hexagon:2611.00490v1

DENSITY-ONE UNIQUENESS FOR MARKOFF–HURWITZ SOLUTIONS

QIYUE TANG

15 November 2026

Abstract. Fix an integer $n \geq 4$. We prove that, as $R \to \infty$, the proportion of sorted positive integer solutions of $\sum_{i=1}^{n} x_i^2 = n \prod_{i=1}^{n} x_i$ with largest coordinate at most R whose largest coordinate is shared by another sorted solution tends to zero.

1. Introduction

Markoff's work on indefinite binary quadratic forms in 1879 and 1880 led to the equation

$$x^2 + y^2 + z^2 = 3xyz.$$

The sorted positive integer solutions are generated from $(1, 1, 1)$ by permutations and Vieta moves, such as $z \mapsto 3xy - z$. Markoff related these triples to forms whose minimum, divided by the square root of the discriminant, exceeds $1/3$ [10, 12]. In 1913 Frobenius asked whether the largest entry determines the sorted triple [7, 1]. The generation of the triples gives a tree, and Frobenius's question asks whether the largest entries label its vertices without repetitions.

Zagier counted the vertices of this tree by the largest entry. He showed that the number of sorted triples with largest entry at most R is asymptotic to $C(\log R)^2$ for an explicit positive constant C [12]. This count assigns one term to each triple, so a largest entry shared by several triples contributes several times. Baragar proved uniqueness for certain arithmetic classes of largest entries. For example, his result applies when either $3z - 2$ or $3z + 2$ is prime, where z is the largest entry [1]. Chen's preprint proves density-one uniqueness in the classical equation, using the action on primitive torus curves and a 3-adic zero-one law [5].

Hurwitz studied positive integral tuples whose sum of squares is a multiple of their product [9]. The specialization in n variables with coefficient n has the all-one solution. The Vieta move again replaces one coordinate by the other root of a quadratic, and descent from the largest entry gives the tree used below. Baragar established the growth exponent for the higher-dimensional equation [2]. Gamburd, Magee and Ronan subsequently obtained an asymptotic formula with positive leading coefficient and growth scale $(\log R)^{\beta_n}$ by comparing the Vieta action in logarithmic coordinates with a linear renewal problem [8]. Their counting theorem provides the scale for asking how often a largest entry is repeated.

Fix an integer $n \geq 4$. For $R \geq 1$, let $T_n(R)$ count the integer tuples satisfying

$$1 \leq x_1 \leq \cdots \leq x_n \leq R, \qquad \sum_{i=1}^{n} x_i^2 = n \prod_{i=1}^{n} x_i.$$

Let $B_n(R)$ count the tuples in this set whose last coordinate equals the last coordinate of a different tuple in the set.

2020 *Mathematics Subject Classification.* Primary 11D45, Secondary 11D72.
Key words and phrases. Markoff–Hurwitz equation, maximum coordinate, density.

Theorem 1. *For every fixed integer $n \geq 4$,*

$$\lim_{R\to\infty} \frac{B_n(R)}{T_n(R)} = 0.$$

If $D_n(R)$ counts the distinct largest entries, then $T_n(R) - B_n(R) \leq D_n(R) \leq T_n(R)$. At most $B_n(R)$ of these entries are repeated. Thus the proportion of distinct largest entries that determine a unique sorted solution also tends to one.

The proof assigns a primitive integer vector to each word of Vieta moves. On most pairs with equal largest entries, these vectors are related by a matrix from a fixed finite set of rational matrices. For a fixed matrix, a positive mass of pairs in the limiting angular distributions forces a local replacement of the initial word. This identifies a common suffix of moves from two fixed seeds. Counting those suffixes on the finite residue orbits reduces the collision estimate to 3-adic separation of the seeds.

2. The Vieta tree and integer addresses

Write $d = n - 1$. All vectors in $\mathbb{R}^d$ are columns unless explicitly written as rows. Superscript t denotes transpose, I denotes the identity matrix, and $\mathbf{1}$ is the column of ones. At coordinate i the Vieta move replaces x_i by $n\prod_{j\neq i} x_j - x_i$ and fixes the other coordinates. The new value is a positive integer because the product of the two roots of the quadratic for x_i is the positive sum of the other squares.

For a sorted positive solution write $y = x_{n-1}$, $z = x_n$, and $P = \prod_{i=1}^{n-2} x_i$. The quadratic f for z satisfies

$$f(y) = \sum_{i=1}^{n-2} x_i^2 + (2 - nP)y^2 \leq n(1 - P)y^2 \leq 0.$$

Equality forces all entries preceding z to equal one, and then z is 1 or $n - 1$. Otherwise the replacement of z is strictly less than y. Also, when z is the unique maximum, the quadratic for any other entry takes the value $\sum_{j\neq i,n} x_j^2 + (2 - n\prod_{j\neq i,n} x_j)z^2$ at z. It is strictly negative because all entries preceding z are strictly smaller. Thus every move at a nonmaximum entry increases the maximum.

We keep the maximum in slot n and start at $a_\varnothing = (1, \ldots, 1, n - 1)$. For $1 \leq i \leq d$, let μ_i swap slots i, n and then apply the Vieta move in slot n. For the finite word $w = i_1 \cdots i_k$ put $a_w = \mu_{i_k} \cdots \mu_{i_1} a_\varnothing$, and let $M(w)$ be its maximum. Descent gives every nonroot sorted solution by sorting some a_w. Moreover, a_w determines w: the largest entry among the first d slots was the previous maximum and identifies i_k, after which the inverse move recovers $a_{i_1\cdots i_{k-1}}$.

With e_i the ith standard column, define

$$R_i = I + e_i(\mathbf{1} - e_i)^t, \quad A_i = R_i^t, \quad V = \mathbf{1}\mathbf{1}^t - (d - 1)I, \quad P_w = A_{i_1} \cdots A_{i_k}.$$

Empty products are identities. Direct multiplication gives $VA_i = R_iV$, $V\mathbf{1} = \mathbf{1}$, and $\mathbf{1}^tV = \mathbf{1}^t$. Set $R_w = R_{i_1} \cdots R_{i_k}$, $v_w = R_w\mathbf{1} = VP_w\mathbf{1}$, and $H(v) = \mathbf{1}^tv$. Each v_w is positive and primitive because each R_i is unimodular. For nonempty w, the coordinate indexed by its first letter exceeds the sum of the others. Inverting that letter and iterating proves that v_w determines w.

Assign degree zero to the parent entries of $a_\varnothing$ and degree one to the maximum. Every new maximum has assigned degree one plus the sum of the degrees of the retained entries. At each stage the maximum degree is one plus the sum of the parent degrees. The vector $q_w = P_w^t\mathbf{1}$ has entries $(n - 2)r_i + 1$, where r_i is the degree in parent slot i: when the maximum becomes parent i, the corresponding entry of q_w is replaced by the sum of all its entries. In particular, if $D(w)$ is the maximum degree, then $H(v_w) = (n - 2)D(w) + 1$. In any outward move the removed entry

is bounded by the product of the retained entries. Hence the new entry lies between $n-1$ and n times that product, with strict upper inequality. Induction gives

$$(n-1)^{D(w)} \le M(w) < n^{D(w)}. \tag{1}$$

3. Counting and the distribution of addresses

Let $\beta = \beta_n > 1$ denote the exponent of [8, Theorems 2 and 13]. For a positive column q and $L > 0$ let $N_q(L)$ count finite words t with $\mathbf{1}^t P_t^t q \le L$. Write $r = \mathbf{1}^t q$ and $s = r - \max_i q_i$. The accelerated scalar renewal theorem of [8, Theorem 13], followed by summation over the initial repetitions of the largest slot, gives

$$N_q(L) \sim c(q) L^\beta, \quad N_q(L) \le C L^\beta c(q), \qquad c(q) \asymp (1 + r/s) r^{-\beta}, \qquad c(q) = \sum_i c(R_i q). \tag{2}$$

The asymptotic is as $L \to \infty$. The relation $\asymp$ denotes upper and lower bounds by positive constant multiples. Constants may depend on the fixed dimension. To see the deacceleration, after sorting the slots the repeated largest-slot moves yield heights $r + As$ for integers $A \ge 0$. Accelerated counts from these seeds are bounded by $C(L/(r+As))^\beta$, and their leading coefficients are bounded above and below by positive multiples of $(r+As)^{-\beta}$ by compactness of the normalized seed simplex and positivity of the GMR eigenfunction. Thus the sum converges and its coefficient is comparable to $\sum_{A \ge 0} (r + As)^{-\beta} \asymp (1 + r/s) r^{-\beta}$. The recursion in (2) follows by partitioning at the first letter. Ties in sorting the seed are resolved by a fixed rule on the slot labels during this formal count.

For words w, t we have $H(v_{wt}) = q_w^t P_t \mathbf{1}$. Thus $c(q_w)/c(\mathbf{1})$ is the limiting fraction of addresses with prefix w. These weights define a probability on infinite words. If a family of prefixes is prefix-free, the total number of finite descendants of height at most L is bounded by CL^β times the sum of its cylinder probabilities. This follows by applying (2) to each prefix.

For a real column z with $\mathbf{1}^t z \ne 0$, write $[z] = z/(\mathbf{1}^t z)$. For a matrix B and a point z of the affine hyperplane $\mathbf{1}^t z = 1$, write $[B]z = [Bz]$ when $\mathbf{1}^t B z \ne 0$. All simplex interiors below are relative to this affine hyperplane. We use $\|z\|_1 = \sum_i |z_i|$, and $\|B\|_1$ for its induced operator norm. Let $\Delta_0 = \{z \ge 0 : \mathbf{1}^t z = 1\}$. The simplices $[R_w]\Delta_0$ shrink uniformly with word length. The ith coordinate of $[R_i]z$ is $(2 - z_i)^{-1}$, and its other coordinates are $z_j/(2 - z_i)$. For a tangent vector v with $\mathbf{1}^t v = 0$, differentiating gives

$$\|D[R_i](z)v\|_1 \le \frac{\sum_{j \ne i} |v_j|}{2 - z_i} + \frac{(1 + \sum_{j \ne i} z_j)|v_i|}{(2 - z_i)^2} = \frac{\|v\|_1}{2 - z_i}.$$

For $i \ne j$, we have $([R_j]z)_i \le 1/2$. Hence a pair of distinct successive letters contracts by at least a factor $2/3$. A constant word of length k has image diameter $2/(k+1)$. The number of disjoint changing pairs or the length of the largest constant run must tend to infinity uniformly. Let $K = \bigcap_{k \ge 0} \bigcup_{|w| = k} [R_w]\Delta_0$, where $|w|$ is the number of letters in w, and let κ be the distribution of the limit under the cylinder law above. Let δ_z denote unit point mass at z. Testing against continuous functions at a fixed cylinder depth and then sending that depth to infinity gives weak convergence

$$L^{-\beta} \sum_{H(v_w) \le L} \delta_{[v_w]} \implies c(\mathbf{1})\kappa. \tag{3}$$

Indeed the finitely many shorter words have vanishing mass, and every other normalized address and its infinite continuations lie in the same shrinking simplex.

Let $\Delta \subset \Delta_0$ be the sorted simplex. For $1 \le j \le d$, let S_j move slot j to slot d preserving the order of the other slots, and set $\gamma_j = S_j R_j$. Then γ_j induces the sorted update, and the accelerated generators are $\gamma_d^A \gamma_j$ with $A \ge 0$ and $j < d$. Denote by ν the accelerated GMR conformal probability at β, and by h its positive eigenfunction normalized by $\int h \, d\nu = 1$.

4. Regeneration and rational overlaps

Put $K_s = K \cap \Delta$. Interiors are relative to the affine hyperplane of coordinate sum one. For a real matrix G, write $[G]z = Gz/(\mathbf{1}^t Gz)$ where the denominator is nonzero. On measures supported in that domain, $[G]_*$ denotes pushforward. For an accelerated letter $a = (A, j)$, write $M_a = \gamma_d^A \gamma_j$.

On Δ, the first map $[\gamma_j]$ is 2/3-Lipschitz in the ℓ^1 norm since $j < d$ and $z_j \le 1/2$. The following A applications of $[\gamma_d]$ are nonexpanding.

The eigenfunction h is C^1 and the eigenvalue at β is 1 [8, Theorem 39]. For a finite product B of accelerated letters put

$$m([B]_{\mathrm{symb}}) = \int \|Bz\|_1^{-\beta} h([B]z)\, d\nu(z), \qquad \mu = h\nu.$$

Here $[B]_{\mathrm{symb}}$ denotes the symbolic cylinder of the specified word. Summation over an additional letter on the right uses the dual eigenmeasure identity and summation on the left uses the eigenfunction identity. Thus m is a stationary symbolic law. Projecting by the limit code gives μ: approximate a continuous test function by its value at any point of each shrinking cylinder and then apply the dual eigenmeasure identity to the resulting integrals. Its support is K_s. Write $\mathrm{sort}(x)$ for increasing rearrangement. Then $\mathrm{sort}([R_i]x) = [\gamma_j]\,\mathrm{sort}(x)$ when slot i has rank j. Every ordinary sorted update is obtained in this way. At each depth the union of sorted word images of Δ is therefore the sorting of the union of full word images of Δ_0. Taking nested intersections gives K_s. A sorted infinite code splits into blocks $d^A j$, $j < d$, whenever it has infinitely many letters other than d. Such codes are dense, and every accelerated cylinder has positive mass.

Lemma 2. *The law m is the concatenation of independent identically distributed nonempty finite words. Every one-letter word has positive probability. For the product B of one such random word, $\mathbb{E}\max(\|B\|_1, \|B^{-1}\|_1)^t < \infty$ for some $t > 0$, where the matrix norm is induced by the vector ℓ^1 norm.*

Proof. Write $p_a(z) = \|M_a z\|_1^{-\beta} h([M_a]z)/h(z)$. The roof is $\log(1 + (A+1)(1 - z_j))$. Its derivative is uniformly bounded since $z_j \le 1/2$. Consequently p_a is bounded by $C(A+2)^{-\beta}$ and $\log p_a$ has a uniform Lipschitz constant. The conditional distribution of a letter given the next k letters is a positive weighted average of p_a on the corresponding cylinder. It follows that the conditional distribution given the infinite future is p_a at its coded point. Its log variation over completions of a k-letter future is $O((2/3)^k)$.

Apply the causal construction of Comets, Fernández and Ferrari [6] to this kernel in reversed time. In their notation let a_k be the infimum over recent histories of length k of the sum of the letterwise infima over their older completions. The log variation estimate gives $a_k \ge \exp(-C(2/3)^k)$. The thresholds are nondecreasing and $b = \prod_{k \ge 0} a_k > 0$. The resulting stationary process X_i is a function of independent uniforms U_j, $j \le i$, and its negative-time sequence $X_{-1}, X_{-2}, \ldots$ has law m by uniqueness in the causal construction.

For integers $s < t$ write $H[s, t-1] = \{U_{s+i} < a_i : 0 \le i < t - s\}$ and let $\mathcal{R}_s = \{U_{s+i} < a_i : i \ge 0\}$ be the event of a cut at s. On $\mathcal{R}_t$, the event $\mathcal{R}_s$ is exactly $H[s, t-1]$, by monotonicity of the thresholds. On $H[s, t-1]$ the symbols in this interval depend only on its uniforms, since the causal look-back never passes s. Thus specifying successive cuts s, t, with no cut between them, and their block symbols, amounts on $\mathcal{R}_t$ to a finite event in the uniforms of $[s, t)$. The cut process is stationary ergodic with intensity $b > 0$, and therefore has infinitely many cuts in both directions. Conditional on $\mathcal{R}_0$, any specified finite list of successive negative blocks has probability equal to the product of their finite-event probabilities: the underlying intervals and the uniforms defining $\mathcal{R}_0$ are disjoint. These negative blocks are identically distributed. Also $\mathcal{R}_0$ is independent of all negative-time symbols. Reversing each complete negative block and reading the blocks backward

therefore gives the original law m. A single-letter block with symbol a has probability equal to the letterwise infimum of the kernel, which is positive by the uniform log variation.

For the gap L between cuts, put $b_k = \prod_{i=0}^{k} a_i$ and $b_{-1} = 1$. On the cut event R_j, the event R_0 is $H[0, j-1]$. Independence of its uniforms from those of R_j, and $\mathbb{P}(R_0) = \mathbb{P}(R_j) = b$, give $\mathbb{P}(R_j \mid R_0) = b_{j-1}$. With $Q(z) = \mathbb{E}[z^L]$ its generating function gives

$$\frac{1-z}{1-Q(z)} = 1 - \sum_{j\geq 1} b_{j-2}(1-a_{j-1})z^j.$$

The series on the right after the minus sign converges beyond 1, and at 1 equals $1-b$. The geometric series in that sum converges at some $z > 1$, proving an exponential tail for L by the generating function of the tails. For a letter a, $c(a) = \log\max(e, \|M_a\|_1, \|M_a^{-1}\|_1)$ is $O(1+\log(A+2))$, since γ_d is a unipotent matrix with square-zero nilpotent part. The conditional expectation of $e^{tc(a)}$ is bounded uniformly by a function $J(t)$ tending to 1 as $t \downarrow 0$. Iterated conditioning bounds the exponential reward of k consecutive letters by $J(t)^k$. Conditioning on the cut at zero costs at most a factor b^{-1}. Cauchy–Schwarz on the event $L = k$ now bounds the conditional exponential reward of a whole block by

$$b^{-1/2} \sum_{k\geq 1} J(2t)^{k/2} \mathbb{P}(L = k \mid \mathcal{R}_0)^{1/2},$$

finite for small positive t. Submultiplicativity gives the claimed matrix moment. □

Lemma 3. *Let μ_{sym} be the average of the coordinate-permutation images of μ. Let $G \in \mathrm{GL}_d(\mathbb{Q})$ and let ρ be a finite measure on K such that $\rho, [G]_*\rho \ll \mu_{\mathrm{sym}}$. For ρ-almost every x there are finite full words u, w and a coordinate permutation Π such that*

$$G = cR_u \Pi R_w^{-1}, \qquad x \in \mathrm{int}([R_w]\Delta_0), \qquad c \in \mathbb{Q}^\times.$$

Proof. It suffices to prove the assertion with K_s and μ in place of K and μ_{sym}. Indeed, the hyperplane-nullity argument below gives zero mass to chamber boundaries. Split the source by the ordering chambers of x and $[G]x$, then conjugate the map by the two sorting permutations on its input and output. The resulting marginals are absolutely continuous with respect to μ, and moving the two permutations through the full words restores the claimed formula.

The block law of Lemma 2 makes μ stationary. Its generated group contains γ_j, $\gamma_d\gamma_j$ for $j < d$, hence $R_d = \gamma_d$, the permutations S_j, and all R_i. To see Zariski density in the determinant-one component, put $X_i = R_i - I$. Write $[X, Y] = XY - YX$ for the matrix commutator. The powers of R_i imply X_i belongs to the Lie algebra of the closure. For distinct indices,

$$Z_{ij} = [X_i, X_j] - X_i + X_j = (e_i + e_j)(e_i - e_j)^t, \qquad [Z_{ij}, X_k] = -2e_k(e_i - e_j)^t.$$

Together with X_k, these give every off-diagonal elementary matrix, whose brackets give all traceless diagonals. If necessary, induce the iid law on first return of the determinant product to $+1$. For a negative support element b, the return products include b^2, cb, bc for any negative c, and a, bab for any positive a. They generate the determinant-one group by the two-coset Schreier generators. The return length has an exponential tail, and the same moment argument as above applies. The classical simple-spectrum theorem for iid products with Zariski-dense generated group and finite logarithmic moment [3, Corollary 9.15(c)] now gives simple exponents. The return induction preserves simplicity for the original law because the singular-value limits along return times scale by their mean length.

The stationary measure gives mass zero to projective hyperplanes. Indeed, if a proper subspace receives positive mass, choose the least dimension of such a subspace and then the subspaces of maximal mass in that dimension. Distinct ones have null intersection, so maximum mass is achieved and its maximizers form a finite nonempty set. Stationarity forces every support

matrix to permute that set, contrary to Zariski density. The same argument applies to stationary flag marginals in the exterior powers, where the determinant-one group acts irreducibly. The independent past and future flags on the bilateral iid space are therefore almost surely opposite. For a fixed j-plane in one flag, failure of transversality with the complementary $(d-j)$-plane is the vanishing of the wedge of their exterior representatives. This is a projective hyperplane condition in the $(d-j)$-th exterior power. At time m, let $F_i^+(m)$ be the dimension-i fast flag of the future right products. Let $F_j^-(m)$ be the annihilator of the dimension-$(d-j)$ fast flag of the transposed past products. The two-sided multiplicative ergodic theorem gives the covariant lines $F_i^+(m) \cap F_{d-i+1}^-(m)$ for $1 \le i \le d$. The top line at time m is the future coding point x_m, by uniform contraction of the positive matrices. The complementary hyperplane is the kernel of the past top transpose line. This line has a strictly positive representative: past transpose limits are nonnegative, and the nearest strictly positive finite past block maps the nonnegative cone into its interior. Such blocks exist almost surely since γ_1^d is positive and its letters have positive one-letter probabilities. In particular the projective chart with this hyperplane at infinity contains all of Δ.

Let x_m be the column on the top line at time m, normalized by coordinate sum one. Choose columns b_j on the other covariant lines at time m, indexed by $1 \le j < d$, and rescale them so that $b_j - (\mathbf{1}^t b_j)x_m$ has Euclidean norm one. This vector is the derivative in the j-th coordinate at zero of the chart $u \mapsto [x_m + \sum_j b_j u_j]$. Choose signs measurably. Write $L_m = B_1 \cdots B_m$ for the source prefix of the iid block matrices $B_1, B_2, \ldots$, with $L_0 = I$. In the charts at m and 0 its projective action is diagonal, with factors $a_j(L_m)$, $1 \le j < d$. These absolute factors are physical derivative norms on the unit tangent directions. With $S_j(m) = -\log|a_j(L_m)|$ we have

$$c_0 \le S_j(m+1) - S_j(m) \le C_m, \qquad \mathbb{E}C_0 < \infty.$$

The lower bound follows from uniform contraction, using a fixed grouping of blocks if necessary to convert norms. For the upper bound let $y = [B]x_{m+1}$ with $B = B_{m+1}$. The derivative of the inverse at y is $\|Bx_{m+1}\|_1(I - x_{m+1}\mathbf{1}^t)B^{-1}$ on the affine tangent space. Its Euclidean norm is at most a dimensional constant times $\|B\|_1\|B^{-1}\|_1$, since $x_{m+1} \in \Delta$. The logarithm of this bound is integrable by Lemma 2. These are stationary clocks with finite positive means χ_j. The target block prefixes N_k have clocks T_i with the same means. For almost every pair $(x, [G]x)$, one can choose typical lifts at both ends by conditioning the iid future on the coded point and then appending an independent iid past. Absolute continuity of both marginals ensures that the bilateral almost sure statements hold for both choices.

The coding is one-to-one off countably many hyperplanes. For a sorted image y, the inverse step inserts $y_d - \sum_{i<d} y_i$ into the sorted list $y_1, \ldots, y_{d-1}$ and normalizes to sum one. The insertion rank j is unique away from ties. Thus the interiors of different ordinary cylinder simplices of the same depth are disjoint, with ambiguities contained in images of tie and coordinate hyperplanes. The same uniqueness holds for accelerated cylinders by the block decomposition. An infinite last-letter run codes e_d, since the simplex images of d^A shrink to that vertex. For an accelerated prefix B, telescoping the conditional kernel gives

$$\mu([B]E) = \int_E \|Bz\|_1^{-\beta} h([B]z)/h(z)\, d\mu(z)$$

for Borel $E \subset K_s$. Divide the weight in the integral by $\mu([B]K_s)$ to obtain the density, relative to μ, of the conditional cylinder law pulled back by $[B]^{-1}$. This density is bounded above and below uniformly in B. To see it, right multiplication by $\gamma_d^A \gamma_j$ sends the column-sum row a to the entries $a_i + (A+1)a_d$, $i < d$, with a_d inserted in slot j. The cone $\max a_i \le a_{d-1} + a_d$ is preserved, and the new row has maximum at most twice its last entry. Thus $\|Bz\|_1 \asymp_d \|B\|_1$ on Δ. If $E \subset K_s$ is a carrier of ρ with $[G]E \subset K_s$, symbolic martingale differentiation now gives $\mu(K_s \setminus [L_m]^{-1}E) \to 0$ for almost every chosen source.

In the initial charts write $[G](u) = Au/(1 + p \cdot u)$, with A invertible. Linearly interpolate the clocks. For a small positive δ and all sufficiently large t set

$$h_*(t) = \max_{A_{ij}\neq 0} S_j^{-1}\big(T_i(t) + \log(|A_{ij}|/\delta)\big).$$

This is an increasing, locally piecewise linear map with $h_*(t)/t \to \max_{A_{ij}\neq 0} \chi_i/\chi_j \leq \max \chi_i / \min \chi_i$. At almost every t, some pair (i, j) attaining the maximum determines its derivative. With $s = h_*(t)$, the derivative of the inverse at s is $S_j'(s)/T_i'(t) \leq C_{\lfloor s \rfloor}/c_0$. Choose a compact set in the regular chart locus where the inverse chart has strictly positive first row on the simplex and the centers lie in the interior of Δ and a cutoff K_0. Call a source interval $[l, l+1)$ bad if its chart at $l+1$ is outside this compact set or if $C_l > K_0$. A target interval $[k, k+1)$ uses its chart at k and the cost at k. Changing variables by h_* bounds the upper density of source-bad selected times by

$$\frac{\max \chi_i}{c_0 \min \chi_i}\mathbb{E}[C_0\, \mathbf{1}_{\{\text{source interval at 0 is bad}\}}].$$

The target bad density is its stationary probability. Both bounds follow by Birkhoff, with the fractional endpoint negligible since stationarity and integrability give $C_l/l \to 0$ by Borel–Cantelli. Compact exhaustion and integrability make both bounds arbitrarily small.

Round the good times to $m = \lceil h_*(t) \rceil$, $k = \lfloor t \rfloor$. The carry $N_k^{-1} G L_m$ in the moving charts is $u \mapsto Bu/(1 + \ell \cdot u)$, where

$$B_{ij} = A_{ij} a_j(L_m)/a_i(N_k), \quad \delta e^{-2K_0} \leq \max_{ij} |B_{ij}| \leq \delta, \qquad \ell_j = p_j a_j(L_m) \longrightarrow 0.$$

The upper bound holds at interpolated times for all entries and equality holds for an active entry. Rounding loses at most $2K_0$ in that entry. On the compact chart set the coordinate ranges of Δ are uniformly bounded and the target chart contains a uniform ball at zero in its interior. Choose δ small enough for the carries to map all of Δ inside Δ with a uniform interior margin. In the moving homogeneous bases, normalize the carry to have top-left entry one, top row $(1, \ell)$ and bottom block $(0, B)$. The compact chart set bounds the basis matrices and their inverses. If $x \in \Delta$ has homogeneous coordinates $r(1, u)$ in the source basis, then r is bounded away from zero uniformly by the choice of charts. Its image in the target basis is $r(1 + \ell \cdot u, Bu)$, whose coordinate sum in the original basis is bounded below for small ℓ and B, since the target center has sum one. Thus the normalized matrices in the original basis range over a compact family with denominators bounded away from zero on Δ. For $\ell = 0$ each such map has positive oscillation on K_s by hyperplane nonconcentration. Compactness gives a uniform positive lower bound, which persists for sufficiently small ℓ. Denote the resulting compact family by $\mathcal{Q}$.

The union Y of the images of K_s under the singular members of $\mathcal{Q}$ that map K_s into itself is compact by continuity, and has zero μ-measure. For the latter assertion, every straight segment of positive length in K follows a rational edge line of a full cylinder. If it crosses from one corner cell to another, their intersection is the shared midpoint of a simplex edge. A line through that point lying on both sides in the simplex must follow the edge. Otherwise descend through cells until a crossing occurs, which must happen by shrinking. Every outer edge belongs to K, since its two corner maps subdivide it at its midpoint. Choose one $W = R_w$ with $[W]\Delta_0 \subset \operatorname{int} \Delta$, possible by a typical sorted point in the chamber interior and full-cylinder shrinking. Then $[W]K \subset K_s$. For any admissible singular H, the image of each outer edge under $[HW]$ is a point or a segment of K, and all nonconstant connecting lines of its column images are rational. In rank two this is the entire range. In higher rank choose three noncollinear column images, which are rational as intersections of two rational lines. Each other column image is one of them or is determined by two rational connecting lines to them. Thus the range is a rational proper subspace. Only countably many ranges are possible, so hyperplane nonconcentration gives $\mu(Y) = 0$.

Delete target times whose tail is in a sufficiently small fixed neighborhood of Y. This costs arbitrarily little density by stationarity, ergodicity and compactness. The compact family was fixed before this deletion, so it still contains the normalized carries on the remaining positive-density set. Were a subsequence to have a singular limit H, the density of $[L_m]^{-1}E$ would imply $[H]K_s \subset K_s$. Indeed the whole-domain interior bound places each image in the interior of the target prefix simplex, where unique sorted decoding implies the target tail belongs to K_s. If a limit image left K_s, continuity and full support would give a relatively open set of positive measure missing those source densities eventually. The actual target tail along this subsequence would consequently approach Y, a contradiction.

Thus the normalized carries are bounded away from singular matrices. If an actual carry is aH for a normalized representative H, then $|a|^d = |\det G|/|\det H|$. Thus the actual carries have bounded entries. All entries lie in one fixed rational lattice since both prefixes are unimodular. An actual carry D repeats infinitely often. On the open domain $U = \operatorname{int}\Delta \cap [D]^{-1}\operatorname{int}\Delta$, the same density-to-inclusion argument gives $[D](K_s \cap U) \subset K_s$. Choose a repetition for which the actual tail is in U, and take a full cylinder containing that tail, with its simplex contained in U.

The coding relation for the full gasket is precisely $wij^\infty \sim wji^\infty$, $i \neq j$: at the first differing letter the two cells meet only at the midpoint, and the two inverse images are the uniquely coded opposite vertices. This is also the coding quotient of the affine Sierpiński gasket, equivariantly for all prefixes and permutations. The self-embedding theorem of Bandt and Retta [4, Proposition 1.2, Theorem 5.1, Corollary 5.2 and Remark 5.7] therefore makes the restriction of $[D]$ to the chosen full cylinder a full prefix replacement followed by a permutation. Two projective transformations that agree on K agree everywhere. Their quotient fixes the vertices and is therefore represented by a diagonal matrix. Fixing each edge midpoint then forces its diagonal entries to be equal. Composing with the original sorted prefixes and moving permutations through the full words gives $G = cR_u \Pi R_w^{-1}$. The chosen input cylinder contains x, and it contains x in its simplex interior after deleting the countable union of boundary hyperplanes. The scalar is rational because the matrices are rational. □

5. Comparison of the two measures

Let ν_{sym} be the average of the coordinate-permutation images of ν. We compare it with κ. For $1 \le j \le d$ write

$$(T_j f)(z) = \|\gamma_j z\|_1^{-\beta} f([\gamma_j]z).$$

The accelerated operator is $\sum_{j<d} T_j \sum_{A\ge 0} T_d^A$, since weighted composition reverses the order of matrix multiplication. Apply the adjoint eigenmeasure equation and subtract its image under T_d^*. The resulting ordinary conformality is $\nu = \sum_j T_j^* \nu$. Symmetrization gives the corresponding equation on the full simplex with the R_i. Indeed, at each permutation of a sorted vector the d possible updated slots correspond bijectively to its d sorted slots, and sorting each updated vector only applies a permutation. The conformal measure is carried by K. Its hyperplane nullity follows from the hyperplane nullity of $\mu = h\nu$ proved in Lemma 3, so the overlaps of the first-level simplices, and all their iterated images, have measure zero.

Write $C = V^{-1} = (\mathbf{1}\mathbf{1}^t - I)/(d-1)$. Conformality and $CR_w = P_w C$ give

$$\nu_{\mathrm{sym}}([R_w]K) = \int_K \|R_w z\|_1^{-\beta}\, d\nu_{\mathrm{sym}}(z) = \int_K (q_w^t C z)^{-\beta}\, d\nu_{\mathrm{sym}}(z).$$

For a positive q choose a slot i with $q_i = \max_j q_j$. Decompose K, except the null point e_i, into the sets $[R_i^A](K\backslash[R_i]K)$ for $A \ge 0$. On the base $K\backslash[R_i]K$ we have $z_i \le 1/2$, since every other first-level simplex has its dominant coordinate in a different slot. The identity $q^t C z = \sum_j q_j(1-z_j)/(d-1)$ gives a lower bound $q_i/(2(d-1)) \ge r/(2d(d-1))$ and an upper bound $r/(d-1)$, where $r = \mathbf{1}^t q$.

Consequently

$$q^tCR_i^Az = q^tCz + \frac{A(1-z_i)s}{d-1} \asymp r + As, \qquad s = r - q_i.$$

The projective denominator cancels the conformal weight in each integral. The base has positive measure uniformly over the finitely many i, since every first-level cylinder has positive mass and the intersections are null. Consequently the integral is comparable to $\sum_{A\geq 0}(r+As)^{-\beta}$, which is comparable to $c(q)$ by (2). The cylinder comparisons extend to Borel sets: at each depth the positive-measure cylinders are disjoint up to null sets, and their diameters tend uniformly to zero. For κ the same null overlap follows because a boundary identification requires an eventually constant address, and each fixed such address has probability zero by (2) along its parabolic run.

$$\kappa \asymp \nu_{\mathrm{sym}}. \tag{4}$$

6. Uniform reduction to rational graphs

For each word w let $z_w = P_w\mathbf{1}$ and $q_w = P_w^t\mathbf{1}$. Thus $z_w = V^{-1}v_w$ and both vectors have ℓ^1 norm $H(w) := H(v_w)$. We use the Euclidean norm and its induced operator and dual norms unless the subscript 1 specifies the sum norm. All norm comparisons have constants depending only on d.

Lemma 4. *There exist $a, b, c > 0$ such that for $L \geq 2$, every nonzero real linear form ℓ and $0 < \delta \leq 1$,*

$$\#\{w : H(w) \leq L, |\ell(z_w)| \leq \delta\|\ell\|\|z_w\|\} \ll L^\beta\delta^a + L^{\beta-b}.$$

For every sublattice $\Lambda \leq \mathbb{Z}^d$ of finite index I,

$$\#\{w : H(w) \leq L, z_w \in \Lambda\} \ll L^\beta \min(I, L^d)^{-c}.$$

Proof. The vectors v_w and q_w range over the same set, since q_w is the v-address of the reversed word. Sorting this set gives the orbit of the maps γ_j from $\mathbf{1}$. Each word in the sorted maps corresponds to exactly one word in the unsorted maps given the stable order of the initial coordinates and the current labels. Unsorting contributes only finitely many permutations to the estimates. In the lattice estimate the change of variables takes a column lattice to the integral lattice $V\Lambda$, of index $|\det V|I$. We therefore prove the estimates for the sorted orbit.

Use the accelerated branches $B_{A,j} = \gamma_d^A\gamma_j$ for $A \geq 0$ and $j < d$. Their norms on the unit sorted simplex are comparable to $A+1$, uniformly in the base point. If B is an accelerated product and a_i are its column sums, then $\max_i a_i \leq a_{d-1} + a_d$ and $\max_i a_i \leq 2a_d$. Indeed right multiplication by $B_{A,j}$ inserts a_d in slot j among the numbers $a_i + (A+1)a_d$, $i < d$. The two assertions now follow by induction from the identity matrix. In particular, for increasing nonnegative x,

$$(2d)^{-1}\|B\|_1\|x\|_1 \leq \|Bx\|_1 \leq \|B\|_1\|x\|_1.$$

There are $O(X^\beta)$ accelerated words of norm at most X by the scalar count from $\mathbf{1}$.

There is a finite collection of accelerated blocks of one common length such that for each unit form ℓ one block sends ℓ under transpose to a form whose coefficients all have the same sign and absolute values bounded below by a fixed positive constant. To construct these blocks, let $v > 0$ be the Perron vector of γ_1. The span of its images under the accelerated semigroup is invariant under γ_1 and $\gamma_d\gamma_1$, hence under γ_d. It contains $(\gamma_d - I)v$, a nonzero multiple of e_d. For $0 \leq k < d$, $\gamma_1^k e_d$ has its first nonzero coordinate in position $d-k$ with entry 1. The resulting d vectors span. Select images D_iv spanning the space. Multiply each D_i on the right by a sufficiently large power of γ_1 to a common length. The Perron limit of each such product has column directions D_iv, which proves the assertion by compactness of the unit sphere.

For real s near β use the scalar positive eigenfunction h_s, eigenmeasure ν_s, and eigenvalue λ_s. On the outer prefix B of length k the probability is

$$\lambda_s^{-k} \int \|Bx\|_1^{-s} h_s([B]x)\, d\nu_s(x), \qquad \nu_s(h_s) = 1.$$

These probabilities are consistent by the eigenmeasure identity, and are comparable to $\lambda_s^{-k}\|B\|_1^{-s}$. At each disjoint trial of the fixed block length, conditional on the outer prefix B, the selected block for $B^t\ell$ has probability bounded below. This follows from norm submultiplicativity and the two-sided cylinder comparison. If a trial succeeds, all later directions lie off $\ker \ell$. More quantitatively their normalized distance from that hyperplane is $\gg \|B\|_1^{-d}$, because $\|(B^t)^{-1}\| \ll \|B\|_1^{d-1}$ by unimodularity and the trial blocks have bounded norm. At $s < \beta$ sufficiently close to β, the geometric decay of the all-failure probabilities dominates λ_s^k. Thus the sum of $\|B\|_1^{-s}$ over products with no completed successful trial is bounded uniformly in ℓ. This gives the hyperplane bound from any sorted seed of height r in the form $O((L/r)^s)$, with the count zero for $L < r$.

At $s = \beta$ the conditional norm ratio of one fixed-length trial has a uniformly bounded positive moment. For one accelerated letter this follows from the cylinder comparison and the tail $(A+1)^{-\beta}$, and iteration gives the fixed-length assertion. By Hölder's inequality, a sufficiently small positive power of the norm ratio multiplied by the indicator of trial failure has conditional expectation at most some $\rho < 1$. Stop at the first completed unsuccessful trial crossing norm T. The probability of such a stop is $O(T^{-a_0})$ for some $a_0 > 0$, by summing the killed norm moments over the number of trials. Take T to be a sufficiently small constant times $\delta^{-1/d}$. Any successful trial starting below T excludes the slab in the lemma. The remaining words either have no completed successful trial, including words terminating inside a trial, or extend an unsuccessful crossing prefix. The former words are covered by the subcritical sum. Each crossing prefix B has at most $O((L/(r\|B\|_1))^\beta)$ inner continuations from a seed of norm r, by the scalar count and the norm comparison. Summing yields $O((L/r)^\beta \delta^{a_0/d} + (L/r)^s)$.

For the lattice assertion take $T = \frac{1}{2} I^{1/d}$ and $R = \sqrt{T}$. For bounded I the scalar count suffices. We may replace T by $\min(T, L)$ and increase constants. Given a sorted seed of norm $r_0 < R$, stop at the first inner prefix taking the norm to at least R. For a stopping vector u with $r \le \|u\|_1 < 2r$ and $\|u\|_1 \le T$, any accelerated outer completion B has norm $O(L/r)$. For each such B, the vectors in $B^{-1}\Lambda$ of norm at most T span a proper subspace, since otherwise their determinant would be nonzero, divisible by I, and smaller than I. The hyperplane bound for the stopping vectors therefore gives a total in this bin of $O(L^\beta r^{s-\beta} r_0^{-s})$.

A stopping vector of norm greater than T requires an acceleration of size $> T/R - 2$. There are $O(1 + \log R)$ steps before crossing R, since every step multiplies the norm by at least $3/2$. Under the forward Gibbs probabilities

$$\|B_{A,j}x\|_1^{-\beta} \frac{h_\beta([B_{A,j}]x)}{h_\beta(x)}$$

the total mass of such stops is $O((1+\log R)(T/R)^{1-\beta})$. A prefix ending at norm r' has probability comparable to $(r'/r_0)^{-\beta}$ and at most $O((L/r')^\beta)$ completions. Thus the overshoot contribution is bounded by its mass times $O((L/r_0)^\beta)$. Words ending below R number $O((R/r_0)^\beta)$. Since $T \le L$ and $R^2 = T$, this is bounded by $O((L/r_0)^\beta R^{-\beta})$.

In all the preceding estimates, remove the initial unaccelerated run by summing over the seeds $\gamma_d^k \mathbf{1}$, whose norms are comparable to $k+1$. Choose $s > 1$, so both seed sums with exponents s and β converge. In the lattice estimate seeds of norm at least R contribute $O(L^\beta R^{1-\beta})$ by the scalar count. Dyadic summation over $r \ge R$ and absorption of the logarithm into a smaller power of R give the desired bound. □

7. Logarithmic interpolation

Set $\sigma = n^{1/(n-2)}$. For the frame $P = P_w$, let $y_i = \sigma x_i$ be its parent labels and let $s = \sigma M(w)$ be its maximum. Define the linear form ℓ_P by $\ell_P(Pe_i) = \log y_i$. The exact Vieta recurrence in these coordinates is $s = \prod_i y_i - a$, where a is the scaled coordinate removed in the preceding Vieta move. At the first child $a = \sigma$. Thus, putting $\Delta_P = \sum_i \log y_i - \log s$,

$$\ell_{PA_i} - \ell_P = -\Delta_P e_i^t P^{-1}, \qquad 0 \le \Delta_P \ll \exp(-c_0 \min_i q_i), \qquad q = P^t \mathbf{1}. \tag{5}$$

Indeed the new parent in slot i is s and its column is the sum of the old columns. Also $a \le \max_i y_i$, and the degree bounds give $\log y_i \asymp q_i$. Cancelling a largest y_i bounds $a/\prod_i y_i$ by $\sigma^{-(d-1)} < 1$ and by $\exp(-c \min_i q_i)$. On this interval $-\log(1-t) \le t/(1-\sigma^{-(d-1)})$, which gives the defect bound. All constants below may depend on the fixed dimension.

We record the growth gap needed to control this variation. Let ξ_m denote the counting exponent for the m-coordinate row system, with the matrices R_i defined in dimension m. In the notation of the counting theorem this exponent is β_{m+1}. The two-coordinate system has exponent 2 by the Euclidean algorithm and the elementary count of primitive pairs. For $m \ge 3$ use the scalar count of [8]. Then $\beta = \xi_d > \xi_{d-1}$. In fact a word in the first $m = d-1$ letters, lifted to dimension d, has operator norm at most a fixed multiple of its original norm. Its upper-right column in the lift is $(B\mathbf{1} - \mathbf{1})/(m-1)$ if B is the original matrix, by multiplication of the block matrices. Appending the extra letter as a delimiter preserves this norm comparison. The scalar lower asymptotic implies divergence of the sum of inverse ξ_m powers of the word norms. Select finitely many delimited blocks with that sum greater than one. The inequality persists at an exponent $\xi_m + \varepsilon$. All concatenations are distinct words and their norms are bounded by the products of block norms. Thus the corresponding Dirichlet series in dimension d diverges at $\xi_m + \varepsilon$, and the scalar upper bound gives $\xi_d \ge \xi_m + \varepsilon$.

From a seed q of height $H = \mathbf{1}^t q$ with all $q_i \ge \theta H$, where $0 < \theta \le 1/d$, let $a_* = \xi_{d-1}$ and $g = \beta - a_* > 0$. The number of continuations of height at most T whose smallest coordinate is at most h is

$$\ll \theta^{-\beta} H^{-\beta} h^g T^{a_*} \qquad (h \ge \theta H). \tag{6}$$

For a final coordinate of value $b \le h$, cut at its last update. Its update makes b the previous total, and the new total lies between b and $2b$. The other $d-1$ coordinates, translated by $b/(d-2)$, thereafter follow the lower-dimensional row system and have initial values at least $b/(d-2)$. In a dyadic bin for b, there are $O(\theta^{-1}(b/H)^\beta)$ cut prefixes and $O((T/b)^{a_*})$ suffixes, by scalar counting and monotonicity from a constant lower seed. Summing uses $g > 0$. If a coordinate is never updated, the bound $O((T/(\theta H))^{a_*})$ suffices, since $h \ge \theta H$.

For a row state q of total r put $\omega(q) = 1 + r/(r - \max q_i)$. The same count weighted by $\omega(q)$ is bounded by

$$O\big(\theta^{-\beta-2} H^{-\beta} h^g T^{a_*} (1 + \log(2T/h))\big) \quad (\theta H \le h \le T).$$

To see this, partition by the start of the terminal run of maximum updates. At the start of any noninitial run the ratio $r/(r - \max q_i)$ is at most 3: after updating a nonmaximum the latest total is less than twice the preceding total and the new sum of the other coordinates is at least half that preceding total. At the initial start the ratio is $O(\theta^{-1})$. The sum of weights for run lengths with final height at most T is $O(\theta^{-2}(T/r)^2)$. The possible run starts in a dyadic bin at height r are bounded by $O(\theta^{-\beta} H^{-\beta} \min(r^\beta, h^g r^{a_*}))$. Since $a_* \ge 2$ and $\beta > 2$, the weighted dyadic sum gives the displayed estimate, with a logarithm only if $a_* = 2$.

The chronological probability of a prefix extending the current seed is $c(q_{\mathrm{new}})/c(q)$, where $c(q) \asymp \omega(q) r^{-\beta}$. Denote this probability by $\mathbb{P}_P$. At the first crossing of $T \ge H$, the total lies between T and $2T$. By the weighted bound, for any fixed $0 < \kappa < g$, the minimum coordinate

m_T there satisfies

$$\mathbb{P}_P(m_T \le h) \ll \theta^{-\beta-2}(h/T)^{\kappa}. \tag{7}$$

The probability vanishes for $h < \theta H$. Enlarging the constant gives the bound for $h > T$ as well.

Choose $0 < p < \min(1, \kappa/d)$. Write the successive frames after P as $P_t = PD_t$, where D_t is the product of the first t continuation matrices. For a frame matrix P write $H(P) = \mathbf{1}^t P \mathbf{1}$. Within the band $2^j H \le H(P_t) < 2^{j+1}H$, nonnegativity gives $\|D_t\| \ll 2^j/\theta$ and unimodularity gives $\|D_t^{-1}\| \ll (2^j/\theta)^{d-1}$. There are $O(2^j/\theta)$ steps in the band, since total increases by at least $(d-1)\theta H$. By monotonicity of the minimum, the defect at every such step is at most $C\exp(-c_0 m_{2^{j-1}H})$, using the initial minimum for the initial band. Integration of (7) gives

$$\mathbb{E}_P e^{-tm_T} \ll_t \theta^{-\beta-2}(H/T)^{\kappa} e^{-t\theta H/2} \quad (t > 0).$$

Multiply (5) on the right by any fixed real matrix U. Subadditivity of the pth power and summation of $2^{j(dp-\kappa)}$ give, for fixed constants $C, B, c_1 > 0$,

$$\mathbb{E}_P \left(\sum_t \|(\ell_{P_{t+1}} - \ell_{P_t})U\| \right)^p \le C\theta^{-B} \|P^{-1}U\|^p e^{-c_1\theta H}. \tag{8}$$

Thus the row forms converge almost surely to a measurable row form ℓ_∞. The bound $\ell_P(P\mathbf{1}) \gg \|P\mathbf{1}\|$ also gives $\|\ell_\infty\| \gg 1$.

We next prove a uniform small-value estimate. A return state means a frame just after updating a nonmaximum parent, so $r/(r - \max q_i) \le 3$. From any such state, a return block consists of A maximum updates followed by one nonmaximum update. Its probability is comparable to $(A+1)^{-\beta}$: the ratio of final to initial height is comparable to $A+1$, and both return states have bounded ω. Therefore the height ratio in any fixed number of return blocks has a uniformly bounded moment of every order strictly below $\beta - 1$.

We show that from every return state there is a uniformly positive-probability escape for every nonzero form argument k. In the frame P, fix the current maximum s and $d-2$ of the old parents, including a largest old parent, leaving two old parents with labels b, c to move. Let $f_1 = s, f_2, \dots, f_{d-1}$ be the frozen labels. Use $f_{i,v}, b_v, c_v$ for their vectors expressed in the original parent coordinates. Put

$$C = \prod_i f_i, \quad S = \sum_i f_i^2, \quad \lambda = (C + \sqrt{C^2 - 4})/2, \quad L_0 = \log\lambda, \quad \delta_0 = \log(C/\lambda), \quad u = \sum_i f_{i,v}.$$

The alternating mutations of the two moving coordinates satisfy

$$y_{j+1} = Cy_j - y_{j-1}, \qquad y_0 = b, \quad y_1 = c, \qquad y_j = A\lambda^j + B'\lambda^{-j}, \qquad AB' = S/(C^2-4).$$

The product identity follows by inserting the solution into $y_0^2 + y_1^2 - Cy_0y_1 = -S$, and $A + B' = b$ proves that both coefficients are positive. The corresponding vectors satisfy $v_j = c_v + (j-1)u$ for $j \ge 1$ and $v_{-j} = b_v + ju$ for $j \ge 0$, by the parent updates. The vectors $f_{i,v}$ and c_v form a unimodular basis with bounded matrix and inverse.

After a long alternating segment in either direction, update every frozen label, choosing f_i first. The limiting linear forms in the original parent coordinates, denoted by $G_{i,+}, G_{i,-}$, satisfy

$$G_{i,\pm}(f_{j,v}) = \log f_j - \delta_0 \mathbf{1}_{i=j}, \quad G_{i,\pm}(u) = L_0, \quad G_{i,+}(c_v) = L_0 + \log A, \quad G_{i,-}(b_v) = \log B'.$$

The difference of consecutive moving vectors is u in the direction of travel, and their logarithmic difference tends to L_0. Thus the first removed frozen vector absorbs the difference $\log C - L_0$. Later updates add parent vectors and, in the limit, add their logs.

Write $r = d-1$ for the number of frozen labels. The estimates $C \ge 2s$, $\lambda \ge s$, $s^{1-2r} \le A, B' \le s$, and $s^{-2r} \le \delta_0 \le s^{-2}$ follow from $f_i \in [\sigma, s]$ and the product identity. Differences of positive-arm

forms give δ_0 times all coordinate differences on the frozen span. For $i \neq 1$ the difference of the opposite arms vanishes on that span and evaluates on c_v to

$$\log\left(\frac{\lambda^2 S}{(C^2-4)s^2}\right) \geq \delta_0/2.$$

Here $b_v + c_v = 2f_{1,v} - u$, and the bound follows from $S \geq s^2 + (r-1)$, $\lambda^2 > C^2 - 4$ and $\delta_0 \leq s^{-2}$. The common value on u is $L_0 > 0$. It follows quantitatively that $\max_{i,\pm} |G_{i,\pm}(x)| \gg \delta_0 \|x\|$: the differences bound the moving coordinate and all frozen coordinate differences, and a single form bounds the remaining coefficient of u, since $\|G_{i,\pm}\| \ll L_0$.

At the end of m consecutive alternating mutations, the absolute indices of the two moving labels are $m + O(1)$, with the smaller absolute index at least $m - 1$. The relative errors from keeping only the corresponding dominant Pell terms are $O(s^{2r}\lambda^{-2m+2})$, since each of A/B' and B'/A is at most s^{2r}. As $\lambda \geq s$ and $r \geq 2$, the errors in the two logs are $O(s^{4r-2m+1})$. Both moving values are $\gg s^{m-2r}$. Each of the next $d-1$ mutations removes a frozen entry at most s while retaining two entries at least as large as the smaller moving value. Hence each nonlinear log defect is $O(s^{4r-2m+1})$. The new maximum column is the sum of the surviving columns. The target agrees with the ideal logs on those columns and therefore agrees with their sum on the new maximum column. Induction over the activations gives terminal residuals $O_d(s^{4r-2m+1})$ on the parent columns. The terminal relative parent matrix has entries $O_d(m)$ and determinant one, so its inverse has norm $O_d(m^{d-1})$ by cofactors. The terminal row error in the original parent coordinates is consequently $O_d(m^{d-1}s^{4r-2m+1})$.

Choose a fixed large m so that this error is smaller than a fixed fraction of the spanning bound. The two moving columns at the cusp have masses comparable to mH, since $H \leq \mathbf{1}^t P u \leq 2H$. Each activated parent is replaced by the preceding maximum column, which also has mass comparable to mH. Thus the terminal row entries lie between $c_d mH$ and $C_d mH$ for some $c_d > 0$. Apply (8) at the terminal frame with $U = P$. Its exponential decreases with mH while $\delta_0 \geq e^{-C_d H}$ by the degree bounds. Increasing m once more makes the probability that subsequent variation exceeds the remaining spanning margin $O(e^{-cH})$. In particular, for each $k \neq 0$, one of these words ensures

$$|\ell_\infty(k)| \geq e^{-C_2 H}\|k\| quad \text{outside conditional mass } O(e^{-cH}).$$

We used $\|P^{-1}k\| \geq \|k\|/\|P\|$. Each chosen word comprises exactly $d+1$ return blocks: two initial nonmaximum updates, a maximum run, then the $d-1$ activations. Its probability is bounded below uniformly, because its finitely many prescribed waiting times are bounded and both endpoints are return states.

At any return trial select the corresponding escape word for the current frame and fixed k. Kill the trial process when the selected word occurs. Hölder's inequality, the positive escape probability and the uniform height ratio moment show that, for some $\alpha > 0$, the conditional expectation of the αth power of the height ratio times the failure indicator is at most $\rho < 1$. Stop upon crossing height $U_0 = t/C_2$. For a failed stop at trial j, condition successively on the preceding trials. The expected αth power of the stopped height divided by H, restricted to failure through trial j, is at most ρ^j. Summing bounds the failed stopping mass by $(H/U_0)^\alpha \sum_{j\geq 1} \rho^j$. If a success occurs before crossing, the probability of exceptional later variation summed over the disjoint first-success cylinders is $O(e^{-cH})$. This proves, from return states of height H,

$$\mathbb{P}_P\big(|\ell_\infty(k)| < e^{-t}\|k\|\big) \ll e^{-cH} + (H/t)^\alpha \qquad (t \geq C_2 H). \tag{9}$$

Lemma 5. *For $0 < \eta \leq 1/2$, $M \geq M_d$ and $0 < E \leq 1$, there is a finite prefix cut with maximal height*

$$H_* \ll \eta^{-S_0}(1 + \log M + \log(1/E))$$

and an exponent $D \ll \eta^{-S_0}$ independent of E. Apart from $O(\eta L^\beta + H_*^\beta)$ words of height at most L, each word w extends a retained cut prefix P and satisfies

$$|\ell_P(k)| \geq cM^{-D}\|k\| \quad (0 < \|k\| \leq M, k \in \mathbb{Z}^d), \qquad |\log(\sigma M(w)) - \ell_P(z_w)| \leq E\|z_w\|.$$

Proof. Fix $0 < \kappa < \min(g, \beta - 1)$ in (7). Take $\varepsilon = (\eta/C_3)^{1/\kappa}$ with C_3 sufficiently large. At the first crossing of height T from the root discard states with minimum at most εT. Then wait for the first nonmaximum update, discarding paths that first make at least $\lceil \varepsilon^{-2} \rceil$ maximum updates. The first discard costs $O(\varepsilon^\kappa)$ by (7). The second costs $O(\varepsilon^{\beta-1})$, since $c(q_A)/c(q) \ll (1 + As/r)^{1-\beta}$ for A consecutive maximum updates from a state of height r with $s = r - \max q_i$, and here $s/r \gg \varepsilon$. Thus the discarded mass is at most $\eta/(100d)$ after fixing C_3. Both discards are unions of prefix cylinders. The retained return cut has minimum at least εT and height between T and $C\varepsilon^{-2}T$, so its balance is at least $c\varepsilon^3$. Words that terminate before the cut or a discarded prefix have height $O(\varepsilon^{-2}T)$.

For a rational subspace W, define its height as the Euclidean norm of the primitive integral wedge spanning its top exterior power. If E is the integer column matrix of a basis of $W \cap \mathbb{Z}^d$, the orthogonal projector is $\Pi_W = E(E^tE)^{-1}E^t$. The squared norm of the primitive wedge is $\det(E^tE)$ by Cauchy–Binet, so this integer clears the denominators of the projector. If two subspaces W, W' have heights at most Q, put $S = 2I - \Pi_W - \Pi_{W'}$. Its kernel is $W \cap W'$ and a denominator $q \leq Q^4$ clears all entries. The product of its positive eigenvalues is at least q^{-d} by the characteristic polynomial. Its norm is at most 2. Consequently

$$\operatorname{dist}(x, W \cap W') \ll Q^{2d}(\operatorname{dist}(x, W) + \operatorname{dist}(x, W')).$$

Taking independent rows of qS, forming their wedge and taking the primitive Hodge dual shows that a nonzero intersection has height $\ll Q^{4d}$.

Exclude neighborhoods of rational subspaces in order of dimension $j = 1, \ldots, d-1$. Use height bounds M^{a_j}, where $a_{d-1} = 1$ and $a_j = 5da_{j+1}$. The neighborhood radius at stage j is M^{-b_j}. Set

$$e_1 = 2da_1 + c_d, \qquad e_j = b_{j-1} + 2da_j + c_d \ (j > 1),$$
$$F = C_4\eta^{-s}, \qquad b_j = F^2 e_j, qquad T_j = Fe_j \log M,$$

where $s > 3/\kappa + 2/\alpha + 3$ and c_d, C_4 are fixed sufficiently large. At the return cut for T_j, apply (8) to discard points with $\|\ell_\infty - \ell_P\| > cM^{-e_j}$. Its conditional cost is at most

$$C\varepsilon^{-3B}(C\varepsilon^{-2}Fe_j \log M)^{p(d-1)} M^{pe_j} \exp(-c\varepsilon Fe_j \log M),$$

using the cofactor bound on P^{-1}. This is at most $\eta/(100d)$ since $e_j = O(F^{2(j-1)})$ and εF is a sufficiently large fixed power of η^{-1}. The remaining directions $\ell_\infty/\|\ell_\infty\|$ in each cylinder lie in a patch of diameter at most M^{-e_j}, by the lower bound on the limiting norm. Their norms are bounded above by $U_j = C(\varepsilon^{-2}Fe_j \log M)^d$, again by the cofactor bound and the parent logarithm estimates.

In a patch containing a direction retained at all previous stages there can be at most one allowed j-plane whose neighborhood meets the patch. Indeed otherwise that retained direction would be within $O(M^{-e_j} + M^{-b_j})$ of two planes. By the separation inequality it would be within $O(M^{2da_j - e_j})$ of their intersection. A zero intersection is impossible for a unit vector. A positive-dimensional intersection has smaller dimension l and height at most M^{a_l} when M_d is large, and its radius M^{-b_l} covers that direction. If the single allowed plane exists, choose a real unit vector k perpendicular to it. A direction in its neighborhood has $|\ell_\infty(k)| \leq U_j M^{-b_j}$. Apply (9) on the cut cylinder at $t = b_j \log M - \log(2U_j)$. The height is at most $C\varepsilon^{-2}Fe_j \log M$ and $t \geq b_j \log M/2$ for large fixed C_4. Thus the conditional loss is $O(e^{-cT_j} + (\varepsilon^{-2}/F)^\alpha) \leq \eta/(100d)$. Sum the cylinder masses and then the stage losses. The scales e_j, b_j are bounded by a fixed power of η^{-1}. Taking $D = b_{d-1}$ and using the hyperplanes orthogonal to primitive integer vectors of norm at most M

shows that the total boundary mass where $|\ell_\infty(k)| < 4cM^{-D}\|k\|$ for some such vector is $O(\eta)$ for a fixed $c > 0$.

Now take a deeper return cut with

$$T = C_5F((D+1)\log M + \log(1/E) + 1).$$

At each retained cylinder, (8) gives $\|\ell_\infty - \ell_P\| \le cM^{-D}$ with conditional probability at least $3/4$. If $|\ell_P(k)| < cM^{-D}\|k\|$ for some tested vector, at least three quarters of the cylinder belongs to the boundary exceptional set. Discarding all such cylinders costs $O(\eta)$ in total. For each remaining P, also discard the first continuation for which the accumulated norm of the increments multiplied on the right by P exceeds $\tau = c'E\theta H$, with $\theta = \min q_i/H$ and $c' > 0$ small. These first exits are prefix-free. By (8) their conditional mass is at most $C\theta^{-B}\tau^{-p}e^{-c_1\theta H} \le \eta$ when C_5 is large. Every remaining endpoint has $z_w = Pz'$ with $z' \ge 0$ and $\|z'\|_1 \le \|z_w\|_1/(\theta H)$. Its partial row variation yields an error at most $C\tau\|z'\|$. The terminal Vieta defect is $O(e^{-c_0\theta H})$, which is at most $E\|z_w\|/2$ upon increasing C_5. This gives the required estimate at the actual endpoint. The uniform scalar continuation bound charges every prefix discard by $O(L^\beta)$ times its boundary mass. The endpoints stopping before the cut and the number of faces are both $O(H_*^\beta)$ by scalar counting. The cut estimates give the required polynomial bounds. □

Proposition 6. *For each $\epsilon > 0$ there is a finite set $\mathcal{F} \subset \mathrm{GL}_d(\mathbb{Q})$ with the following property. In any collection of pairs of words of height at most L with equal maxima and both projections injective, at most $\epsilon L^\beta + o(L^\beta)$ pairs lie off the graphs $z_{w'} = Gz_w$, $G \in \mathcal{F}$. The remainder is uniform in the collection.*

Proof. Write a rational matrix uniquely as A/q with $q > 0$ and the entries of A and q having joint gcd one. Its height is $Q = \max(q, \max_{i,j}|A_{ij}|)$. For a graph of height $Q \le CL^d$, Lemma 4 gives the uniform bound

$$\#\{z : Gz \text{ and } z \text{ are addresses of height at most } L\} \ll L^\beta Q^{-b_1} + L^{\beta-b_2} \tag{10}$$

for fixed $b_1, b_2 > 0$. If $q \ge Q^{1/2}$, the integrality lattice of G has index divisible by q: the image of $\mathbb{Z}^d$ in $(\mathbb{Q}/\mathbb{Z})^d$ under G has exponent q. The lattice bound applies. Otherwise a row of G has norm at least $Q^{1/2}$. Remove the ball of radius $LQ^{-1/4}$ by scalar counting and apply the slab bound to that row on the remaining addresses, using $\|Gz\| \ll L$.

Fix a small loss η and take $M = L^{d+2}$ in Lemma 5. Choose $E = L^{-A-1}$ with $A > (d+2)D+d+2$, where D is independent of E. Outside $O(\eta L^\beta) + o(L^\beta)$ pairs, both endpoints are in retained faces and have absolute interpolation errors $O(L^{-A})$. For each face pair let W be the span of its actual paired vectors. If the first projection has rank $r < \dim W$, select $r+1$ independent pairs, the first r of whose left vectors are independent. Signed r-minors give a nonzero integral relation among the left vectors, with coefficients $O(L^r)$. The same combination on the right gives a nonzero integer vector k by independence of the pairs. It satisfies $\|k\| = O(L^{r+1})$ and $|\ell_{P'}(k)| = O(L^{r-A})$ by equality of the maxima. This contradicts the nonresonance bound. The reverse projection is also injective on W. Thus $\dim W \le d$. If it is less than d, the left endpoints lie in a real hyperplane, so the slab bound with $\delta \downarrow 0$ makes the piece negligible. Otherwise W is the graph of an invertible rational matrix of height $O(L^d)$ by Cramer's rule. There are $O_\eta((\log L)^{2\beta})$ face pairs. By (10), all graphs with height above $(\log L)^K$ have total negligible count for sufficiently large fixed K.

Partition the remaining graphs into dyadic bands $Q_j/2 < Q \le Q_j$, where $Q_j = 2^j$. For these bands use new face cuts. Fix $B > d^2 + 2$. The slab bound and scalar small-ball bound imply that any subset of at least L^β/Q_j^B addresses has d vectors with smallest singular value at least $cLQ_j^{-K_1 d}$ for a fixed large K_1. In fact each hyperplane has at most $CL^\beta t^a + CL^{\beta-b}$ points in the absolute slab of width tL after possibly decreasing a by splitting off the radius $\sqrt{t}L$. For $t = Q_j^{-K_1}$ and j sufficiently large this bound is smaller than the subset size uniformly over $Q_j \le 2(\log L)^K$.

Choose the vectors successively at distance at least tL from the previous span. Their determinant has absolute value at least $(tL)^d$, and all other singular values are $O(L)$.

Use losses $\eta_j = j^{-2}$, normal bound $M_j = C_d Q_j^2$, nonresonance threshold $\omega_j = cM_j^{-D_j}$, and relative error $E_j = c'\omega_j Q_j^{-K_1 d-4}$ in Lemma 5. The heights and numbers of faces are polynomial in j and independent of L. Delete endpoints failing their band cut, at total cost $O(L^\beta \sum_{j\ge j_0} j^{-2}) + o(L^\beta)$ by injectivity and $j = O(\log\log L)$. For each graph and new face pair merge all occurrences inherited from the initial cut. Call the resulting piece large if its cardinality is at least L^β/Q_j^B. On a large piece, the singular value bound and the two relative error bounds yield $\|\ell_P - \ell_{P'}G\| \ll E_j Q_j^{K_1 d}$. Two distinct large graphs $G = A/q$ and $G' = A'/q'$ in one face pair would give a nonzero integer column $k = (q'A - qA')e_i$ of norm at most M_j, with

$$|\ell_{P'}(k)| \ll E_j Q_j^{K_1 d+2} < \omega_j$$

for sufficiently small fixed c'. This contradicts nonresonance. There is thus at most one large graph per face pair.

There are $O(Q_j^{d^2+1})$ possible matrices in the band. Hence the small pieces contribute at most $L^\beta Q_j^{d^2+1-B}$ times a polynomial in j. The large pieces contribute at most a polynomial in j times the bound in (10). The exponential band tails in both estimates are summable, and the $L^{\beta-b_2}$ error remains negligible after summation over $j = O(\log\log L)$. First choose η and then j_0 so that the total normalized loss is less than ϵ. The finite set consists of all invertible rational matrices below the resulting fixed height threshold. Its choice is independent of L and of the paired collection. □

8. Equal maxima from two seeds

Lemma 7. *Let $a = a_w$ be the framed solution of a finite word w, and let b be a distinct positive integral solution with its unique maximum in the last slot. Apply the same word of fixed-slot outward moves to both. The number of words with equal terminal maxima at most R is $o((\log R)^\beta)$.*

Proof. Let Γ be the group of polynomial automorphisms generated by the Vieta involutions and coordinate permutations. Put

$$J_9 = \ker\big(\Gamma \longrightarrow \mathrm{Sym}((\mathbb{Z}/9\mathbb{Z})^n)\big).$$

On an ambient ball $z + 3\mathbb{Z}_3^n$, with $z \in \{0,1,2\}^n$, the map $g \in J_9$ becomes $(g(z+3u) - z)/3$. This is congruent to the identity coefficientwise modulo 3. The constant term is divisible by 3, evaluation at $u = e_i$ determines the linear congruences, and every term of degree at least two has a factor of 3. Poonen's interpolation theorem [11] gives a restricted analytic parameterization of its iterates for times in $\mathbb{Z}_3$.

On the closure of a diagonal J_9 orbit of a pair, let μ_{orb} be the inverse limit of the uniform probabilities on its finite residue orbits. This measure has full support and is ergodic: the projection of any invariant function onto the functions on a finite quotient is constant by transitivity, and the finite quotient projections converge in L^2. The interpolated maps preserve the measure and satisfy the addition law for their time parameters, by continuity from integer iterates. These facts hold on each invariant residue ball and hence on the orbit closure. Let E be the set on the closure where last coordinates agree. On each interpolated g orbit its indicator is almost everywhere constant in the time variable, since a nonzero restricted analytic function has only finitely many zeros in $\mathbb{Z}_3$. Write $\mathbf{1}_E$ for its indicator and dt for normalized Haar measure on $\mathbb{Z}_3$. Then

$$\int_{\mathbb{Z}_3}\int |\mathbf{1}_E(g^t x) - \mathbf{1}_E(g^{t+1}x)|\, d\mu_{\mathrm{orb}}(x)\, dt = 0.$$

By invariance the inner integral equals $\int |\mathbf{1}_E(x) - \mathbf{1}_E(gx)|\, d\mu_{\mathrm{orb}}(x)$ for every t. Applying this to each $g \in J_9$ proves that E has measure zero or one.

To rule out measure one, write m_i for the fixed-slot outward move on slots i, n. It acts on these two coordinates by the matrix

$$U(C_i) = \begin{pmatrix} 0 & 1 \\ -1 & C_i \end{pmatrix}, \qquad C_i = n \prod_{j \neq i,n} x_j.$$

The value C_i is unchanged by iteration. For a direct check modulo 9, set $F_0 = 0$, $F_1 = 1$ and $F_{m+1} = CF_m - F_{m-1}$. Cayley–Hamilton gives $U(C)^m = F_m U(C) - F_{m-1} I$ for $m \geq 1$. Iterating the scalar recurrence gives $(F_{35}, F_{36}) = (8, 0)$ modulo 9 for each $C = 0, \ldots, 8$. Thus $m_i^{36} \in J_9$. If the two pairs have matching last coordinates at times $0, 36, 72$ under m_i, the Cayley–Hamilton identity and positivity of the middle value imply equality of the two traces of $U(C_i)^{36}$. This trace equals $r^{36} + r^{-36}$ where $r = (C_i + \sqrt{C_i^2 - 4})/2$. Since $C_i > 2$, equality forces equality of C_i. The nonzero lower-left entry of $U(C_i)^{36}$ then forces equality of coordinate i. Doing this for all i would make the seeds equal. This argument holds at any translated pair because the maps are invertible and preserve positivity of integral solutions. Since E is closed, full support excludes measure one. The Γ orbit closure is the union of finitely many J_9 orbit closures. For each component, the proportion of its reduction modulo 3^k on which the last coordinates agree tends to zero by continuity from above of the compatible measure. The fraction for the full orbit is bounded by the sum of these finitely many component fractions, since the size of each component reduction is at most the size of the full orbit reduction.

Fix the modulus 3^k. Order the parent slots of both seeds by increasing entries of q_w, breaking ties by their slot indices. During a common fixed-slot move, each tuple replaces the selected parent by its old maximum and places the new maximum last. The corresponding row state replaces the selected entry by the sum of the row entries. This sum is strictly larger than every old row entry, so the order of the row state is maintained by appending the replaced parent at the end. Keep the parents of both tuples in this same row-state order, whether or not their entries increase in that order. With this convention the slot-j update of the row state is γ_j, and the simultaneous slot-j update of the tuples is λ_j, which moves slot j last and then applies the Vieta involution there. The maxima remain in the last slot of each tuple. These updates give a bijection between the formal fixed-slot suffixes and the formal ordered suffixes, by tracking the parent-slot permutation at every step. The identities

$$\lambda_j^{-1} \lambda_{j+1} = (j, j+1) \quad (j < d), \qquad \lambda_d^{-1} (i, d) \lambda_d = (i, n) \quad (i < d)$$

give all coordinate permutations. Since λ_d is the swap (d, n) followed by the last Vieta involution, these moves also give that involution and its conjugates. They therefore generate Γ. In the finite paired residue orbit $\mathcal{C}$, the accelerated moves $\lambda_d^A \lambda_j$ for $j < d$ generate a transitive permutation group, since $A = 0, 1$ suffice to recover the ordinary generators. A semigroup of permutations of a finite set contains their inverses.

Let $\phi_{A,j} = [\gamma_d^A \gamma_j]$ on the sorted unit simplex and $\tau_{A,j}(x) = \log \|(\gamma_d^A \gamma_j) x\|_1$. Write $\pi_{A,j}$ for the corresponding permutation of $\mathcal{C}$. The transfer operator is

$$L_s^{\mathcal{C}} f(x, c) = \sum_{A \geq 0,\, j < d} e^{-s\tau_{A,j}(x)} f(\phi_{A,j} x, \pi_{A,j} c).$$

Its leading real eigenfunction is the scalar h_s independently of color, and its dual eigenmeasure is the scalar eigenmeasure times the uniform probability on $\mathcal{C}$. Write L_u for the scalar transfer operator at $u = \operatorname{Re} s$. For a vector of functions f, set $F_0(x) = \max_c |f(x, c)|$ and $F_1(x) = \max_c \|Df(x, c)\|$. For the jth iterate, each branch contracts by a factor at most $(2/3)^j$. Its total roof is the sum of the individual roofs evaluated after the successive branch compositions.

The individual roof derivatives are uniformly bounded, so the total roof derivatives are bounded by a convergent geometric sum. The color permutations have zero derivative in x. Termwise differentiation and the chain rule therefore give

$$\max_c |(L_s^{\mathcal{C}})^j f(x,c)| \le L_u^j F_0(x),$$
$$\max_c \|D(L_s^{\mathcal{C}})^j f(x,c)\| \le C_s L_u^j F_0(x) + C(2/3)^j L_u^j F_1(x).$$

The branch sum and its derivatives in s converge on compact subsets of $\operatorname{Re} s > 1$ by the bound $O((A+2)^{-\operatorname{Re} s})$ and its logarithmic derivatives. The compact inclusion of C^1 into C^0 for the finite tuple of functions gives, by the two-norm argument of [8, Proposition 45 and Lemma 46], an essential spectral radius strictly below λ_u, the leading eigenvalue of L_u.

The branch orbit from every (x,c) is dense in the scalar attractor times $\mathcal{C}$: use an outer block placing the simplex in a desired cylinder, and use an inner block to adjust the color to its preimage under that outer block. For an eigenfunction with eigenvalue 1 at $s = \beta + it$, normalize by h_β. The maximum-modulus inequality, density of these orbits and uniform contraction imply that its modulus is constant and nonzero on the attractor times $\mathcal{C}$. Indeed iteration places all projective orbits uniformly near the attractor, and on the attractor equality propagates the maximum along every branch. Equality in the triangle inequality then gives a phase relation for each branch. For $a_A = \gamma_d^A \gamma_{d-1}$ the subspace where the first $d-2$ entries vanish is invariant. On the last two coordinates the matrix is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & A+1 \end{smallmatrix}\right)$. Its positive Perron eigenvector is sorted and projects to the branch fixed point. The roof at this point is

$$\log\left(\frac{A+1+\sqrt{(A+1)^2+4}}{2}\right).$$

Along any progression in A with step equal to the order of the color permutation induced by λ_d, the branch color permutation is constant. Raising the branch to the order m of that permutation shows that mt times every such roof belongs to $2\pi\mathbb{Z}$. Successive roof differences are positive and tend to zero, so $t = 0$. At $t = 0$ the same branch relation and density show that the eigenfunction is constant after normalization on the attractor. Uniform contraction and iteration then show constancy everywhere. The normalized operator contracts the sup norm, which excludes a Jordan block at 1. For the scalar leading eigenvalue, differentiation of the eigenvalue equation and application of the dual eigenmeasure give

$$\lambda'_\beta = -\int_\Delta \sum_{A\ge 0,\, j<d} \tau_{A,j}(x) e^{-\beta\tau_{A,j}(x)} h([\gamma_d^A \gamma_j]x)\, d\nu(x) < 0.$$

The derivative for the product eigenvalue is the same because the color-constant subspace carries the scalar operator. Thus the resolvent has a simple pole at β with the product eigenprojection and is holomorphic near every other point of the line $\operatorname{Re} s = \beta$.

For a positive sorted seed q and a starting color c, the resolvent applied to the indicator of terminal color set E is the Laplace transform of the counting measure of the logarithms of the height ratios of accelerated words. Its residue at β is $|E|/|\mathcal{C}|$ times the scalar residue because the eigenprojection is the scalar eigenprojection tensored with the uniform probability. Apply the renewal argument of [8, Section 4.3] to this positive counting measure. The pole and the absence of any other singularity on the line $\operatorname{Re} s = \beta$ give, for each fixed q, c,

$$\lim_{L\to\infty} L^{-\beta} \#\{B : \|Bq\|_1 \le L,\ \pi_B c \in E\} = \frac{|E|}{|\mathcal{C}|} c_{\mathrm{acc}}(q),$$

where B ranges over accelerated words including the identity, π_B is the corresponding color permutation, and $c_{\mathrm{acc}}(q)$ is the scalar coefficient. In particular the coefficient is bounded by $C\|q\|_1^{-\beta}$.

Every ordinary word consists, in order of application, of A repetitions of the last-letter move followed by an accelerated word, for a unique $A \geq 0$. If the initial sorted row seed is q, its norm after that run is $r + As$, with $r = \|q\|_1$ and $s = r - q_d > 0$. The color after the run remains in the same finite orbit. The scalar upper estimate bounds each normalized colored summand by $C(r + As)^{-\beta}$, independently of the modulus. Summation and dominated convergence therefore give the limit for ordinary words as $|E|/|\mathcal{C}|$ times their scalar coefficient $c(q)$ from (2).

For the two seeds, take q to be the sorted version of q_w and take E to be the set of colors where the final coordinates agree. For a suffix t, its final row-state norm equals $H(v_{wt})$ by the continuation identity, since sorting preserves the sum of the entries. By (1), every suffix counted in the lemma has row-state norm at most $L_R = 1 + (n-2)\log R/\log(n-1)$ and terminal color in E. For this fixed modulus, its count divided by $(\log R)^\beta$ thus has upper limit at most

$$c(q)\left(\frac{n-2}{\log(n-1)}\right)^\beta \frac{|E|}{|\mathcal{C}|}.$$

The factor preceding the residue fraction depends on w and n only. Sending $k \to \infty$ makes the residue fraction vanish and proves the lemma. □

9. Proof of Theorem 1

Proof of Theorem 1. Fix an invertible rational matrix G, and count pairs of addresses v_w, v_u of height at most L with $v_u = Gv_w$, equal maxima, and distinct sorted tuples. If this count is not $o(L^\beta)$, along a sequence of L the measures assigning mass $L^{-\beta}$ to each pair $([v_w], [v_u])$ have a nonzero weak limit. Each address occurs at most once in each marginal because G is invertible. By (3) the limiting marginals are dominated by $c(\mathbf{1})\kappa$. The limit is supported on the projective graph of G with positive scale: for simplex points x, y on that graph, $Gx = ay$ with $a = \|Gx\|_1$ bounded below by a positive constant. Thus for the first marginal ρ, both ρ and $[G]_*\rho$ are absolutely continuous with respect to ν_{sym} by (4).

Lemma 3 covers almost every such source point by open prefix simplices on each of which

$$G = cR_{u_0}\Pi R_{w_0}^{-1},$$

with Π a coordinate permutation and $c \neq 0$. Any actual finite address in the interior of $[R_{w_0}]\Delta_0$ has prefix w_0: at each inverse step the strict dominant-coordinate inequality forces the next letter, and $[\mathbf{1}]$ cannot lie in a first-level simplex. If the patch contains an actual collision pair, primitivity of both integral vectors forces $|c| = 1$, and positivity of the two vectors forces $c = 1$. For the remaining tail t the two addresses are therefore $v_{w_0 t}$ and $v_{u_0\Pi(t)}$, where $\Pi(t)$ acts on the letters of t. Their maxima are obtained by applying the common tail t to the framed seeds a_{w_0} and $\Pi^{-1}a_{u_0}$, with the permutation acting on the parent slots. If the seeds coincide, word recovery gives $u_0 = \Pi(w_0)$, so $G = \Pi$ and the sorted tuples are equal. If the seeds are distinct, Lemma 7 and (1) bound the patch count by $o(L^\beta)$, since the maxima of its addresses are at most $\exp(L\log n/(n-2))$. The open-set inequality for weak convergence gives zero ρ-mass to each patch. The countable cover contradicts the choice of ρ. This proves the required $o(L^\beta)$ estimate for every fixed G.

For each maximum occurring at least twice among the sorted tuples of largest entry at most R, form a directed cycle through its distinct tuples. It has one pair per tuple and both projections are injective after choosing one framed address per sorted tuple. By (1) both address heights are at most

$$L_R = 1 + \frac{n-2}{\log(n-1)}\log R.$$

Given $\varepsilon > 0$, Proposition 6, conjugated by V from column addresses $z_w = P_w\mathbf{1}$ to normal addresses v_w, puts all but $\varepsilon L_R^\beta + o_\varepsilon(L_R^\beta)$ of these pairs on a fixed finite set of rational graphs. Each such

graph contributes $o(L_R^\beta)$ by the preceding argument. Letting $R \to \infty$ and then $\varepsilon \downarrow 0$ shows $B_n(R) = o((\log R)^\beta)$.

The scalar word count in (2) gives the required lower bound for $T_n(R)$. For $H(v_w) \le (n-2)\log R/\log n$, equation (1) gives $M(w) < R$. A sorted tuple has at most $d!$ framed representatives, since its largest entry is in the last slot and its other entries can be permuted in at most $d!$ ways. Each framed representative determines its word. Thus

$$T_n(R) \ge \frac{1}{d!} N_{\mathbf{1}}\left(\frac{n-2}{\log n}\log R\right) \gg (\log R)^\beta.$$

Together with the bound for $B_n(R)$, this proves the theorem. □

References

[1] A. Baragar, On the unicity conjecture for Markoff numbers, *Canad. Math. Bull.* **39** (1996), 3–9.
[2] A. Baragar, The exponent for the Markoff–Hurwitz equations, *Pacific J. Math.* **182** (1998), 1–21.
[3] Y. Benoist and J.-F. Quint, *Random walks on reductive groups*, Springer, Cham, 2016.
[4] C. Bandt and T. Retta, Topological spaces admitting a unique fractal structure, *Fund. Math.* **141** (1992), 257–268.
[5] W. Y. Chen, Almost all Markoff numbers are unique, Hexagon preprint 2610.00281v2, 2026.
[6] F. Comets, R. Fernández and P. A. Ferrari, Processes with long memory: regenerative construction and perfect simulation, *Ann. Appl. Probab.* **12** (2002), 921–943.
[7] G. Frobenius, Über die Markoffschen Zahlen, *Sitzungsber. Königl. Preuss. Akad. Wiss. Berlin* (1913), 458–487.
[8] A. Gamburd, M. Magee and R. Ronan, An asymptotic formula for integer points on Markoff–Hurwitz varieties, *Ann. of Math.* (2) **190** (2019), 751–809.
[9] A. Hurwitz, Über eine Aufgabe der unbestimmten Analysis, *Arch. Math. Phys.* (3) **11** (1907), 185–196.
[10] A. Markoff, Sur les formes quadratiques binaires indéfinies, *Math. Ann.* **15** (1879), 381–406, and **17** (1880), 379–399.
[11] B. Poonen, p-adic interpolation of iterates, *Bull. Lond. Math. Soc.* **46** (2014), 525–527.
[12] D. Zagier, On the number of Markoff numbers below a given bound, *Math. Comp.* **39** (1982), 709–723.

Zhili College, Tsinghua University, Beijing 100084, China
Email address: tangqy24@mails.tsinghua.edu.cn