. A minimizer exists: the configuration space is compact, and the energy diverges at collisions. All limits below are as $q \rightarrow \infty$ through real values.

The energy of (1) is

$$F_q(t) = 4(1+t)^{-q} + (4+2^{1-q})(1-t^2)^{-q}. \quad (4)$$

Its derivative vanishes precisely when (2) holds. The left side of (2) strictly decreases from 1 to 0, whereas its right side strictly increases from 0. The derivative of F_q is negative before the crossing and positive after it.

Lemma 2. *The pyramid $P(t_q)$ is stationary for E_q in the full configuration space modulo rotations. Furthermore,*

$$t_q \rightarrow 0, \quad qt_q \rightarrow \infty, \quad t_q = O\left(\frac{\log q}{q}\right), \quad qt_q^2 \rightarrow 0, \quad qt_q \sim \log q. \quad (5)$$

Every global minimizer satisfies $E_q \leq 4 + o(1)$.

Proof. Equation (2) first gives $t_q \rightarrow 0$. If qt_q had a bounded subsequence, its left side would stay bounded away from zero while its right side tended to zero, a contradiction. Taking logarithms and using $-\log(1-t) \geq t$ gives

$$(q+1)t_q \leq -\log((2+2^{-q})t_q) \leq \log(1/t_q).$$

Splitting into $t_q \geq 1/q$ and $t_q < 1/q$ proves the stated upper bound, and hence $qt_q^2 \rightarrow 0$. The logarithm of (2) then gives $qt_q = \log q - \log(qt_q) - \log 2 + o(1)$, proving the equivalence in (5).

At the symmetric configuration all angular derivatives vanish by reflection. The four height derivatives agree by rotational symmetry; each is one quarter of $F'_q(t_q)$ and therefore vanishes. These are all seven shape derivatives in the coordinates introduced below. Finally, substituting $t = q^{-2/3}$ in (4) gives $F_q(t) = 4 + o(1)$, which bounds the minimum. $\square$

3 Equilibrium near the limiting packing

Choose an arbitrary sequence of global minimizers with $q \rightarrow \infty$. The convergence theorem [2, Theorem 7] permits congruences and cyclic labels for which the apex is e_3 and the other points are

$$b_i = (r_i \cos \theta_i, r_i \sin \theta_i, -t_i), \quad r_i = \sqrt{1 - t_i^2}, \quad 0 \leq i \leq 3, \quad (6)$$

where $t_i \rightarrow 0$, $\theta_0 = 0$ and $\theta_i \rightarrow i\pi/2$. The four heights and three free angles form a smooth seven-dimensional coordinate chart modulo rotations. Set $\theta_4 = 2\pi$ and $\delta_i = \theta_{i+1} - \theta_i$, using cyclic indices thereafter. Thus $\sum_i \delta_i = 2\pi$ and $\delta_i \rightarrow \pi/2$.

The apex-to-base and base-to-base normalized distances are

$$a_i = 1 + t_i, \tag{7}$$

$$d_{ij} = 1 - t_i t_j - r_i r_j \cos(\theta_j - \theta_i), \quad i \neq j. \tag{8}$$

Write $d_i = d_{i,i+1}$ for adjacent distances and $o_i = d_{i,i+2}$ for opposite distances. Define

$$v_i = a_i^{-q-1}, \quad w_i = d_i^{-q-1}, \quad u_i = o_i^{-q-1}.$$

Since $o_i \rightarrow 2$, a fixed $0 < \rho < 1$ bounds $u_i = O(\rho^q)$. Constants in this section can be chosen uniformly in a fixed small coordinate neighborhood of the limiting packing; the $o(1)$ terms initially refer to the chosen sequence.

Lemma 3. *For the angular currents*

$$J_i = r_i r_{i+1} \sin \delta_i w_i$$

and $J = J_0$, stationarity gives

$$J_i = J + O(\rho^q). \tag{9}$$

The height equations are exactly

$$(1 - t_i^2)v_i = \sum_{j \neq i} d_{ij}^{-q-1} [t_j - t_i(1 - d_{ij})]. \tag{10}$$

Proof. The angular equation at b_i says that $J_i - J_{i-1}$ equals the opposite-pair contribution, which is $O(\rho^q)$ after dividing by q . The equation for θ_0 follows from invariance under a common angular shift, even though that angle is fixed in the chart. Summing at most three consecutive equations gives (9).

Differentiating with respect to t_i and dividing by q gives

$$v_i = \sum_{j \neq i} d_{ij}^{-q-1} \left(t_j - t_i \frac{r_j}{r_i} \cos(\theta_j - \theta_i) \right).$$

Multiply by $1 - t_i^2 = r_i^2$ and substitute (8) to obtain (10). □

Lemma 4. *Along the chosen sequence,*

$$d_i^{-q} \rightarrow 1, \quad J \rightarrow 1, \quad a_i^{-q} \rightarrow 0 \quad (0 \leq i \leq 3). \tag{11}$$

In particular, all t_i are eventually positive and $q \min_i t_i \rightarrow \infty$.

Proof. Initially let $T = \max_i |t_i| = o(1)$. Lemma 2 implies $a_i^{-q} \leq 5$ eventually, and therefore

$$t_i \geq 5^{-1/q} - 1 \geq -\frac{\log 5}{q}.$$

Some gap δ_k is at most $\pi/2$. Hence

$$d_k \leq 1 - t_k t_{k+1} \leq 1 + CT/q,$$

because a negative height has absolute value at most $(\log 5)/q$. It follows that $w_k \geq 1 - o(1)$. Its angular prefactor tends to one, so $J_k \geq 1 - o(1)$, and (9) gives this bound for every J_i . Since $r_i r_{i+1} \sin \delta_i \leq 1$, we also have $w_i \geq 1 - o(1)$. Thus each adjacent energy $d_i^{-q} = d_i w_i$ is at least $1 - o(1)$.

These four terms sum to at least $4 - o(1)$, whereas the total energy is at most $4 + o(1)$ and all other terms are positive. Each adjacent term therefore tends to one and every spoke term tends to zero. This proves (11). In particular $t_i > 0$ eventually, and $q \log(1 + t_i) \rightarrow \infty$ implies $qt_i \rightarrow \infty$. $\square$

4 Relative height balance and moving-center localization

From now on set $T = \max_i t_i > 0$. The positivity established above precedes any logarithmic upper bound on T ; this order matters in the relative estimates below. Put

$$D = \max_i |\delta_i - \pi/2|, \quad \epsilon = \max_i |d_i - d_0|, \quad \Delta = 1 - d_0.$$

Lemma 5. *The angular spread and distance variation satisfy*

$$D = O(T^2 + \rho^q/q), \quad \epsilon = O((T^2 + \rho^q)/q). \quad (12)$$

If $S_i = t_{i-1} + t_{i+1}$, then

$$\frac{v_i}{J} = S_i(1 + o(1)) \quad (0 \leq i \leq 3). \quad (13)$$

Proof. By (9) and $J \rightarrow 1$, the logarithms of the currents differ by $O(\rho^q)$. In the identity

$$\log J_i = \log r_i + \log r_{i+1} + \log \sin \delta_i - (q+1) \log d_i,$$

the radius terms are $O(T^2)$ and the sine term is $O(D^2)$. All d_i tend to one. It follows that

$$\epsilon \leq C(T^2 + D^2 + \rho^q)/q. \quad (14)$$

There are gaps on both sides of $\pi/2$, allowing equality. Using (8) on these gaps gives

$$\min_i t_i t_{i+1} - \epsilon \leq \Delta \leq \max_i t_i t_{i+1} + \epsilon. \quad (15)$$

Consequently $|\cos \delta_i| \leq C(T^2 + \epsilon)$ and $D \leq C(T^2 + \epsilon)$. Substitute (14) into this inequality. Since $D = o(1)$, the term CD^2/q can be absorbed into the left side. This proves the first bound in (12), and substitution into (14) proves the second.

For each fixed i , every adjacent height product is at most TS_i : every edge joins the two parity classes, and one endpoint belongs to $\{i-1, i+1\}$. Positivity and (15) therefore give

$$-\epsilon \leq \Delta \leq TS_i + \epsilon. \quad (16)$$

Each adjacent correction in (10) satisfies

$$|t_i(1 - d_{i,\pm 1})| \leq T^2 S_i + 2T\epsilon.$$

By Lemma 4, $qS_i \rightarrow \infty$. Hence

$$\frac{T\epsilon}{S_i} \leq \frac{CT(T^2 + \rho^q)}{qS_i} \rightarrow 0.$$

Also $w_i/J \rightarrow 1$ by (9) and convergence of the angular prefactors. The opposite contribution in (10), divided by JS_i , is $O(\rho^q T/S_i) = o(1)$. Dividing that equation by JS_i proves (13). $\square$

The error in (13) is relative to S_i , not merely an additive $O(T^3)$. An additive bound alone would not control a parity class whose heights were much smaller than T .

Proposition 6. *Uniformly over all global minimizers, after congruence and relabeling,*

$$\max_i |t_i - t_q| + \max_{1 \leq i \leq 3} |\theta_i - i\pi/2| = o(1/q). \quad (17)$$

Proof. Opposite indices have the same S_i . Taking logarithms in (13) and using the mean-value theorem gives $q(t_0 - t_2) = o(1)$ and $q(t_1 - t_3) = o(1)$. Write $a = t_0$, $b = t_1$. Since $q \min_i t_i \rightarrow \infty$, these additive errors are relative $o(1)$ errors in the positive heights. Thus

$$v_0 = 2Jb(1 + o(1)), \quad v_1 = 2Ja(1 + o(1)).$$

Dividing and taking logarithms yields

$$(q+1)(\log(1+a) - \log(1+b)) = \log(a/b) + o(1). \quad (18)$$

If $a \neq b$, the difference between the two displayed main terms is

$$(a-b) \left(\frac{q+1}{1+\xi} - \frac{1}{\eta} \right),$$

where ξ, η lie between a, b . Since $\xi \rightarrow 0$ and $q\eta \rightarrow \infty$, the factor in parentheses is $q(1 + o(1))$. Equation (18) gives $q(a-b) = o(1)$, also when $a = b$. All four heights are therefore $a + o(1/q)$.

Now $J \rightarrow 1$ and (13) give

$$H_q(a) = o(1), \quad H_q(t) = (q+1)\log(1+t) + \log(2t). \quad (19)$$

We have $a \geq 1/q$ eventually. Using $\log(1+a) \geq a/2$ in (19) gives $a = O((\log q)/q)$, and hence $qa^2 \rightarrow 0$. On the other hand, the exact candidate equation gives

$$H_q(t_q) = (q+1)\log(1-t_q^2) + \log \frac{2}{2+2^{-q}} = o(1). \quad (20)$$

Throughout the small positive interval joining a and t_q , $H'_q(t) \geq q/2$. Equations (19) and (20) therefore imply $q|a - t_q| \rightarrow 0$. Finally (12) gives $qD \rightarrow 0$, because $T = O((\log q)/q)$. Summing at most three angular gaps proves (17) along the chosen sequence.

This conclusion is uniform over all minimizers. Otherwise a sequence violating it would, after the congruences and labels supplied by [2, Theorem 7], contradict the preceding argument. $\square$

5 Strict convexity near the moving pyramid

Lemma 7. *Fix $c > 0$. For every sufficiently large q , the energy E_q is strictly convex in the seven coordinates on the box*

$$|t_i - t_q| < c/q \ (0 \leq i \leq 3), \quad |\theta_i - i\pi/2| < c/q \ (1 \leq i \leq 3), \quad \theta_0 = 0. \quad (21)$$

Proof. All heights in the box are positive for large q , since $qt_q \rightarrow \infty$. The chart is regular and all pair distances stay uniformly away from zero. Constants below may depend on c but not on q or the point in the box. The weights satisfy

$$a_i^{-q-2} \geq c_1 t_q, \quad c_2 \leq d_i^{-q-2} \leq c_3, \quad \sum_{i < j} d_{ij}^{-q-1} \leq C, \quad (22)$$

where the last sum is over the six base pairs. For the first bound, (2) implies the exact identity

$$(1 + t_q)^{-q-2} = \frac{(2 + 2^{-q})t_q}{1 + t_q}(1 - t_q^2)^{-q-1}.$$

This is comparable to t_q by $qt_q^2 \rightarrow 0$; changing t_q by at most c/q changes the weight by a bounded factor. Adjacent distances are $1 + O(1/q) + O(t_q^2)$, so their weights are bounded above and below. Opposite distances tend uniformly to two. These observations establish all of (22).

Let (h, k) be a coordinate variation, with four height components h_i and angle components k_i , where $k_0 = 0$ and $k_4 = k_0$. Differentiating an adjacent distance gives

$$Dd_i[h, k] = A_i(k_{i+1} - k_i) + B_i[h], \quad A_i = r_i r_{i+1} \sin \delta_i \rightarrow 1,$$

where $|B_i[h]| \leq C(t_q + 1/q)|h| \leq Ct_q|h|$ for large q . The cycle difference map $k \mapsto (k_{i+1} - k_i)_{i=0}^3$ is injective when $k_0 = 0$. The inequality $(u + v)^2 \geq u^2/2 - v^2$ therefore gives

$$\sum_i (Dd_i[h, k])^2 \geq c_4|k|^2 - Ct_q^2|h|^2. \quad (23)$$

Every second derivative of a normalized base distance satisfies $|D^2d_{ij}[h, k]| \leq C(|h|^2 + |k|^2)$ uniformly on the box.

Direct differentiation of (3) gives

$$\begin{aligned} D^2E_q[h, k] = q(q+1) & \left(\sum_i a_i^{-q-2} h_i^2 + \sum_{i < j} d_{ij}^{-q-2} (Dd_{ij}[h, k])^2 \right) \\ & - q \sum_{i < j} d_{ij}^{-q-1} D^2d_{ij}[h, k], \end{aligned}$$

with both pair sums restricted to base pairs. Retain only adjacent pairs in the positive sum. Equations (22) and (23) give a positive height coefficient $c_1 t_q - Ct_q^2$, and a fixed positive angle coefficient. Absorbing the t_q^2 term and bounding the last sum yields constants $c_5, C > 0$ for which

$$D^2E_q[h, k] \geq q(c_5(q+1)t_q - C)(|h|^2 + |k|^2). \quad (24)$$

Since $qt_q \rightarrow \infty$, this is positive for every nonzero variation and all sufficiently large q . The coordinate box is convex, which proves the lemma. $\square$

Proof of Theorem 1. Fix $c > 0$. Proposition 6 places every global minimizer, after congruence and labeling, inside (21) for all sufficiently large q . Lemma 2 makes its center $P(t_q)$ stationary. By Lemma 7, this box contains at most one stationary point. Every global minimizer is stationary and must therefore equal $P(t_q)$ in these coordinates. Taking S to be twice a sufficiently large value of q proves the assertion. $\square$

6 Scope and supplementary checks

The proof is asymptotic and supplies no numerical value of S . In particular it does not prove optimality for every exponent above the transition in [3]. The distinction between convergence to a packing and exact optimality at large finite exponents is resolved here by the moving-center localization in Proposition 6.

An accompanying exact-rational checker compares two derivative implementations on sixteen rational configurations, including 5,472 distance-derivative and 2,736 energy-derivative entries, and verifies the equilibrium identities and limiting contact ranks. These are finite regression checks, not proofs of an asymptotic assertion. An optional floating-point Hessian regression likewise gives no certified threshold: some corners of the fixed-width $1/q$ box have a negative Hessian even at $q = 10,000$. The analytic proof requires only eventual positivity, as supplied by (24).

This is an unrefereed manuscript; no absolute priority claim is made. AI assistance was used in literature comparison, mathematical exploration, drafting and computational checks. The author is responsible for the manuscript.

References

- [1] American Institute of Mathematics, *Discrete geometry and automorphic forms*, Problem 1.38, <http://aimpl.org/discreteaf/1/>.
- [2] A. V. Bondarenko, D. P. Hardin and E. B. Saff, *Mesh ratios for best-packing and limits of minimal energy configurations*, Acta Math. Hungar. **142** (2014), 118–131. doi:10.1007/s10474-013-0339-3. arXiv:1212.6211.
- [3] R. E. Schwartz, *Divide and Conquer: A Distributed Approach to Five Point Energy Minimization*, arXiv:2301.05090v4 (2026).
- [4] L. Kryvonos, L. Liehr and M. A. Taylor, *Energy minimization for eight points on the sphere*, arXiv:2609.22077v1 (2026).