

Hankel determinants and the length of finite Pierce expansions

Jason Davies*

October 27, 2026

Abstract

For integers $a > b > 0$, let $P(a, b)$ be the number of steps the iteration $x \mapsto a \bmod x$ takes to reach 0 from b ; it is the length of the Pierce expansion of b/a . We prove

$$P(a, b) \ll_{\varepsilon} a^{\beta_* + \varepsilon}, \quad \beta_* = \frac{41301 - 3\sqrt{82849}}{137480} = 0.2941336553 \dots,$$

improving the exponent $1/3 - 2/177$ of Chase and Pandey. The tail of an orbit carries a sequence of positive integers, the Hankel determinants of a natural moment sequence, and these are small where the drops of the orbit are small. Grouping orbit states by such integers places the quotient differences of each group on curves of bounded degree, whose integer points we count with the uniform bounds of Castryck, Cluckers, Dittmann and Nguyen. The resulting exponent approaches $5/17$ when these curves have large degree relative to the logarithmic size of the box. We raise the degree by adjoining further integer invariants, by an effective Hilbert irreducibility sieve, and by an exact identity for the direction of the curvature vector, and come within $1.61 \cdot 10^{-5}$ of $5/17$. We also give an elementary proof of the exponent $3/10 + \varepsilon$ and a bound for orbits whose quotient increments rarely rise. All results are formalised in Lean 4 with Mathlib. Shallit’s conjectured $O((\log a)^2)$ bound remains open.

AI transparency statement. This paper was produced entirely by artificial intelligence, including the discovery of the main results, their formalisation in Lean, and the writing of the manuscript. The human author initiated the research and directed the AI systems, but did not write the text. Details are given at the end of the paper.

1 Introduction

Let $a > b > 0$ be integers; we call a the *dividend*. Put $b_0 = b$ and $b_{i+1} = a \bmod b_i$ while $b_i > 0$. The sequence decreases until it reaches 0; if $b_n = 0$ we write $P(a, b) = n$. For example, $P(35, 22) = 7$, from 22, 13, 9, 8, 3, 2, 1, 0. This is the length of the Pierce expansion of b/a [13, 14], since

$$\frac{b}{a} = \frac{1}{q_0} - \frac{1}{q_0 q_1} + \frac{1}{q_0 q_1 q_2} - \dots, \quad q_i = \lfloor a/b_i \rfloor.$$

Shallit [12] asked whether

$$P(a, b) = O((\log a)^2). \tag{1}$$

Shallit [14] proved $P(a, b) \ll a^{1/2}$ by “Archimedean” arguments. Erdős and Shallit [6] combined these with a divisor argument to obtain $P(a, b) \ll_{\varepsilon} a^{1/3+\varepsilon}$, and showed that $\max_b P(a, b)/\log a$ does

*jason@jasondavies.com

not tend to 0. Chase and Pandey [5] improved the upper exponent to $1/3 - 2/177$, using exponential sum estimates, and proved $\max_b P(a, b) \gg \log a / \log \log a$ for all large a . Baraskar and Vukusic [1] study the related sets of iterated remainders at fixed iteration depth. Shallit's talk pages [15] report a bound $n^{3/10}$, described as “LLM, unpublished so far”. We have not seen that argument; our elementary bound below has the same exponent up to ε , and we make no claim about how the arguments are related.

Our main result is the following uniform bound.

Theorem 1.1. *For every $\varepsilon > 0$ there is $C_\varepsilon > 0$, independent of a, b , such that*

$$P(a, b) \leq C_\varepsilon a^{\beta_* + \varepsilon} \quad (0 < b < a), \quad \beta_* = \frac{41301 - 3\sqrt{82849}}{137480}.$$

Here $\beta_* = 0.2941336553\dots$. The proof depends on two fixed positive integers m, n , the powers of the adjoined coordinates (see “Raising the degree” below), and β_* is the limit of the resulting exponents as n/m tends to an irrational ratio. For instance $(m, n) = (3, 2)$ gives the rational exponent $4317/14677 = 0.2941336785\dots$ (Theorem 10.9).

The curve-counting bounds of this paper (Theorems 1.1, 1.2 and 1.3) all lie just above $5/17 = 0.2941176\dots$, and Theorem 1.1 is within $1.61 \cdot 10^{-5}$ of it. We first introduce the integer invariants of an orbit, and then explain where $5/17$ comes from, since it organises the whole argument.

Integer invariants of an orbit

Write $k_i = q_{i+1} - q_i$ for the quotient increments and $d_i = b_i - b_{i+1}$ for the drops. The *curvature* $D_i = d_i k_{i+1} - d_{i+1} k_i$ of four consecutive terms satisfies

$$D_i b_{i+2} = d_i d_{i+2} + k_i d_i d_{i+1} + (k_i - 1) d_{i+1}^2. \quad (2)$$

Thus D_i is a positive integer, although its defining terms largely cancel, and it is small when the drops are small. A six-term identity supplies a second positive integer L , and with it the bound $q_i \leq 12K^3 \max(D_{i+1}, D_{i+2})$ when the next four increments are at most K (Section 4).

These integers belong to one hierarchy. Each tail carries a finite positive measure ν_i of mass b_i/a , and

$$I_r(i) = a \det \left(\int x^{j+l} d\nu_i \right)_{j,l < r}$$

is a positive integer when the tail has at least r positive terms. The first four invariants are b_i, d_i, D_i, L_i , in the index ranges of Theorem 6.2, which stay away from the end of the orbit.

Where the exponent $5/17$ comes from

We count the T orbit terms in a remainder window $[A, 2A)$, then sum over dyadic A . The total drop across the window is less than A , and Erdős and Shallit's divisor argument shows that a drop value can repeat only $a^{o(1)}$ times; hence $A \geq T^{2-o(1)}$. Similarly the quotients increase by at most a/A across the window. Call an orbit index i a *start* when we regard it as the first index of the block $b_i, b_{i+1}, \dots$. A fixed proportion of starts in the window then have their next few increments bounded by $H \asymp a/(AT) \leq a^{1+o(1)}/T^3$. For such a start, form the 3×5 matrix whose rows are the next five drops, the increments, and a quadratic expression in the increments. Its ten maximal minors, among them $I_4(i)$, $I_4(i+1)$ and $I_4(i+2)$, are integers. To leading order the drop row is a combination of the other two rows, since the Pierce expansion gives $d_j/a = k_j/q_i^2 - (r_j + k_j)/q_i^3 + O(H^3/q_i^4)$

with r_j the quadratic row (Proposition 7.8). Subtracting that combination leaves entries of size $O(aH^3/q_i^4)$; so the minors have absolute value at most

$$X \asymp \frac{A^4 H^6}{a^3} \asymp \frac{a^3}{A^2 T^6} \leq \frac{a^{3+o(1)}}{T^{10}}.$$

Their vector is the *pattern* of the start. It is an integral point of the affine cone over the Grassmannian $\text{Gr}(3, 5)$, and there are $X^{5+o(1)}$ such points (Section 7); the exponent 5 is the Fano index of $\text{Gr}(3, 5)$, so this count of all integral points of the cone cannot be improved (orbit patterns could in principle be sparser). If every pattern were shared by $O(1)$ starts, then $T \ll X^5$ would give $T^{51} \leq a^{15+o(1)}$, that is, $T \leq a^{5/17+o(1)}$.

The rest of the argument bounds the number of starts sharing a pattern. For a fixed pattern, three nodes $s_j = q_{i+j} - q_i$, $j = 1, 2, 3$, determine the increments (the pattern supplies two linear relations, which give k_3 and k_4). The pattern also supplies a quadratic equation in the nodes. A second equation, of degree at most twelve, says that $I_4(i)$ times the resultant of the orthogonal quartic of state i and x times that of state $i+1$ equals a times the fourth power of the squared norm of the first quartic; it involves the dividend a . The nodes therefore lie on a union of curves of total degree at most 24. The coefficient-uniform bound of Castryck, Cluckers, Dittmann and Nguyen [3] says that an integral curve over $\mathbb{Q}$ of degree δ has at most $B^{1/\delta+o(1)}$ integer points in a box of size B , whatever its coefficients. So a pattern carries $B^{1/\delta+o(1)}$ starts. If the degree could be made large with the box fixed, the factor $B^{1/\delta}$ would become $a^{o(1)}$ and the exponent would tend to $5/17$. A component has degree at most 24, and the components of smallest degree govern the bound; but the curves carry further integer-valued functions. Adjoining powers of them as coordinates raises δ and $\log B$ together, and what matters is the ratio $\log B/\delta$ (Remark 8.16). Every refinement below either gains more degree per unit of height or shrinks B , for almost all patterns.

Raising the degree

We adjoin three integer curvature minors $W = (W_0, W_1, W_2)$ and the fifth Hankel invariant $J = I_5(i)$, whose sizes satisfy

$$|W_j| \ll S := A^3 H^3 / a^2, \quad |J| \ll Y := A^5 H^{10} / a^4.$$

On each component they are rational functions of the nodes: J is a polynomial, and W is obtained by recovering the moments of ν_i from the two orthogonal polynomials and the division equations, which gives the curvature $I_3(i)$ as a rational function of the nodes (its *reconstruction*). On a component of degree d , J has a pole of total order $5d/2$ at infinity, while the W_j vanish there to total order $3d/2$. A suitable combination of the W_j also vanishes at the node triple of one orbit state, so it has at least $3d/2 + 1$ zeros, hence as many poles, all affine. For fixed positive integers m, n , consider the graph obtained by adjoining W_0^m, W_1^m, W_2^m, J^n to the nodes. The degree of an affine curve is the pole degree of a generic linear combination of its coordinates, and on this graph J^n has poles only at infinity and the W_j^m only at affine points, so their pole degrees add. The graph therefore has degree at least $\frac{5n}{2}d + m(\frac{3d}{2} + 1)$, and the box grows only to $\max(H, S^m, Y^n)$. Conics and components on which the reconstruction fails are counted separately. This gives Theorem 1.2 below.

The pattern also determines a rank-three lattice (the rows satisfy congruences modulo the largest minor). If λ is the smallest sup norm of a nonzero lattice vector, a reduced basis gives integer increment and curvature coordinates of sizes $O(H/\lambda)$ and $O(S/\lambda^2)$. Averaged bounds for short lattice vectors shrink the box for most patterns without changing the graph degree (Section 9, Theorem 1.3).

The polar improvement

Two further facts give Theorem 1.1. First, an effective Hilbert irreducibility sieve (the *Hilbert sieve*) shows that, outside a sparse set of patterns, the pattern curve is irreducible over $\mathbb{Q}$, of the maximal degree 24. Second, an exact identity, $JW = N_0^2(G\mathbf{k} + \ell)$, expresses the curvature vector through three affine-linear forms in the increments $\mathbf{k} = (k_0, k_1, k_2)$. Here N_0 is a pattern coordinate, and $G\mathbf{k} + \ell$ arises by polarising the pattern quadratic $f(\mathbf{k}) = \mathbf{k} \cdot (G\mathbf{k} + \ell)$. Since $f(\mathbf{k}) = 0$, the increment vector $\mathbf{k}$ is orthogonal to $G\mathbf{k} + \ell$; the division equations say exactly that the drop vector $\mathbf{d}$ is orthogonal to it too, so that $W = \mathbf{d} \times \mathbf{k}$ is parallel to it.

On a component of degree 24, J has 60 zeros, all affine. Since $W = N_0^2(G\mathbf{k} + \ell)/J$, each zero of J is a pole of W unless the three linear forms also vanish there, and outside a fixed proper algebraic set of *degenerate* patterns this happens at most four times, counted with multiplicity. So W has at least $60 - 4 = 56$ affine poles, and the powered graph has degree

$$\deg \Gamma \geq \frac{5n}{2} \cdot 24 + m(60 - 4) = 60n + 56m.$$

Without the sieve the general bound can use only $d \geq 4$, which gives $10n + 7m$; forcing $d = 24$ would give $60n + 37m$, and the polar identity raises the affine pole count from 37 to 56.

The sieve and the degenerate set together discard $O_\eta(a^\eta X^{14/3})$ patterns. These are counted with the unpowered graph of Section 8, of degree at least 17, which gives the exponents 717/2440 and 732/2491, both below 5/17. Applying the curve bound in the reduced lattice coordinates and optimising the ratio n/m gives β_* (Section 10).

Intermediate bounds

The two curve-counting stages before the polar improvement give useful bounds without the Hilbert sieve. Put

$$\alpha_* = \frac{1310 - \sqrt{655}}{4365} = 0.2942513\dots, \quad \gamma_* = \frac{471}{1559} - \frac{6\sqrt{105}}{7795} = 0.2942294\dots$$

Theorem 1.2 (Joint-invariant bound). *For every $\varepsilon > 0$, $P(a, b) \ll_\varepsilon a^{\alpha_* + \varepsilon}$ uniformly for $0 < b < a$.*

Theorem 1.3 (Exterior lattice bound). *For every $\varepsilon > 0$, $P(a, b) \ll_\varepsilon a^{\gamma_* + \varepsilon}$ uniformly for $0 < b < a$.*

Fixed powers give the rational instances 128/435 and 2201/7480 of the joint-invariant bound, and 153/520 of the exterior lattice bound; see Corollaries 8.15 and 9.11. The table records how the exponent approached 5/17 during this work. The first four rows are superseded stages, recorded for orientation only; their proofs are formalised in the repository but omitted here.

Exponent	Exponent – 5/17	Main input
23/77	$4.6 \cdot 10^{-3}$	I_4 as an integer label
11/37	$3.2 \cdot 10^{-3}$	a ternary quadratic and norm equations on slices
8/27	$2.2 \cdot 10^{-3}$	the $X^{5+o(1)}$ count of minor vectors
45/152	$1.9 \cdot 10^{-3}$	lattices of pivot increments and their short vectors
α_*	$1.3 \cdot 10^{-4}$	adjoining curvature and rank-five invariants
γ_*	$1.1 \cdot 10^{-4}$	reducing the joint coordinates in a lattice basis
β_*	$1.6 \cdot 10^{-5}$	the Hilbert sieve and the polar identity

An elementary bound and a rise-sensitive bound

The initial identities already yield an elementary improvement over the exponent $1/3$.

Theorem 1.4. *For every $\varepsilon > 0$, $P(a, b) \ll_{\varepsilon} a^{3/10+\varepsilon}$ uniformly for $0 < b < a$.*

The proof uses only division equations and the divisor estimate $\tau(n) \ll_{\eta} n^{\eta}$. For a window of T remainders in $[A, 2A)$,

$$T(T-1) \leq 4A\Delta, \quad (3)$$

$$T^6 A^2 \ll a^3, \quad (4)$$

where $\Delta \ll_{\eta} a^{\eta}$ is a maximum of the divisor function. The first is the Erdős–Shallit drop argument. The second follows from the six-term identity: some block of six terms has total drop at most $10A/T$, which bounds its increments and curvatures, and these bound the quotient $q \asymp a/A$, with no divisor loss. Multiplying the second bound by the square of the first eliminates A and gives $T^{10} \ll a^3 \Delta^2$ (Section 5). The six-term bound is also used in all the later arguments.

There is also a bound sensitive to the order of the increments. Let

$$R = 1 + \#\{i : 0 \leq i \leq P(a, b) - 3, k_i < k_{i+1}\}.$$

Theorem 1.5. *$P(a, b) \ll a^{1/4} \sqrt{R}$ uniformly for $0 < b < a$. In particular, nonincreasing quotient increments give $P(a, b) \ll a^{1/4}$.*

Curvature positivity makes the drops decrease between rises. The six-term identity limits repeated equal increments, and summing the resulting window bounds proves the theorem. This argument is independent of the curve-counting refinements.

Reading guide and formal verification

Sections 3–5 establish the elementary identities and prove Theorem 1.4. Sections 6–7 construct the integer invariants and their patterns. Section 8 builds the enlarged curves and proves Theorem 1.2, and Section 9 reduces their lattice coordinates and proves Theorem 1.3. Section 10 proves Theorem 1.1 by rerunning the window bound of Section 9 with a larger graph degree, so it depends on all of Sections 6–9. The independent rise bound, Theorem 1.5, is proved in Section 11 and uses only Sections 3 and 4. The conic count, graph-degree upper bound, and three-dimensional lattice reduction are proved in Appendices A, B and C. Appendix D proves the effective Hilbert irreducibility estimate, Appendix E records the degree certificates used to exclude lines, and Appendix F gives the certificates for the polar bound.

All the unconditional results above, including their curve-counting and specialisation inputs, are formalised in Lean 4 [8] with Mathlib [9]. Section 12 identifies the proof artefact and its verification scope. Section 13 explains the present 5/17 barrier.

2 Preliminaries

2.1 Notation

Throughout, $a > b > 0$ are integers, $b_0 = b$, $b_{i+1} = a \bmod b_i$, and $P = P(a, b)$, so that $a > b_0 > b_1 > \dots > b_{P-1} > 0 = b_P$. Write $q_i = \lfloor a/b_i \rfloor$, so

$$a = q_i b_i + b_{i+1}, \quad 0 \leq b_{i+1} < b_i \quad (0 \leq i < P).$$

An index i of the orbit is also called a *state*; quantities “at state i ” are computed from $b_i, b_{i+1}, \dots$ and $q_i, q_{i+1}, \dots$, and are defined whenever the orbit terms they use exist. When we count blocks $b_i, b_{i+1}, \dots$ in a window, we call their first index i a *start*; a start is thus just a state, named for its role in the count. We write $k_i = q_{i+1} - q_i$ for the quotient *increments* and $d_i = b_i - b_{i+1}$ for the *drops*. Let $\tau(n)$ be the number of positive divisors of n , and $\tau_{\max}(Y) = \max_{1 \leq j \leq Y} \tau(j)$. We write $F \ll_\eta G$ when $F \leq C_\eta G$, with the constant uniform in the dividend and orbit. A factor denoted $a^{o(1)}$ is bounded by $C_\eta a^\eta$ for every fixed $\eta > 0$; such factors include the divisor losses below. We write $F \asymp G$ when both $F \ll G$ and $G \ll F$ hold. Absolute constants play no role in the exponents, and we do not track them: implied constants are absolute unless their dependence is indicated, and a letter c or C with a descriptive subscript denotes a suitable absolute constant introduced where it is needed. Explicit values of all constants are recorded in the Lean development; a few of them are sharper than the ones derived here.

Lemma 2.1 (Quotients increase). *If $b_{i+1} > 0$ then $q_{i+1} \geq q_i + 1$.*

Proof. Since $b_{i+1} < b_i$ we have $(q_i + 1)b_{i+1} \leq q_i b_i + b_{i+1} = a$, so $q_{i+1} = \lfloor a/b_{i+1} \rfloor \geq q_i + 1$. $\square$

Lemma 2.2 (Divisor bound). *For every integer $k \geq 1$ there is C_k with $\tau(n)^k \leq C_k n$ for all $n \geq 1$.*

Proof. This is standard: $\tau(n)^k/n = \prod_{p^\alpha \parallel n} (\alpha + 1)^k/p^\alpha$. Each factor with $p \geq 2^k$ is at most $(\alpha + 1)^k/2^{k\alpha} \leq 1$. Each of the fewer than 2^k factors with $p < 2^k$ is at most $\sup_\alpha (\alpha + 1)^k/2^\alpha < \infty$. $\square$

2.2 Window runs

Definition 2.3. Let $A \geq 1$ be an integer. A *window run* of length T is a sequence $x_0, x_1, \dots, x_{T-1}$ of consecutive terms of the sequence (b_i) , so $x_{j+1} = a \bmod x_j$, all lying in $[A, 2A)$.

In a window run, $x_0 > x_1 > \dots > x_{T-1}$.

Lemma 2.4 (Quotients in a window). *If $0 < b_i < 2A$ then $a < 4Aq_i$.*

Proof. $a = q_i b_i + b_{i+1} < (q_i + 1)b_i < 2A(q_i + 1) \leq 4Aq_i$, since $q_i \geq 1$. $\square$

3 The curvature of three consecutive steps

Fix four consecutive terms

$$x = b_i, \quad y = b_{i+1} = a \bmod x, \quad z = b_{i+2} = a \bmod y, \quad w = b_{i+3} = a \bmod z,$$

with $x, y, z > 0$ (and $w \geq 0$). Write

$$q = \left\lfloor \frac{a}{x} \right\rfloor, \quad u = \left\lfloor \frac{a}{y} \right\rfloor - q, \quad v = \left\lfloor \frac{a}{z} \right\rfloor - \left\lfloor \frac{a}{y} \right\rfloor, \quad d = x - y, \quad e = y - z, \quad f = z - w.$$

In the notation of the introduction, $q = q_i$, $u = k_i$, $v = k_{i+1}$ and $d, e, f = d_i, d_{i+1}, d_{i+2}$; the short letters keep the identities below readable. By Lemma 2.1, $u, v \geq 1$, and clearly $d, e, f \geq 1$. The division equations read

$$xq + y = a, \quad y(q + u) + z = a, \quad z(q + u + v) + w = a. \quad (5)$$

Definition 3.1. The *curvature* of the block is $D = dv - eu$.

Lemma 3.2 (Curvature identity).

$$Dz = df + ude + (u - 1)e^2.$$

In particular $D \geq 1$.

Proof. Subtracting the first two equations of (5) gives $uy = qd + e$, hence $uz = uy - ue = qd + (1 - u)e$. Subtracting the last two gives $vz = (q + u)e + f$. Therefore

$$(dv - eu)z = d((q + u)e + f) - e(qd + (1 - u)e) = df + ude + (u - 1)e^2.$$

The right-hand side is at least $df \geq 1$, and $z > 0$, so $D > 0$. $\square$

Lemma 3.3 (Size of the parameters). *We have*

$$u yx \leq a(d + e), \quad v zy \leq a(e + f), \quad Dz \leq (u + 1)(d + e + f)^2.$$

Proof. From $uy = qd + e$ we get $uyx = (qx)d + ex \leq ad + ae$, since $qx \leq a$ and $x < a$. Similarly $vz = (q + u)e + f$ with $(q + u)y \leq a$ and $y \leq a$ gives the second bound. For the third, by Lemma 3.2, $Dz = df + ude + (u - 1)e^2 \leq df + ue(d + e) \leq (1 + u)(d + e + f)^2$. $\square$

4 Six consecutive terms

Fix six consecutive terms $x_0 > x_1 > \dots > x_5 \geq 0$ of the sequence (b_i) with $x_4 > 0$. Write $q = \lfloor a/x_0 \rfloor$ and let $k_0, \dots, k_3$ be the next four quotient increments, so that

$$\lfloor a/x_j \rfloor = q + k_0 + \dots + k_{j-1} \quad (0 \leq j \leq 4).$$

By Lemma 2.1, $k_0, \dots, k_3 \geq 1$. Let $f_j = x_j - x_{j+1} \geq 1$ be the drops, and let

$$D = f_0k_1 - f_1k_0, \quad E = f_1k_2 - f_2k_1, \quad F = f_2k_3 - f_3k_2$$

be the curvatures of the blocks starting at x_0, x_1 and x_2 . Write $\mathcal{E}_j$ for the division equation $x_jq_j + x_{j+1} - a = 0$, where $q_j = \lfloor a/x_j \rfloor$. Define polynomials

$$\begin{aligned} \sigma(y_1, y_2, y_3) &= y_1^2y_2 + (y_1 - 1)y_2^2 + y_1y_3, \\ \omega(y_1, y_2, y_3, y_4) &= (y_1^2 + (y_1 - 1)y_2)y_4 + y_3(y_1^3 + y_1y_2(2y_1 - 1) + (y_1 - 1)y_2^2) \\ &\quad + y_3^2((y_1 - 1)(y_1 + y_2) + 1), \end{aligned}$$

and put

$$L = (k_1k_2 + k_2^2 + k_3)D - (k_0^2 + (k_0 - 1)k_1)E - k_2(k_2f_0 - k_0f_2).$$

Lemma 4.1 (The six-term identity). 1. $\sigma(k_1, k_2, k_3)D - \sigma(k_0, k_1, k_2)E = k_1L$.

2. $qk_2L = \omega(k_0, k_1, k_2, k_3)E + \sigma(k_0, k_1, k_2)F$.

Proof. Part (1) is a polynomial identity in the drops and increments. Modulo k_1 we have $\sigma(k_1, k_2, k_3) \equiv -k_2^2$ and $\sigma(k_0, k_1, k_2) \equiv k_0k_2$, and $k_2D + k_0E = k_1(k_2f_0 - k_0f_2)$; so the left side is divisible by k_1 , and expanding the quotient gives L . Part (2) holds modulo the division equations: $qk_2L - \omega E - \sigma(k_0, k_1, k_2)F = k_2(c_0\mathcal{E}_0 - c_1\mathcal{E}_1 + c_2\mathcal{E}_2 - c_3\mathcal{E}_3)$, where

$$\begin{aligned} c_0 &= k_1^2k_2 + k_1k_2^2 + k_1k_3 - k_2^2, \\ c_1 &= k_0^2k_2 + 2k_0k_1k_2 + k_0k_2^2 + k_0k_3 + k_1^2k_2 + k_1k_2^2 - k_1k_2 + k_1k_3 - k_2^2, \\ c_2 &= k_0^2k_1 + k_0^2k_2 + k_0k_1^2 + 2k_0k_1k_2 + k_0k_2^2 + k_0k_2 + k_0k_3 - k_1^2 - k_1k_2, \\ c_3 &= k_0^2k_1 + k_0k_1^2 + k_0k_2 - k_1^2. \end{aligned}$$

This is an identity of polynomials in a, q , the x_j and the increments. $\square$

Corollary 4.2. *Suppose $k_0, \dots, k_3 \leq K$. Then $E, F \geq 1$, $L \geq 1$, and*

$$q \leq 9K^3E + 3K^3F.$$

Proof. Lemma 3.2, applied to the blocks starting at x_1 and x_2 , gives $E, F \geq 1$. Clearly $\omega > 0$ and $\sigma \geq 0$, so by Lemma 4.1(2) $qk_2L > 0$. Since $q \geq 1$ (as $x_0 < a$) and L is an integer, $L \geq 1$ and $qk_2 \leq \omega E + \sigma F$. Bounding each monomial, using $k_2 \geq 1$, gives $\omega \leq 9K^3k_2$ and $\sigma(k_0, k_1, k_2) \leq 3K^3k_2$. Dividing by k_2 gives the claim. $\square$

So the quotient q , which is about a/A when the six terms lie in a window $[A, 2A)$, is controlled by the small parameters of a single block of six terms. No counting over many blocks is needed.

Proposition 4.3 (Six-term window bound). *For a window run of length $T \geq 10$,*

$$T^6 A^2 \ll a^3.$$

Proof. For $0 \leq j \leq T-6$ let $s_j = x_j - x_{j+5}$. By telescoping, $\sum_{j \leq T-6} s_j \leq 5(2A-1) - 5A = 5A-5$. So some j has $s_j(T-5) \leq 5A-5$, and since $T \leq 2(T-5)$, we get $s := s_j$ with $sT \leq 10A$. Apply Corollary 4.2 to $x_j, \dots, x_{j+5}$, all of which lie in $[A, 2A)$, with $K = \max(k_0, \dots, k_3)$.

By Lemma 3.3 at x_j, x_{j+1} and x_{j+2} , each increment k satisfies $kA^2 \leq as$. Multiplying by T gives $KAT \leq 10a$. By Lemma 3.3 again, $EA \leq Ex_{j+3} \leq (k_1+1)s^2 \leq 2Ks^2$, and likewise $FA \leq 2Ks^2$. Hence $qA \leq 24K^4s^2$. Also $a < 4Aq$ by Lemma 2.4. Therefore

$$qA \cdot A^4 T^6 \leq 24(KAT)^4 (sT)^2 \ll a^4 A^2,$$

so $qA^3 T^6 \ll a^4$, and $aA^2 T^6 \leq 4qA^3 T^6 \ll a^4$. $\square$

5 Proof of Theorem 1.4

We bound the number T of orbit terms in a window $[A, 2A)$ in two ways. Write $B = a/A$, so that the quotients in the window are about B . The drop argument (Proposition 5.1) gives $T \ll a^{1/2+o(1)} B^{-1/2}$, which is good when the quotients are large, and the six-term bound (Proposition 4.3) gives $T \ll a^{1/6} B^{1/3}$, which is good when they are small. The two agree at $B = a^{2/5}$, where both equal $a^{3/10+o(1)}$. Multiplying suitable powers of the two bounds eliminates A (Corollary 5.2), so no case split is needed. Throughout,

$$\Delta = \Delta(a) = \tau_{\max}(2a) \ll_{\eta} a^{\eta}$$

by Lemma 2.2. In this section $x_0, \dots, x_{T-1}$ is a window run in $[A, 2A)$, and $q_j = \lfloor a/x_j \rfloor$ are its quotients.

Proposition 5.1 (Drop bound; after Erdős–Shallit). *If $T \geq 2$ then $T(T-1) \leq 4A\Delta$.*

Proof. Let $h_j = x_j - x_{j+1} \geq 1$ for $0 \leq j \leq T-2$. The total drop is $\sum_j h_j = x_0 - x_{T-1} \leq A-1$. Thus fewer than $(T-1)/2$ drops exceed $2A/(T-1)$. The remaining set $\mathcal{S}$ has $2|\mathcal{S}| \geq T$.

For $j \in \mathcal{S}$ we have $1 \leq h_j \leq \bar{h} := \lfloor 2A/(T-1) \rfloor$. Also $a+h_j = q_j x_j + x_{j+1} + x_j - x_{j+1} = (q_j+1)x_j$, so $x_j \mid a+h_j$. Moreover $a+h_j \leq 2a$. Since $j \mapsto x_j$ is injective, at most $\tau(a+h) \leq \Delta$ indices in $\mathcal{S}$ have $h_j = h$. Hence $|\mathcal{S}| \leq \bar{h}\Delta$, and

$$T(T-1) \leq 2|\mathcal{S}|(T-1) \leq 2\Delta \bar{h}(T-1) \leq 4A\Delta. \quad \square$$

Corollary 5.2 (Window bound). *If $T \geq 10$ then $T^{10} \ll a^3 \Delta^2$.*

Proof. The drop bound and $T \leq 2(T-1)$ give $T^4 \leq 64A^2\Delta^2$. Multiply by the six-term bound $T^6A^2 \ll a^3$ and cancel A^2 . $\square$

Proof of Theorem 1.4. Let $0 < b < a$. For $0 \leq k \leq \log_2 a$ let T_k be the number of indices $i < P$ with $2^k \leq b_i < 2^{k+1}$. Every b_i with $i < P$ lies in $[1, a)$, so $P = \sum_{0 \leq k \leq \log_2 a} T_k$. Since (b_i) is decreasing, the indices counted by T_k are consecutive. So when $T_k > 0$ they form a window run in $[A, 2A)$ with $A = 2^k$. Inserting $\Delta \ll_\eta a^\eta$ with $\eta = 5\varepsilon/2$ into Corollary 5.2 yields $T_k \ll_\varepsilon a^{3/10+\varepsilon/2}$; the windows with $T_k < 10$ satisfy the same estimate. There are $O(\log a) \ll_\varepsilon a^{\varepsilon/2}$ windows, so summing proves the theorem. $\square$

6 The Hankel hierarchy

From now on we write $N = P(a, b)$, freeing the letter P for polynomials, so that $b_0 > \dots > b_{N-1} > 0 = b_N$, and $k_i = q_{i+1} - q_i$. The division equation gives $b_i/a = (1 - b_{i+1}/a)/q_i$, and iterating it expresses b_i/a as the alternating Pierce series in the tail quotients $q_i, q_{i+1}, \dots$. The construction below rewrites this series as a Stieltjes transform,

$$\frac{b_i}{a} = \int \frac{d\mu_i(x)}{q_i + x},$$

of a probability measure μ_i on the nodes $q_{i+j} - q_i$ (Lemma 6.1(2)). The measure $\nu_i = \mu_i/(q_i + x)$ then has total mass b_i/a , and the Hankel determinants of its moments are the integer invariants we need.

Put $p_i(j) = q_{i+j} - q_i$ for $0 \leq j < N - i$; the symbols p_i, w_i are local to this construction. Define the weights backwards from $w_{N-1}(0) = 1$ by

$$w_i(0) = 1 - \sum_{j=0}^{N-i-2} \frac{w_{i+1}(j)}{p_i(j+1)}, \quad w_i(j+1) = \frac{w_{i+1}(j)}{p_i(j+1)} \quad (0 \leq j < N - i - 1).$$

Let μ_i be the measure with mass $w_i(j)$ at $p_i(j)$, and let

$$d\nu_i(x) = \frac{d\mu_i(x)}{q_i + x}, \quad h_i(s) = \int x^s d\nu_i = \sum_{j=0}^{N-i-1} \frac{w_i(j)p_i(j)^s}{q_{i+j}}.$$

The recursion is chosen to reproduce the division equation $q_i h_i(0) + h_{i+1}(0) = 1$ for $i+1 < N$.

Lemma 6.1 (The moment model). *1. The weights are positive and sum to 1.*

2. $h_i(0) = b_i/a$, that is, $b_i/a = \mathbb{E}[1/(q_i + X_i)]$ for $X_i \sim \mu_i$; and $h_{N-1}(s) = 0$ for $s \geq 1$.

3. $h_i(s+1) = \sum_{t \leq s} \binom{s}{t} k_i^{s-t} h_{i+1}(t)$ for $i+1 < N$.

4. $h_i(s) = (-q_i)^s b_i/a + z$ with $z \in \mathbb{Z}$.

Proof. (1) Since $p_i(j+1) \geq 1$, the sum defining $w_i(0)$ is at most 1. It is strictly less: if $N-i-1 \geq 2$, the term $j=1$ has $p_i(2) \geq 2$; if $i = N-2$, the last increment satisfies $q_{N-1} \geq q_{N-2} + 2$, because $b_{N-1} \mid a$ gives $q_{N-1} - q_{N-2} - 1 = q_{N-2}(b_{N-2} - b_{N-1})/b_{N-1} \geq 1$. (2) Since $p_i(j+1) = q_{i+j+1} - q_i$, the definitions give $q_i h_i(0) + h_{i+1}(0) = 1$, which is the normalised division equation; and $h_{N-1}(0) = 1/q_{N-1} = b_{N-1}/a$; since $\mu_{N-1} = \delta_0$, also $h_{N-1}(s) = 0$ for $s \geq 1$. (3) Use $p_i(j+1) = k_i + p_{i+1}(j)$ and the binomial theorem. (4) For $i = N-1$ this holds with $z = -(-q_{N-1})^s b_{N-1}/a = -(-1)^s q_{N-1}^{s-1}$ for $s \geq 1$, since $q_{N-1} b_{N-1} = a$. Then induct backwards on i using (3), $b_{i+1}/a = 1 - q_i b_i/a$ and $k_i - q_{i+1} = -q_i$. $\square$

The weights also explain the role of the Newton polynomials below. Write $\mu_i(g) = \int g d\mu_i$. For $i+1 < N$ and any polynomial g , the zero atom contributes nothing to $\mu_i(xg)$, and $p_i(j+1)w_i(j+1) = w_{i+1}(j)$ gives

$$\mu_i(xg(x)) = \mu_{i+1}(g(k_i + x)).$$

With Lemma 6.1(1) this gives the low moments

$$\mu_j(1) = \mu_j(x) = 1, \quad \mu_j(x^2) = k_j + 1, \quad \mu_j(x^3) = k_j^2 + 2k_j + k_{j+1} + 1 \quad (j+3 < N). \quad (6)$$

Once the state i is fixed we drop it from the notation and write $s_j = p_i(j) = q_{i+j} - q_i$; put $F_j(x) = \prod_{0 \leq l < j} (x - s_l)$, with $F_0 = 1$. Iterating the identity $\mu_i(xg(x)) = \mu_{i+1}(g(k_i + x))$ gives the *shift rule*

$$\mu_i(F_j(x)g(x)) = \mu_{i+j}(g(x + s_j)) \quad (i+j < N). \quad (7)$$

Indeed, each step removes the first factor of F_j and translates the remaining nodes by the next increment. Thus multiplication by a Newton polynomial advances along the orbit; this is why this basis is useful for bounding its Hankel determinants.

Theorem 6.2 (Hankel determinants). *For $1 \leq r \leq N-i$, the number $I_r(i) = a \det(h_i(j+l))_{0 \leq j,l < r}$ is a positive integer. Moreover $I_1(i) = b_i$, $I_2(i) = d_i$ (if $i+1 < N$), $I_3(i) = D_i$ (if $i+3 < N$), and $I_4(i) = L_i$ (if $i+5 < N$), where D_i and L_i are the curvature and the six-term integer of the blocks starting at b_i .*

Proof. Integrality. By Lemma 6.1(4), $h_i(s+1) + q_i h_i(s) \in \mathbb{Z}$. Multiply the Hankel matrix on the left by the unit lower triangular matrix that adds q_i times row $j-1$ to row j . This makes rows $1, \dots, r-1$ integral and leaves the determinant unchanged. Expand along row 0, whose entries are $(-q_i)^l b_i/a + (\text{integer})$; then $a \det$ is an integer.

Positivity. The Hankel matrix is the Gram matrix of $1, x, \dots, x^{r-1}$ in $L^2(\nu_i)$. Its quadratic form is $\sum_j w_i(j) q_{i+j}^{-1} (\sum_l c_l p_i(j)^l)^2$. If this vanishes, the polynomial $\sum_l c_l x^l$ vanishes at the $N-i \geq r$ distinct points $p_i(j)$, so it is zero, by the invertibility of the Vandermonde matrix. So the matrix is positive definite.

The first four. First, $I_1(i) = a h_i(0) = b_i$. For $i+1 < N$, write $y_j = b_{i+j}/a$. Then $h_i(0) = y_0$, $h_i(1) = y_1$ and $h_i(2) = k_i y_1 + y_2$. Hence

$$I_2(i) = a(y_0(k_i y_1 + y_2) - y_1^2) = a(y_0 - y_1) = d_i,$$

using $q_i y_0 + y_1 = q_{i+1} y_1 + y_2 = 1$. For the larger determinants we use the Newton basis. By Lemma 6.1(3), $\nu_i(xg(x)) = \nu_{i+1}(g(x + k_i))$, so ν_i obeys the same shift rule (7) as μ_i , and for $j \leq l$

$$\nu_i(F_j F_l) = \nu_{i+l}(F_j(x + s_l)) = \nu_{i+l}\left(\prod_{m < j} (x + s_l - s_m)\right).$$

Since $F_0, \dots, F_{r-1}$ are monic of degrees $0, \dots, r-1$, the Gram matrix $(\nu_i(F_j F_l))_{j,l < r}$ has the same determinant as the Hankel matrix. Its entries are explicit: $\nu_{i+l}(1) = y_l$, and higher moments of ν_{i+l} follow from the same rule. For $r = 3$ the matrix is

$$\begin{pmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 + k_i y_1 & y_3 + s_2 y_2 \\ y_2 & y_3 + s_2 y_2 & y_4 + (s_2 + k_{i+1} + k_{i+2}) y_3 + s_2 k_{i+1} y_2 \end{pmatrix}.$$

Using $y_{j+1} = 1 - q_{i+j} y_j$ for $0 \leq j \leq 3$ to express every entry in terms of y_0, q_i and the increments, and expanding, gives a times its determinant as $(b_i - b_{i+1})k_{i+1} - (b_{i+1} - b_{i+2})k_i = D_i$. At the end of the orbit this computation uses $h_{N-1}(s) = 0$ for $s \geq 1$, that is, $b_N = 0$.

For $r = 4$ we avoid the analogous 4×4 expansion by passing through the orthogonal cubic of ν_i . Let $i + 5 < N$, and for $m \leq 2$ write $k_m = k_{i+m}$, $d_m = d_{i+m}$, $r_m = k_m(s_m + s_{m+1}) + k_{m+1}$ and $G_m = F_{m+1} - F_m = F_m(x)(x - s_m - 1)$. The shift rule, $\nu_{i+m}(1) = b_{i+m}/a$, $\nu_{i+m}(x) = b_{i+m+1}/a$ and the low moments (6) at $j = i + m$ give

$$a \nu_i(G_m) = -d_m, \quad \mu_i(G_m) = 0, \quad \mu_i(xG_m) = k_m, \quad \mu_i(x^2G_m) = r_m + k_m \quad (m \leq 2).$$

Put $\mathbf{d} = (d_0, d_1, d_2)$, $\mathbf{k} = (k_0, k_1, k_2)$, $\mathbf{r} = (r_0, r_1, r_2)$ and $W = \mathbf{d} \times \mathbf{k}$, whose last coordinate is $W_2 = d_0k_1 - d_1k_0 = D_i \geq 1$. The monic cubic $\pi = \sum_{m \leq 2} (W_m/W_2)G_m$ satisfies $\nu_i(\pi) = -\mathbf{d} \cdot \mathbf{W}/(aW_2) = 0$, $\mu_i(\pi) = 0$ and $\mu_i(x\pi) = \mathbf{k} \cdot \mathbf{W}/W_2 = 0$. Since $\mu_i = (q_i + x)\nu_i$, we have $\nu_i(xg) = \mu_i(g) - q_i\nu_i(g)$, so $\nu_i(x\pi) = \nu_i(x^2\pi) = 0$: π is the monic orthogonal cubic of ν_i . Hence

$$\nu_i(\pi^2) = \nu_i(x^3\pi) = \mu_i(x^2\pi) - q_i\nu_i(x^2\pi) = \frac{\mathbf{r} \cdot \mathbf{W} + \mathbf{k} \cdot \mathbf{W}}{W_2} = \frac{\det(\mathbf{d}, \mathbf{k}, \mathbf{r})}{D_i}.$$

In the basis $1, x, x^2, \pi$ the Gram matrix of ν_i is block diagonal and has the same determinant as the Hankel matrix, so $I_4(i) = I_3(i) \nu_i(\pi^2) = \det(\mathbf{d}, \mathbf{k}, \mathbf{r})$, using $I_3(i) = D_i$. Finally $\det(\mathbf{d}, \mathbf{k}, \mathbf{r}) = L_i$. Expanding, with the nodes cancelling,

$$k_1r_2 - k_2r_1 = \sigma(k_1, k_2, k_3), \quad k_0r_1 - k_1r_0 = \sigma(k_0, k_1, k_2),$$

and $k_1(k_0r_2 - k_2r_0) = k_0(k_1r_2 - k_2r_1) + k_2(k_0r_1 - k_1r_0)$. Multiplying the expansion of $\det(\mathbf{d}, \mathbf{k}, \mathbf{r})$ along its first row by k_1 therefore gives

$$k_1 \det(\mathbf{d}, \mathbf{k}, \mathbf{r}) = (d_0k_1 - d_1k_0) \sigma(k_1, k_2, k_3) - (d_1k_2 - d_2k_1) \sigma(k_0, k_1, k_2) = k_1L_i$$

by Lemma 4.1(1), since $d_0k_1 - d_1k_0$ and $d_1k_2 - d_2k_1$ are the curvatures D and E of that lemma. $\square$

The ranges $i + 3 < N$ and $i + 5 < N$ cannot be relaxed to all i at which D_i or L_i is defined. For $a = 926$ and $b = 229$ the orbit is $229, 10, 6, 2$, so $N = 4$, and $I_3(1) = 984$ while $D_1 = 988$.

Remark 6.3. In these index ranges, the positivity of D and L , proved earlier by explicit identities, is an instance of the positivity of Gram determinants. The integrality of L (Lemma 4.1(1)) is an instance of the integrality of I_4 . A priori $I_r(i)$ involves the moments $h_i(s)$ for $s \leq 2r - 2$, and hence $b_i, \dots, b_{i+2r-2}$. In fact $I_3(i)$ depends only on b_i, b_{i+1}, b_{i+2} and q_i, q_{i+1}, q_{i+2} . Similarly $I_4(i)$ depends only on $b_i, \dots, b_{i+3}$ and $q_i, \dots, q_{i+4}$: the later remainders cancel.

7 Patterns

In this section we develop the patterns used in Section 8, together with the bounds on their size and number. We keep the notation of Section 6: $b_0 > b_1 > \dots > b_{N-1} > 0$ is the positive part of the orbit, $N = P(a, b)$, and $q_i = \lfloor a/b_i \rfloor$.

7.1 Heuristics

The exponent $3/10$ of Theorem 1.4 comes from a single balanced configuration: $T \asymp a^{3/10}$ remainders in a window $[A, 2A)$ with $A \asymp a^{3/5}$, quotients $q \asymp a^{2/5}$ and typical increments $k \asymp a^{1/10}$. At this scale the rank-four Hankel invariant $I_4(i) = L_i$ has size $\asymp aH^6/q^4 \asymp 1$, where H bounds the increments. The idea is that a positive integer of bounded size carries information. The labels $I_4(i), I_4(i+1), I_4(i+2)$, together with seven further 3×3 minors of the same size, form a *pattern* (made precise in Section 7.6). States with a given pattern satisfy two linear equations and one

quadratic equation in five consecutive increments (Lemma 7.1), and the quotient q_i is a root of a quartic fixed by the pattern and the increments (Lemma 7.3). So the problem becomes one of counting patterns, and the states belonging to each pattern. Counting the latter with these equations alone already improves on 3/10 (the first rows of the table in Section 1). To go further, the linear and quadratic relations alone do not suffice: for some patterns they admit constant increments $k_j = t$ for every t (an example follows Lemma 8.1). Section 8 therefore adds the fixed-dividend equation, an equation in the nodes alone that involves a ; it excludes such families.

The rest of this section builds the pattern matrix and identifies its cofactors with the orthogonal quartic of ν_i (Section 7.2), ties q_i to the pattern by a quartic equation (Section 7.3), proves the signs and the supply of good starts used throughout Section 8 (Sections 7.4 and 7.5), and bounds the size and number of patterns (Sections 7.6 and 7.7).

7.2 The pattern matrix

Fix a state i with $i + 6 < N$ and write, for $m \geq 0$,

$$\begin{aligned} k_m &= q_{i+m+1} - q_{i+m}, & d_m &= b_{i+m} - b_{i+m+1}, \\ s_m &= k_0 + \cdots + k_{m-1}, & r_m &= k_m(s_m + s_{m+1}) + k_{m+1}. \end{aligned}$$

All indices here are local to the chosen state, and the definitions apply whenever the indicated orbit entries exist; the empty sum gives $s_0 = 0$. So $s_m = q_{i+m} - q_i$, $r_0 = k_0^2 + k_1$, $r_1 = k_1(2k_0 + k_1) + k_2$, and so on. In this local notation the low moments (6) at state $i + m$ read $\mu_{i+m}(x^2) = k_m + 1$ and $\mu_{i+m}(x^3) = k_m^2 + 2k_m + k_{m+1} + 1$.

The *pattern* $p(i)$ of the state is the vector of the ten 3×3 minors of the 3×5 matrix with rows $(d_0, \dots, d_4)$, $(k_0, \dots, k_4)$ and $(r_0, \dots, r_4)$ (Section 7.6). Every object built from the pattern below (the pattern integers, the normalised parameters, the Newton coefficients γ and, in Section 10, G and ℓ) is a function of $p(i)$. Consider the 3×4 integer matrix

$$\mathcal{M}_i = \begin{pmatrix} d_0 & d_1 & d_2 & d_3 \\ k_0 & k_1 & k_2 & k_3 \\ r_0 & r_1 & r_2 & r_3 \end{pmatrix},$$

and let $U_m(i) = (-1)^{m+1} \det(\mathcal{M}_i \text{ with column } m \text{ deleted})$ for $0 \leq m \leq 3$. The three rows are, up to sign and the factor a , the moments of $F_{m+1} - F_m$ against ν_i , $x\mu_i$ and $(x^2 - x)\mu_i$ (proof of Lemma 7.2); this is why the cofactor vector describes an orthogonal polynomial.

Lemma 7.1 (Cofactor relations). *1. $U_3(i) = L_i = I_4(i)$ and $U_0(i) = -L_{i+1} = -I_4(i+1)$, where L is the six-term integer of Lemma 4.1.*

2. $\sum_m U_m(i) d_m = \sum_m U_m(i) k_m = \sum_m U_m(i) r_m = 0$.

Proof. (2) is the Laplace expansion of a 4×4 determinant with a repeated row. For (1), the identity $\det(\text{columns } 0, 1, 2) = L_i = I_4(i)$ was proved in the case $r = 4$ of Theorem 6.2. Passing from i to $i + 1$ shifts the d - and k -rows by one column, and, since $s_m(i + 1) = s_{m+1}(i) - k_0$, it replaces r_m by $r_{m+1} - 2k_0k_{m+1}$. This row operation, which subtracts $2k_0$ times the k -row, does not change the determinant, so $\det(\text{columns } 1, 2, 3) = L_{i+1}$. $\square$

For example, the positive part of the orbit with $(a, b) = (95, 61)$ is 61, 34, 27, 14, 11, 7, 4, 3, 2, 1. Its first pattern matrix and cofactor vector are

$$\mathcal{M}_0 = \begin{pmatrix} 27 & 7 & 13 & 3 \\ 1 & 1 & 3 & 2 \\ 2 & 6 & 23 & 29 \end{pmatrix}, \quad U(0) = (-81, 833, -296, 68).$$

Thus $L_0 = 68$, $L_1 = 81$, and each row of $\mathcal{M}_0$ is orthogonal to $U(0)$.

The cofactor vector has a meaning, which is the source of everything that follows. Recall the Newton polynomials $F_m(x) = \prod_{l < m} (x - s_l)$ and the measures μ_i, ν_i of Section 6.

Lemma 7.2 (The orthogonal quartic). *Let $i + 6 < N$ and $C_m = U_m(i)/U_3(i)$. Then*

$$\pi_i = \sum_{m=0}^3 C_m (F_{m+1} - F_m)$$

is the monic orthogonal polynomial of degree four for ν_i . Writing $\pi_i = \sum_{l \leq 4} \gamma_l F_l$, with Newton coefficients γ_l , one has $U_m(i)/I_4(i) = \gamma_{m+1} + \dots + \gamma_4$ for $0 \leq m \leq 3$, and $\gamma_0 = I_4(i+1)/I_4(i)$.

Proof. Note that $F_{m+1} - F_m = F_m(x)(x - s_m - 1)$. The shift rule (7) holds for ν_i as well as for μ_i (proof of Theorem 6.2). Together with $\nu_{i+m}(1) = b_{i+m}/a$, $\nu_{i+m}(x) = b_{i+m+1}/a$ and the low moments (6) at $j = i + m$, it gives, for $m \leq 3$,

$$\begin{aligned} \nu_i(F_m(x)(x - s_m - 1)) &= \nu_{i+m}(x - 1) = -d_m/a, \\ \mu_i(F_m(x)(x - s_m - 1)) &= \mu_{i+m}(x - 1) = 0, \\ \mu_i(xF_m(x)(x - s_m - 1)) &= \mu_{i+m}((x + s_m)(x - 1)) = k_m, \\ \mu_i(x^2F_m(x)(x - s_m - 1)) &= \mu_{i+m}((x + s_m)^2(x - 1)) = r_m + k_m. \end{aligned}$$

By Lemma 7.1(2), the three rows of $\mathcal{M}_i$ are orthogonal to (C_m) . Hence $\nu_i(\pi_i) = \mu_i(\pi_i) = \mu_i(x\pi_i) = \mu_i(x^2\pi_i) = 0$. Since $\mu_i = (q_i + x)\nu_i$, we have $\nu_i(xg) = \mu_i(g) - q_i\nu_i(g)$, and therefore $\nu_i(x^j\pi_i) = 0$ for $j \leq 3$, inductively in j . The coefficient of F_4 in π_i is $C_3 = 1$, so π_i is monic of degree four. Finally $\gamma_l = C_{l-1} - C_l$ with $C_{-1} = C_4 = 0$, which gives the tail sums, and $\gamma_0 = -C_0 = I_4(i+1)/I_4(i)$ by Lemma 7.1(1). $\square$

The *pattern integers* of i are the seven integers

$$\theta(i) = (L_i, L_{i+1}, L_{i+2}, U_1(i), U_2(i), U_1(i+1), U_2(i+1)),$$

and we write $\theta(i) = (N_0, N_1, N_2, u_1, u_2, v_1, v_2)$; the subscripted N_j have nothing to do with the orbit length N . By Lemma 7.1, $\theta(i)$ determines $U(i) = (-N_1, u_1, u_2, N_0)$ and $U(i+1) = (-N_2, v_1, v_2, N_1)$ completely, and the k -relations at i and $i+1$ read

$$N_1k_0 = u_1k_1 + u_2k_2 + N_0k_3, \quad N_1k_4 = N_2k_1 - v_1k_2 - v_2k_3.$$

7.3 The normalisation identity and the label bound

Lemma 7.3 (The normalisation identity). *Let $i + 6 < N$, let π_i be as in Lemma 7.2, and let $H_4 = \det(\mu_i(x^{j+l}))_{j,l < 4}$, a polynomial in $k_0, \dots, k_4$ independent of i . Then*

$$I_4(i) \pi_i(-q_i) = a H_4(k_0, \dots, k_4).$$

As a polynomial in q , $I_4(i)\pi_i(-q)$ has degree four and leading coefficient $I_4(i) = L_i \geq 1$.

This is Christoffel's formula for the modification $\mu_i = (q_i + x)\nu_i$ of the measure; we include the short proof.

Proof. Write $h_s = h_i(s) = \nu_i(x^s)$ and $M_s = \mu_i(x^s)$, so that $h_{s+1} + q_i h_s = M_s$. Expanding along the last row shows that

$$\det(h_{j+l})_{j,l < 4} \pi_i(x) = \det \begin{pmatrix} h_0 & h_1 & \cdots & h_4 \\ \vdots & & & \vdots \\ h_3 & h_4 & \cdots & h_7 \\ 1 & x & \cdots & x^4 \end{pmatrix},$$

since the right side is a polynomial of degree four with leading coefficient $\det(h_{j+l})_{j,l < 4}$ that is orthogonal to $1, x, x^2, x^3$ (replacing the last row by a moment row gives a repeated row). Put $x = -q_i$ and, for $l = 4, 3, 2, 1$ in turn, add q_i times column $l - 1$ to column l . The last row becomes $(1, 0, 0, 0)$, and the entry in row $j < 4$ and column $l \geq 1$ becomes $h_{j+l} + q_i h_{j+l-1} = M_{j+l-1}$. Expanding along the last row gives $\det(h_{j+l}) \pi_i(-q_i) = \det(M_{j+l})_{j,l < 4} = H_4$. Multiply by a and use $a \det(h_{j+l}) = I_4(i)$. The leading coefficient is that of the monic π_i . $\square$

Lemma 7.4 (Labels). *Let $i + 6 < N$ and put $\mathcal{G} = \prod_{j < 4} \mu_i(F_j^2)$. Explicitly,*

$$\mathcal{G}(k_0, \dots, k_4) = (k_0 + 1)(k_1 s_2 + s_2 + k_1 + k_2 + 1)(k_2(k_1 + k_2)s_3 + g_3),$$

where $g_3 = s_3(k_1 + k_2) + s_3 k_2 + (k_1 + k_2)k_2 + (k_0 + 2k_1 + 3k_2)(k_3 + 1) + k_3^2 + 2k_3 + k_4 + 1$. Then $L_i q_i^4 \leq a \mathcal{G}(k_0, \dots, k_4)$.

Proof. The Newton polynomials $F_0, \dots, F_3$ are monic of degrees $0, \dots, 3$. So the matrix $(\nu_i(F_j F_l))$ is $\mathcal{U}\mathcal{H}\mathcal{U}^\top$, where $\mathcal{H}$ is the Hankel matrix of ν_i and $\mathcal{U}$ is unit lower triangular. Hence it has determinant $I_4(i)/a$, and it is positive definite, as in the proof of Theorem 6.2. By Hadamard's inequality,

$$\frac{I_4(i)}{a} \leq \prod_{j < 4} \nu_i(F_j^2) \leq \prod_{j < 4} \frac{\mu_i(F_j^2)}{q_i},$$

since $\nu_i = \mu_i/(q_i + x) \leq \mu_i/q_i$ on the nonnegative support. By the shift rule (7) and the low moments (6), the product $\prod_{j < 4} \mu_i(F_j^2)$ expands to the displayed formula. $\square$

7.4 Signs

Throughout this subsection the pattern and increments arise from the same state, so the cofactor relations of Lemma 7.1 remain in force. Besides the polynomial σ of Section 4, put $\tilde{\sigma}(y_1, y_2, y_3, y_4) = y_1^2 y_3 + (2y_1 - 1)y_2 y_3 + y_1 y_3^2 + y_1 y_4$. Both σ and $\tilde{\sigma}$ are positive at positive integers.

Lemma 7.5 (Signs). *If $N_0, N_1 \geq 1$ and $k_0, \dots, k_4 \geq 1$, then $u_1 > 0 > u_2$.*

Proof. The 2×2 minors of the k - and r -rows do not involve the nodes: expanding $r_m = k_m(s_m + s_{m+1}) + k_{m+1}$ with $s_{m+1} = s_m + k_m$ gives

$$k_m r_{m+1} - k_{m+1} r_m = \sigma(k_m, k_{m+1}, k_{m+2}), \quad k_m r_{m+2} - k_{m+2} r_m = \tilde{\sigma}(k_m, k_{m+1}, k_{m+2}, k_{m+3}).$$

The k - and r -relations of Lemma 7.1 read $u_1 k_1 + u_2 k_2 = N_1 k_0 - N_0 k_3$ and $u_1 r_1 + u_2 r_2 = N_1 r_0 - N_0 r_3$. Solving them for (u_1, u_2) by Cramer's rule and using these minors gives the exact identities

$$\begin{aligned} u_1 \sigma(k_1, k_2, k_3) &= N_1 \tilde{\sigma}(k_0, k_1, k_2, k_3) + N_0 \sigma(k_2, k_3, k_4), \\ -u_2 \sigma(k_1, k_2, k_3) &= N_1 \sigma(k_0, k_1, k_2) + N_0 \tilde{\sigma}(k_1, k_2, k_3, k_4), \end{aligned}$$

since the Cramer determinant is $k_1 r_2 - k_2 r_1 = \sigma(k_1, k_2, k_3)$. $\square$

7.5 Good starts

Lemma 7.6 (Good starts). *Let $b_{i_0}, \dots, b_{i_0+T-1}$ be a window run in $[A, 2A)$ with $T \geq 36$, and put $H = \lfloor 28a/(AT) \rfloor + 1$. Call a start i good if $i_0 \leq i$, $i+9 \leq i_0+T$ and $q_{j+1} - q_j \leq H$ for $i \leq j \leq i+6$, that is, $k_0, \dots, k_6 \leq H$ at state i . Then at least $T/2$ starts are good, and every good start satisfies $i+8 < N$ and $b_i, b_{i+1}, b_{i+2} < 2A$.*

Proof. Let $\mathcal{B}$ be the set of $j \in [i_0, i_0+T-1)$ with $q_{j+1} - q_j > H$. Then $|\mathcal{B}|(H+1) \leq q_{i_0+T-1} - q_{i_0} \leq a/A$. Since $H \cdot AT > 28a$, this gives $28|\mathcal{B}| \leq T$. Each $j \in \mathcal{B}$ spoils at most seven starts, so at least $T - 9 - 7|\mathcal{B}| \geq T/2$ starts are good, as $T \geq 36$. The last assertion is immediate, since the states $i, \dots, i+8$ lie in the window. $\square$

7.6 The full minor vector

The seven pattern integers are not independent: they are coordinates of a single vector of 3×3 minors. If all ten minors are at most X in absolute value, there are only $X^{5+o(1)}$ such vectors (Proposition 7.9), rather than the X^7 choices of seven independent integers. Let $\widehat{\mathcal{M}}_i$ be the 3×5 matrix with rows $(d_0, \dots, d_4)$, $(k_0, \dots, k_4)$ and $(r_0, \dots, r_4)$, and let $p_{c_1 c_2 c_3}$ be its minor on columns $c_1 < c_2 < c_3$. We define the *pattern* of a state i to be the vector $p(i)$ of these ten minors. Expanding the cofactors above shows that it determines the pattern integers:

$$\theta(i) = (p_{012}, p_{123}, p_{234}, p_{023}, -p_{013}, p_{134}, -p_{124}).$$

Lemma 7.7 (Expansion). *Let $q = q_i$ and $s_j = q_{i+j} - q$. If $i+j+3 < N$ then*

$$\frac{b_{i+j}}{a} = \frac{1}{q} - \frac{1+s_j}{q^2} + \frac{s_j^2 + 2s_j + k_j + 1}{q^3} - e_j, \quad 0 \leq e_j \leq \frac{\mu_{i+j}((x+s_j)^3)}{q^4}.$$

Proof. Apply μ_{i+j} to $\frac{1}{q+y} = \frac{1}{q} - \frac{y}{q^2} + \frac{y^2}{q^3} - \frac{y^3}{q^3(q+y)}$ with $y = s_j + x \geq 0$, using $b_{i+j}/a = \nu_{i+j}(1) = \mu_{i+j}(1/(q+s_j+x))$ and (6). $\square$

Proposition 7.8 (All minors). *Let $i+8 < N$, $b_i < 2A$ and $k_0, \dots, k_6 \leq H$ at state i . Then every minor of $\widehat{\mathcal{M}}_i$ satisfies $|p| \leq c_X A^4 H^6 / a^3$, for an absolute constant c_X .*

Proof. Subtracting the expansions of Lemma 7.7 for j and $j+1$ gives $d_j/a = k_j/q^2 - (r_j + k_j)/q^3 + e_{j+1} - e_j$. So replacing the first row by $d'_j = d_j - a(q^{-2} - q^{-3})k_j + aq^{-3}r_j = a(e_{j+1} - e_j)$ changes no minor. Now $|d'_j| \ll aH^3/q^4$ by the error bound of Lemma 7.7, while $|k_j| \leq H$ and $|r_j| \leq 10H^2$, so $|p| \ll aH^6/q^4$. Finally $a < 4Aq$ (Lemma 2.4). $\square$

7.7 Counting minor vectors

Proposition 7.9 (Minor vectors). *The number of vectors of minors of integer 3×5 matrices with all coordinates in $[-X, X]$ is $\ll 1 + \sum_{\kappa \leq X} \kappa^4 \tau(\kappa) \ll X^{5+o(1)}$.*

The minor vectors are the integral points of the affine cone over the Grassmannian $\text{Gr}(3, 5)$ in its Plücker embedding, which we call the *Plücker cone*. This cone has dimension seven, but its integral points of height at most X number about X^5 , as for rational points of bounded height on $\text{Gr}(3, 5)$, whose anticanonical class is five times the hyperplane class [7]. So the exponent 5 is optimal. The proof uses the standard affine chart of the Grassmannian.

Proof. For a nonzero vector p choose a pivot π with $\kappa = |p_\pi|$ maximal, and let c_4, c_5 be the two columns outside π . Let $\mathbf{u}_r$ ($\mathbf{v}_r$) be the minor with the r -th pivot column replaced by c_4 (c_5), with the sign of the column permutation. Write $\mathcal{M} = \mathcal{M}_\pi(I | B)$ up to a permutation of columns, where $\mathcal{M}_\pi$ is the pivot block; by Cramer's rule the two columns of B are $\mathbf{u}/p_\pi$ and $\mathbf{v}/p_\pi$. The minors of $(I | B)$ are ± 1 , the entries of B , and the 2×2 minors of B ; multiplying by $\det \mathcal{M}_\pi = p_\pi$ shows that the three minors p_{w_r} using both c_4 and c_5 , where w_r consists of c_4 , c_5 and the r th pivot column, satisfy $p_\pi p_{w_r} = \pm(\mathbf{u} \times \mathbf{v})_r$. So p is determined by $(\pi, \text{sgn } p_\pi, \kappa, \mathbf{u}, \mathbf{v})$, where $|\mathbf{u}_r|, |\mathbf{v}_r| \leq \kappa$ and κ divides every coordinate of $\mathbf{u} \times \mathbf{v}$.

Fix $\kappa \geq 1$ and $\mathbf{u}$, and let $g = \gcd(\kappa, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, $\ell = \kappa/g$, $\mathbf{u} = g\mathbf{u}_0$. By Bézout there are $z \in \mathbb{Z}^3$ and $t \in \mathbb{Z}$ with $z \cdot \mathbf{u}_0 + t\ell = 1$. If $\kappa \mid \mathbf{u} \times \mathbf{v}$, then $\ell \mid \mathbf{u}_0 \times \mathbf{v}$, and the identity $\mathbf{v} - (z \cdot \mathbf{v})\mathbf{u}_0 = t\ell \mathbf{v} + z \times (\mathbf{v} \times \mathbf{u}_0)$ shows $\mathbf{v} = c\mathbf{u}_0 + \ell \mathbf{y}$ with $c = z \cdot \mathbf{v}$ and $\mathbf{y} \in \mathbb{Z}^3$. Hence $\mathbf{v}$ lies in one of ℓ classes modulo ℓ , each meeting $[-\kappa, \kappa]^3$ in at most $(3g)^3$ points: at most $27\kappa g^2$ choices. Summing over $\mathbf{u}$, grouped by $g = e \mid \kappa$, gives $\sum_{\mathbf{u}} g^2 \leq \sum_{e \mid \kappa} e^2 (3\kappa/e)^3 \leq 27\kappa^3 \tau(\kappa)$, so at most $729\kappa^4 \tau(\kappa)$ pairs $(\mathbf{u}, \mathbf{v})$. There are ten pivots and two signs. $\square$

8 Joint integer invariants and the optimised bound

We now prove Theorem 1.2. Fix a window run in $[A, 2A)$ with $T \geq 36$ terms. Throughout this section constants may depend on fixed positive integers m, n (the powers of Section 8.8) and on a loss $\eta > 0$, but never on the dividend, the orbit or the pattern. The argument has six steps.

1. For a fixed pattern, the nodes of every good start lie on at most 24 curves (Section 8.1).
2. On these curves the quotient q_i and the curvature $I_3(i)$ are rational functions of the nodes. Components on which the reconstruction identities fail contribute $O(1)$ starts (Section 8.2).
3. The directions at infinity of the curves are controlled by the Newton coefficients of the pattern. This makes nonexceptional components have even degree (Section 8.3), so that the orders $3e/2$ and $5e/2$ below are integers and every component left after the conics has degree at least four; it also shows that exceptional patterns are few, and excludes lines (Section 8.4).
4. The reconstructed functions, and hence two families of new integer coordinates W and J , have prescribed orders at infinity (Sections 8.5 and 8.6).
5. Each nonexceptional conic contributes $a^{o(1)}$ starts (Section 8.7).
6. On the remaining components, adjoining W_j^m and J^n raises the degree. The curve-counting theorem and a balancing of heights then give the exponent (Sections 8.8 and 8.9).

Notation. In this section

- $\theta = (N_0, N_1, N_2, u_1, u_2, v_1, v_2)$ are the pattern integers of Section 7, with $N_0 = I_4(i)$, and ρ, φ, χ and ρ', φ', χ' the normalised parameters (8);
- s_j are the nodes, k_j the increments, P_4, R_4 the two orthogonal quartics, Ψ the quartic norm and f the pattern quadratic;
- $D_{jl} = d_j k_l - d_l k_j$ are the 2×2 minors of drops and increments, so that $D_{01} = I_3(i)$ is the curvature of Section 3;
- Q and E_0 are the reconstructed quotient and curvature, $W = (W_0, W_1, W_2)$ and J the joint coordinates, and S, Y their height bounds;

- ord_P is the order at a point P , positive at zeros and negative at poles, e is a multiplicity at infinity, d the degree of a component (not to be confused with the drops d_j), δ the degree of a graph and $\delta_{\min} = 10n + 7m$ its lower bound (Lemma 8.13).

8.1 Good starts and the pattern curve

Put

$$H = \left\lfloor \frac{28a}{AT} \right\rfloor + 1, \quad X = \left\lfloor \frac{c_X A^4 H^6}{a^3} \right\rfloor,$$

with c_X the constant of Proposition 7.8. Let the good starts be as in Lemma 7.6; at least $T/2$ starts are good. By Proposition 7.8, the minor vector of a good start has all coordinates at most X , and $X \geq 1$ because its coordinate $p_{012} = L_i$ is a positive integer. The pattern of a good start (Section 7.6) therefore has height at most X , and it determines $\theta(i)$. By Proposition 7.9, there are $O_\eta(a^\eta X^5)$ patterns.

The *nodes* of state i are $s_j = q_{i+j} - q_i$. They are the support points of the measure μ_i of Section 6, and the interpolation nodes of the Newton polynomials of Section 7. The three *free nodes* s_1, s_2, s_3 of a good start are integers in $[0, 3H]$. Distinct good starts with the same pattern have distinct free-node triples: the pattern and the free nodes determine q_i , which strictly increases along the orbit. (That q_i is determined is shown in Section 8.2, where it is the rational function Q ; the argument there uses only Lemma 8.1(1) below.)

We now write, as functions of the three free nodes, the objects that at a good start are the orthogonal quartics of states i and $i+1$ and the norm $\nu_i(P_4^2)$; the pattern curve is cut out by the pattern quadratic and a resultant identity among them. For the pattern integers θ , put

$$(\rho, \varphi, \chi) = \frac{(N_1, u_1, u_2)}{N_0}, \quad (\rho', \varphi', \chi') = \frac{(N_2, v_1, v_2)}{N_1}, \quad (8)$$

the *normalised parameters* of the pattern. Then $\rho > 0$, and $\varphi > 0 > \chi$ by Lemma 7.5. Treat the free nodes s_1, s_2, s_3 as coordinates and define

$$s_0 = 0, \quad s_4 = (\rho + \varphi)s_1 - (\varphi - \chi)s_2 - (\chi - 1)s_3, \quad k_j = s_{j+1} - s_j \quad (0 \leq j \leq 3),$$

$$k_4 = \rho'k_1 - \varphi'k_2 - \chi'k_3, \quad s_5 = s_4 + k_4.$$

At a good start these are the actual nodes and increments, by the two k -relations of Lemma 7.1; the first is equivalent to $\rho k_0 = \varphi k_1 + \chi k_2 + k_3$. Let

$$\gamma = (\rho, -\rho - \varphi, \varphi - \chi, \chi - 1, 1),$$

and let γ' be the same expression in ρ', φ', χ' . Put

$$P_4(x) = \sum_{j=0}^4 \gamma_j \prod_{\ell=0}^{j-1} (x - s_\ell), \quad R_4(x) = \sum_{j=0}^4 \gamma'_j \prod_{\ell=1}^j (x - s_\ell).$$

By Lemma 7.2, at a good start P_4 is the monic orthogonal polynomial π_i of degree four for ν_i , and $R_4(x) = \pi_{i+1}(x - k_0)$ is the corresponding polynomial of the next state, written in the variable of the current state. Put $Z = xR_4$ and $f = \sum_{j=0}^4 \gamma_j(s_j^2 + k_j)$, the *pattern quadratic*, and let

$$\Psi = \sum_{j=0}^4 \gamma_j(s_j^3 + 3s_j^2 + 3s_j(1 + k_j) + (1 + k_j)^2) + \sum_{j=0}^3 \gamma_j k_{j+1}$$

$$+ \rho' r_1 - \varphi' r_2 - \chi' r_3 - k_4(2s_4 + k_4),$$

where $r_j = k_j(s_j + s_{j+1}) + k_{j+1}$ for $j = 1, 2, 3$. This is a cubic polynomial in the free nodes, the *quartic norm*: at a good start it is the squared norm $\nu_i(P_4^2)$ of the orthogonal quartic (Lemma 8.1).

The degree of an affine curve is that of its projective closure, equivalently the leading coefficient of its affine Hilbert function. “Integral” means that the defining ideal is prime over $\mathbb{Q}$; it does not assert geometric irreducibility or a rational parametrisation. For a monic polynomial g we write $\text{Res}(g, h) = \prod_{g(\theta)=0} h(\theta)$, the product over the roots of g with multiplicity. For monic g, h of degrees m, n this is $(-1)^{mn} \text{Res}(h, g)$, and $\text{Res}(g, h) = \text{Res}(g, h + ug)$ for every polynomial u .

Lemma 8.1 (The pattern curve). *1. At a good start i , $\Psi = \nu_i(P_4^2) > 0$, and the free nodes satisfy*

$$f = 0, \quad N_0 \text{Res}(P_4, Z) = a\Psi^4. \quad (9)$$

We call the second equation the fixed-dividend equation.

2. For every pattern with $\rho, \varphi > 0 > \chi$, the two equations (9) have degrees two and at most twelve and no common factor. Their common zero set in $\mathbb{A}^3$ is the union of at most twenty-four integral affine curves over $\mathbb{Q}$, of total degree at most twenty-four.

Proof. (1) Since $r_m = s_{m+1}^2 - s_m^2 + k_{m+1}$, summation by parts and the k -relation at state i show that $N_0 f$ is the r -relation $\sum_m U_m(i) r_m = 0$ of Lemma 7.1. Next, the shift rule $\mu_i(g \prod_{\ell < j} (x - s_\ell)) = \mu_{i+j}(g(x + s_j))$ and the low moments (6) give $\mu_i(x^3 P_4) = \sum_j \gamma_j (s_j^3 + 3s_j^2 + 3s_j(1 + k_j) + (1 + k_j)^2 + k_{j+1})$. Eliminating k_5 with the r -relation at state $i + 1$, together with the k -relation at $i + 1$ that defines k_4 , turns this into Ψ . Since $\nu_i(xg) = \mu_i(g) - q_i \nu_i(g)$ and P_4 is orthogonal to x^3 , $\nu_i(P_4^2) = \nu_i(x^4 P_4) = \mu_i(x^3 P_4) = \Psi$, and this is positive as in the proof of Theorem 6.2.

For the resultant, Lemma 6.1(3) gives $\nu_i(xg(x)) = \nu_{i+1}(g(x + k_0))$. Hence, for $j < 4$, $\nu_i(x^j Z) = \nu_{i+1}((x + k_0)^j \pi_{i+1}(x)) = 0$. So the monic quintic Z differs from the monic orthogonal polynomial $\pi^{(5)}$ of degree five for ν_i by a multiple of P_4 , and $\text{Res}(P_4, Z) = \text{Res}(P_4, \pi^{(5)})$. Let $\|\pi^{(j)}\|^2$ be the squared ν_i -norm of the monic orthogonal polynomial $\pi^{(j)}$ of degree j , so $\pi^{(4)} = P_4$. The three-term recurrence $\pi^{(j+1)} = (x - \mathbf{a}_j)\pi^{(j)} - \mathbf{b}_j\pi^{(j-1)}$, with $\mathbf{b}_j = \|\pi^{(j)}\|^2 / \|\pi^{(j-1)}\|^2$, gives $\pi^{(j+1)} \equiv -\mathbf{b}_j\pi^{(j-1)} \pmod{\pi^{(j)}}$, so $\text{Res}(\pi^{(j)}, \pi^{(j+1)}) = (-\mathbf{b}_j)^j (-1)^{j(j-1)} \text{Res}(\pi^{(j-1)}, \pi^{(j)})$. Reducing from $j = 4$ down to $j = 1$, the signs multiply to $(-1)^{4+3+2+1} = 1$, and

$$\text{Res}(P_4, \pi^{(5)}) = \prod_{j=1}^4 \mathbf{b}_j^j = \frac{\|P_4\|^8}{\prod_{j<4} \|\pi^{(j)}\|^2} = \frac{\Psi^4}{\det(\nu_i(x^{j+l}))_{j,l<4}} = \frac{a\Psi^4}{I_4(i)}.$$

Since $N_0 = I_4(i)$, this is the second equation.

(2) The first equation has degree two. For the second, Ψ is cubic, and we show that the resultant has total degree at most ten in the nodes. Since $Z = xR_4$ and the product of the roots of P_4 is $P_4(0) = \gamma_0$ (every term with $j \geq 1$ has the factor $x - s_0 = x$), $\text{Res}(P_4, Z) = \gamma_0 \text{Res}(P_4, R_4)$.

We use the Newton basis $N_\alpha(x) = \prod_{1 \leq l \leq \alpha} (x - s_l)$ at the nodes $s_1, \dots, s_4$. From $F_{j+1} = xN_j = N_{j+1} + s_{j+1}N_j$,

$$P_4 = \sum_{j \leq 4} a_j N_j, \quad R_4 - P_4 = \sum_{\alpha \leq 3} c_\alpha N_\alpha, \quad a_j = \gamma_j + \gamma_{j+1}s_{j+1}, \quad c_\alpha = \gamma'_\alpha - a_\alpha,$$

with $\gamma_5 = 0$, so $a_4 = 1$; every a_j and c_α has degree at most one in the nodes, with top-degree parts $[a_j]_1 = \gamma_{j+1}s_{j+1}$ and $[c_\alpha]_1 = -\gamma_{\alpha+1}s_{\alpha+1}$. The Bézout matrix $\mathcal{B} = (\mathcal{B}_{\alpha\beta})_{\alpha,\beta < 4}$ is defined by

$$\frac{P_4(x)R_4(y) - P_4(y)R_4(x)}{x - y} = \sum_{\alpha,\beta < 4} \mathcal{B}_{\alpha\beta} N_\alpha(x) N_\beta(y). \quad (10)$$

Its determinant is $\text{Res}(P_4, R_4)$: in the monomial basis this is the classical Bézout formula for two monic quartics, and the change to the Newton basis is unitriangular in both variables. Since $R_4 = P_4 + \sum_{\alpha} c_{\alpha} N_{\alpha}$, the left side of (10) is $\sum_{j>k} (a_j c_k - a_k c_j) K_{jk}$, where $c_4 = 0$ and $K_{jk} = (N_j(x)N_k(y) - N_j(y)N_k(x))/(x - y)$. For $j > k$, telescoping $\prod_{l=k+1}^j (x - s_l) - \prod_{l=k+1}^j (y - s_l)$ gives

$$K_{jk} = \sum_{m=k+1}^j N_{m-1}(x) Y_{kmj}(y), \quad Y_{kmj}(y) = N_k(y) \prod_{l=m+1}^j (y - s_l). \quad (11)$$

The polynomial Y_{kmj} is monic of degree $k + j - m$, homogeneous of that degree in y and the nodes, and divisible by N_k . So its coefficient of N_{β} is homogeneous of degree $k + j - m - \beta$ in the nodes, and vanishes unless $\beta \geq k$. Thus the triple (j, k, m) contributes to $\mathcal{B}_{\alpha\beta}$ only when $\alpha = m - 1$ and $k \leq \min(\alpha, \beta)$, and its contribution has degree at most $\deg(a_j c_k - a_k c_j) + k + j - 1 - \alpha - \beta$. For $j = 4$ we have $a_4 c_k - a_k c_4 = c_k$, of degree at most one, and otherwise $j \leq 3$; in both cases the degree is at most $4 + k - \alpha - \beta \leq 4 - \max(\alpha, \beta)$. Hence

$$\deg \mathcal{B}_{\alpha\beta} \leq 4 - \max(\alpha, \beta), \quad (12)$$

and each term $\prod_{\alpha} \mathcal{B}_{\alpha\sigma(\alpha)}$ of the determinant has degree at most $\sum_{\alpha} (4 - \max(\alpha, \sigma(\alpha))) \leq 10$, with equality possible only for the identity permutation. Note that $\gamma_1 < 0 < \gamma_2$, $\gamma_3 = \chi - 1 < -1$ and $\gamma_0, \gamma_4 > 0$. Since $s_0 = 0$ and $\sum_j \gamma_j s_j = 0$, the quadratic part of f is

$$f_2 = \gamma_1 s_1^2 + \gamma_2 s_2^2 + \gamma_3 s_3^2 + (\gamma_1 s_1 + \gamma_2 s_2 + \gamma_3 s_3)^2,$$

whose Hessian determinant is $8\gamma_1\gamma_2\gamma_3(1 + \gamma_1 + \gamma_2 + \gamma_3) = -8\gamma_0\gamma_1\gamma_2\gamma_3 \neq 0$. A ternary quadratic form of rank three is not a product of linear forms, so f_2 , and hence f , is irreducible. The degree-twelve part of the second equation is $-ac^4$ with $c = \sum_j \gamma_j s_j^3$. If f divided the second equation, the irreducible form f_2 would divide c^4 , hence c . Here $s_4 = -(\gamma_1 s_1 + \gamma_2 s_2 + \gamma_3 s_3)$. Divide c by f_2 as polynomials in s_3 ; the s_3^2 coefficient of f_2 is $\gamma_3(1 + \gamma_3) \neq 0$. The remainder is $\mathbf{r}_1 s_3 + \mathbf{r}_0$, with $\mathbf{r}_0, \mathbf{r}_1$ forms in s_1, s_2 and

$$\mathbf{r}_1 = \frac{\gamma_3 - 1}{\gamma_3 + 1} (\gamma_1(\gamma_1 + \gamma_3 + 1)s_1^2 + 2\gamma_1\gamma_2 s_1 s_2 + \gamma_2(\gamma_2 + \gamma_3 + 1)s_2^2),$$

as one checks by multiplying out the quotient $-((\gamma_3 + 2)(\gamma_1 s_1 + \gamma_2 s_2) + (\gamma_3^2 - 1)s_3)/(\gamma_3 + 1)$. The $s_1 s_2$ coefficient $2\gamma_1\gamma_2(\gamma_3 - 1)/(\gamma_3 + 1)$ of $\mathbf{r}_1$ is nonzero by the signs above, so $f_2 \nmid c$. So the two equations have no common factor, their common zero set is a curve, and by Bézout's theorem its components over $\mathbb{Q}$ have total degree at most $2 \cdot 12$. $\square$

The following example shows why the fixed-dividend equation is needed. If $(\rho, \varphi, \chi) = (\rho', \varphi', \chi') = (\vartheta, 2\vartheta + 1, -\vartheta - 2)$ with $\vartheta > 0$, then $\sum_l \gamma_l z^l = (z - 1)^3(z - \vartheta)$, and constant increments $k_j = t$ satisfy the linear and quadratic relations for every t : since $z = 1$ is a triple root, $\sum_l \gamma_l l^r = 0$ for $r \leq 2$, and with $s_j = jt$ these sums are exactly the k - and r -relations.

8.2 Reconstruction and retained components

We need q_i and $I_3(i)$ as rational functions on each component, with computable orders at infinity. This works only on components where two further identities hold identically; on any other component they cut out boundedly many good starts, which we count directly. The idea is that the polynomials P_4 and Z , both orthogonal to $1, x, x^2, x^3$ in $L^2(\nu_i)$, determine the moment vector of ν_i up to scale, and the two normalised division equations fix the scale; this recovers q_i and

$I_3(i)$ as rational functions of the free nodes. Concretely, form the 8×9 matrix of coefficients of $P_4, xP_4, x^2P_4, x^3P_4, Z, xZ, x^2Z, x^3Z$ in the basis $1, x, \dots, x^8$. Let c_j be $(-1)^j$ times the minor obtained by deleting column j , and define

$$Q = \frac{c_1 - c_2}{c_1 - c_0}, \quad E_0 = a \frac{\det(c_{r+s})_{0 \leq r, s < 3} (c_1 - c_0)^3}{(c_1^2 - c_0 c_2)^3}.$$

At a good start the matrix annihilates the moment vector $(h_i(0), \dots, h_i(8))$ and has rank eight. Indeed, the resultant in (9) is positive, so P_4 and Z are coprime. So the kernel is spanned by (c_j) , and $h_i(j) = \lambda c_j$ for some λ . The normalised division equations $q_i h_i(0) + h_i(1) = q_i h_i(1) + h_i(2) = 1$ then determine q_i and λ , and $h_i(0)h_i(2) > h_i(1)^2$ shows that the denominators below are nonzero. This gives

$$h_i(j) = \frac{c_j(c_1 - c_0)}{c_1^2 - c_0 c_2} \quad (0 \leq j \leq 8), \quad Q = q_i, \quad E_0 = I_3(i).$$

In particular, all displayed denominators are nonzero at good starts.

To state the identities as equations on the curve, use the raw-moment polynomials

$$M_0 = 1, \quad M_{r+1}(k_0, k_1, \dots) = \sum_{t=0}^r \binom{r}{t} k_0^{r-t} M_t(k_1, k_2, \dots).$$

Only $M_0, \dots, M_6$ are needed; they involve $k_0, \dots, k_4$. At good starts, $M_r = \mu_i(x^r)$ for $0 \leq r \leq 6$, by the weight recursion in Section 6. Set

$$H_3 = \det(M_{r+s})_{0 \leq r, s < 3}, \quad H_4 = \det(M_{r+s})_{0 \leq r, s < 4}, \quad H_{013} = \det(M_{r+s})_{\substack{r=0,1,2 \\ s=0,1,3}}.$$

This H_4 is the polynomial of Lemma 7.3. With $p_3 = \chi - 1 - s_1 - s_2 - s_3$, the coefficient of x^3 in P_4 , consider the *normalisation* and *resolvent identities*

$$N_0 P_4(-Q) = a H_4, \quad E_0 P_4(-Q) = a((Q - p_3)H_3 - H_{013}). \quad (13)$$

The first is Lemma 7.3 at good starts, where $Q = q_i$ and $P_4 = \pi_i$; the second is the following analogue for I_3 .

Lemma 8.2 (The resolvent identity). *At a good start, $I_3(i) P_4(-q_i) = a((q_i - p_3)H_3 - H_{013})$.*

Proof. This is the bordering argument of Lemma 7.3, one size smaller. Write $h_j = h_i(j)$, $M_j = \mu_i(x^j) = q_i h_j + h_{j+1}$ and $P_4(x) = x^4 + p_3 x^3 + p_2 x^2 + p_1 x + p_0$. Since P_4 is orthogonal to $1, x, x^2$, expanding along the last row gives

$$\det(h_{r+s})_{r, s < 3} P_4(x) = \det \begin{pmatrix} h_j & h_{j+1} & h_{j+2} & h_{j+4} + p_3 h_{j+3} \\ 1 & x & x^2 & x^4 + p_3 x^3 \end{pmatrix}_{j=0,1,2}.$$

At $x = -q_i$, add q_i times column 1 to column 2, then q_i times column 0 to column 1; the first three columns of row j become h_j, M_j, M_{j+1} , and the last row becomes $(1, 0, 0, q_i^4 - p_3 q_i^3)$. Using $h_{s+1} = M_s - q_i h_s$ repeatedly, the last column equals $M_{j+3} + (p_3 - q_i)M_{j+2}$ plus a combination of the columns M_j, M_{j+1} and $q_i^3(q_i - p_3)$ times the column $(h_j; 1)$; removing these leaves $(M_{j+3} + (p_3 - q_i)M_{j+2}; 0)$. Expanding along the last row gives

$$\det(h_{r+s})_{r, s < 3} P_4(-q_i) = (q_i - p_3)H_3 - H_{013},$$

and $a \det(h_{r+s})_{r, s < 3} = I_3(i)$. □

So both identities (13) hold at good starts. Clearing denominators turns them into polynomial equations $\mathcal{C}_N = 0$ and $\mathcal{C}_R = 0$ in the free nodes, the *cuts* (written out in the proof below). A component containing a good start is *retained* if both cuts vanish identically on it.

Lemma 8.3 (Removing components by polynomial equations). *The cuts have degree at most an absolute constant C_{cut} in the free nodes. On a degree- d component, each cut either vanishes identically or vanishes at no more than $C_{\text{cut}}d$ good starts. Consequently each component that is not retained contributes $O(1)$ good starts.*

Proof. Each entry of the 8×9 coefficient matrix has degree at most four in the nodes, so the cofactors c_j have bounded degree. Write $Q = Q_n/Q_d$ and $E_0 = E_n/E_d$ using the formulas above, and clear denominators:

$$\begin{aligned}\mathcal{C}_N &= N_0 \hat{P} - a H_4 Q_d^4, \\ \mathcal{C}_R &= E_n \hat{P} - a((Q_n - p_3 Q_d) H_3 - H_{013} Q_d) Q_d^3 E_d,\end{aligned}$$

where $\hat{P} = Q_d^4 P_4(-Q_n/Q_d)$. Every factor here is a fixed polynomial expression in the nodes and the pattern coordinates, so both have bounded degree in the nodes.

If a polynomial of degree r does not vanish identically on an integral curve of degree d , its zero set on that curve has at most rd distinct rational points, by Bézout's theorem. Every good start satisfies both cleared equations, and distinct good starts of this pattern give distinct free-node triples (Section 8.1). Applying the point bound to whichever equation is not an identity proves the claim; the bound $d \leq 24$ makes it uniform. $\square$

8.3 Directions at infinity

At a point at infinity the leading nodes satisfy $\sum_j \gamma_j z_j^r = 0$ for $r \leq 3$, and if some of them coincide, the γ_j in each group of equal leading nodes sum to zero (Lemma 8.4). Excluding vanishing pair sums therefore makes the five leading nodes distinct, which fixes the orders at infinity of everything below. Call a pattern *exceptional* if two of the five Newton coefficients γ_j sum to zero. Since $\rho > 0$, $\varphi > 0$ and $\chi < 0$, every γ_j is nonzero, and a pair sum can vanish only in one of the four cases

$$\rho + \chi = 1, \quad \rho + \chi = 0, \quad \varphi = 1, \quad \rho + \varphi = 1. \quad (14)$$

Equivalently, the integer pattern coordinates satisfy $N_1 + u_2 = N_0$, $N_1 + u_2 = 0$, $u_1 = N_0$ or $N_1 + u_1 = N_0$.

Lemma 8.4 (Leading nodes and determinant orders). *Let C be an integral component of positive degree of (9). Let P be a closed point above infinity on the normalisation of the projective closure of C , write $\text{ord} = \text{ord}_P$, and let $e > 0$ be the multiplicity of the hyperplane at infinity at P ; equivalently, $e = \max_{1 \leq j \leq 3} (-\text{ord } s_j)$. For a local parameter (uniformiser) ϖ at P , let $z_j \in \mathbb{Q}(P)$ be the value at P of the regular function $\varpi^e s_j$, $0 \leq j \leq 4$; we call z_j the leading node. Then the z_j are not all equal and satisfy*

$$\sum_{j=0}^4 \gamma_j z_j^r = 0 \quad (0 \leq r \leq 3). \quad (15)$$

If two of these five leading nodes coincide, the pattern is exceptional. For a nonexceptional pattern, therefore,

$$\text{ord}(s_l - s_j) = -e \quad (0 \leq j < l \leq 4), \quad (16)$$

$$\text{ord Res}(P_4, Z) = -10e, \quad \text{ord } H_3 = -3e, \quad \text{ord } H_4 = -6e. \quad (17)$$

Proof. All nonzero rational constants have order zero. By the description of e in the statement, each $\varpi^e s_j$ is regular at P , and s_4 is a linear combination of s_1, s_2, s_3 . Thus a polynomial of total node degree at most r has order at least $-re$. The values z_j lie in the residue field $\mathbb{Q}(P)$; we do not require P to be rational. We have $z_0 = 0$, and at least one of z_1, z_2, z_3 is nonzero.

The leading nodes are distinct. The definitions of the nodes give the moment equations of orders zero and one. The highest-degree part of $f = 0$ gives the equation of order two. Finally, for the equation of order three, multiply the fixed-dividend equation by ϖ^{12e} and evaluate at P : the left side vanishes because the resultant has total degree at most ten, and the right side is $a(\sum_j \gamma_j z_j^3)^4$, since the cubic part of Ψ is $\sum_j \gamma_j s_j^3$. This gives (15).

If two nodes coincided, there would be at most four distinct values among the z_j . Grouping the sums in (15) by those values and using the corresponding Vandermonde matrix shows that the sum of the γ_j in each group is zero. Since there are five indices and at least two groups, some group has at most two elements. It cannot have one, because every γ_j is nonzero; it must therefore give a vanishing pair sum. This proves the collision assertion. For a nonexceptional pattern it is impossible, so all five leading nodes are distinct, which is (16).

The determinant orders. Write $\mathcal{R} = \text{Res}(P_4, Z)$. The leading homogeneous parts needed below are

$$[\mathcal{R}]_{10} = \gamma_0 \gamma_1 \gamma_2 \gamma_3 \prod_{0 \leq j < l \leq 4} (s_l - s_j), \quad (18)$$

$$\begin{aligned} [H_3]_3 &= k_0 k_1 (k_0 + k_1), \\ [H_4]_6 &= k_0 k_1 k_2 (k_0 + k_1) (k_1 + k_2) (k_0 + k_1 + k_2). \end{aligned} \quad (19)$$

Here $[F]_r$ denotes the homogeneous part of total node degree r .

For the degree-ten formula we use the Bézout matrix of the proof of Lemma 8.1(2). By (12), only the diagonal term of its determinant can have degree ten, so $[\text{Res}(P_4, R_4)]_{10} = \prod_{\alpha < 4} [\mathcal{B}_{\alpha\alpha}]_{4-\alpha}$. By the degree count there, a contribution of degree $4 - \alpha$ to $\mathcal{B}_{\alpha\alpha}$ needs $k = \alpha$, hence $m = \alpha + 1$, and $j \in \{3, 4\}$. The coefficient of N_α in $Y_{\alpha, \alpha+1, j} = N_\alpha \prod_{l=\alpha+2}^j (y - s_l)$ is $\prod_{l=\alpha+2}^j (s_{\alpha+1} - s_l)$. For $j = 3$ the degree-two part of $a_3 c_\alpha - a_\alpha c_3$ is $s_4(-\gamma_{\alpha+1} s_{\alpha+1}) - \gamma_{\alpha+1} s_{\alpha+1}(-s_4) = 0$, so only $j = 4$ contributes, and

$$[\mathcal{B}_{\alpha\alpha}]_{4-\alpha} = [c_\alpha]_1 \prod_{l=\alpha+2}^4 (s_{\alpha+1} - s_l) = -\gamma_{\alpha+1} s_{\alpha+1} \prod_{l=\alpha+2}^4 (s_{\alpha+1} - s_l).$$

The product of these four entries is $\gamma_1 \gamma_2 \gamma_3 \prod_{0 \leq j < l \leq 4} (s_l - s_j)$, recalling $s_0 = 0$ and $\gamma_4 = 1$, and $\text{Res}(P_4, Z) = \gamma_0 \text{Res}(P_4, R_4)$ gives (18).

For the Hankel polynomials, let $\mathcal{L}^{(l)}$ be the linear functional on polynomials in x with $\mathcal{L}^{(l)}(x^r) = M_r(k_l, k_{l+1}, \dots)$. The recurrence defining the M_r says exactly that $\mathcal{L}^{(l)}(xg(x)) = \mathcal{L}^{(l+1)}(g(x + k_l))$, so, as for μ_i in Section 6, $\mathcal{L}^{(0)}(F_j F_l) = \mathcal{L}^{(l)}(\prod_{m < j} (x + s_l - s_m))$ for $j \leq l$. Since M_r has degree $r - 1$ for $r \geq 1$ and $M_0 = 1$, the coefficient of x^t contributes degree at most $j - 1$ when $t \geq 1$, and

$$\deg \mathcal{L}^{(0)}(F_j F_l) \leq \min(j, l), \quad [\mathcal{L}^{(0)}(F_j F_l)]_j = \prod_{m < j} (s_l - s_m) \quad (j \leq l). \quad (20)$$

The unitriangular change from $1, x, x^2, x^3$ to $F_0, \dots, F_3$ shows that $H_4 = \det(\mathcal{L}^{(0)}(F_j F_l))_{j, l < 4}$. By (20), only the diagonal term can have degree six, so $[H_4]_6 = \prod_{l < 4} \prod_{m < l} (s_l - s_m)$, which is the displayed product since $s_1 = k_0$, $s_2 = k_0 + k_1$ and $s_3 = k_0 + k_1 + k_2$. For H_3 the same Gram matrix

with $j, l < 3$ is explicit,

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 + s_1 & 1 + s_2 \\ 1 & 1 + s_2 & k_2 + 1 + 2s_2 - s_1 + s_2(s_2 - s_1) \end{pmatrix}, \quad \text{so} \quad H_3 = k_0 k_1 (k_0 + k_1) + k_0 k_2 - k_1^2.$$

Finally, $x^3 = F_3 + (s_1 + s_2)x^2 - s_1 s_2 x$; in the minor H_{013} the multiple of the column of x can be removed, so $H_{013} = (s_1 + s_2)H_3 + \det(\mathcal{L}^{(0)}(F_j F_l))_{j < 3, l \in \{0, 1, 3\}}$, and by (20) the last determinant has degree at most three. Hence

$$[H_{013}]_4 = (s_1 + s_2)[H_3]_3 = (2k_0 + k_1)k_0 k_1 (k_0 + k_1). \quad (21)$$

Every factor in these leading products has order $-e$ by (16). The lower-degree terms have strictly larger order, so they cannot cancel the leading terms, which proves (17). $\square$

Corollary 8.5 (Even component degrees). *Every positive-degree nonexceptional component of (9) has even degree. At each point above infinity its hyperplane multiplicity e is even, and $\text{ord } \Psi = -5e/2$; so the function $J = N_0 \Psi$ of Section 8.6 also has order $-5e/2$.*

Proof. Taking orders in the fixed-dividend equation and using Lemma 8.4 gives $4 \text{ord } \Psi = -10e$. Since orders are integers and $N_0 \neq 0$, this proves both local assertions. The degree of the curve is the sum of the multiplicities e , weighted by the residue-field degrees, so it is even. The retained identities are not needed here. $\square$

8.4 Exceptional patterns and line components

Exceptional patterns may have components of odd degree, including lines, but there are fewer of them. Lines must be excluded because the curve bound saves nothing on them. We exclude them only in the range (22) below. Outside that range, $c_L A^4 H^5 X \geq a^3$ together with $X \asymp A^4 H^6 / a^3$ and $H \asymp a / (AT)$ already gives $T^{17} \leq a^{5+o(1)}$ (first case in the proof of Corollary 8.15).

Lemma 8.6 (Exceptional patterns and lines). *There are $O(X^4 \log(2X)) = O_\eta(a^\eta X^4)$ exceptional patterns. In the range*

$$c_L A^4 H^5 X < a^3, \quad (22)$$

where c_L is a suitable absolute constant, no line component contains a good start.

Proof. Each relation in (14) becomes a vanishing maximal minor after a fixed invertible integer column change. The transformed minor vector still has height $O(X)$. For example, if the original columns are $C_0, \dots, C_4$, then

$$\det(C_1, C_2 + C_0, C_3 - C_0) = p_{123} - p_{013} - p_{012} = N_1 + u_2 - N_0.$$

The other three relations follow by the same column additions and permutations.

A fixed zero minor reduces the Plücker relations to a simple shape by Grassmann duality. Up to sign, the ten minors $p_{c_1 c_2 c_3}$ are the Plücker coordinates $p_{c_4 c_5}^\perp$, $\{c_1, \dots, c_5\} = \{0, \dots, 4\}$, of the orthogonal complement of the row space, a 2-plane in $\mathbb{Q}^5$. After permuting columns we may take the vanishing minor to be $p_{34}^\perp = 0$. The Plücker relations $p_{c_1 c_2}^\perp p_{c_3 c_4}^\perp - p_{c_1 c_3}^\perp p_{c_2 c_4}^\perp + p_{c_1 c_4}^\perp p_{c_2 c_3}^\perp = 0$ then say exactly that the integer vectors

$$\mathbf{x} = (p_{03}^\perp, p_{13}^\perp, p_{23}^\perp), \quad \mathbf{y} = (p_{04}^\perp, p_{14}^\perp, p_{24}^\perp), \quad \mathbf{z} = (p_{12}^\perp, -p_{02}^\perp, p_{01}^\perp)$$

satisfy $\mathbf{x} \times \mathbf{y} = 0$ and $\mathbf{x} \cdot \mathbf{z} = \mathbf{y} \cdot \mathbf{z} = 0$. They have height $O(X)$ and are signed rearrangements of the nine remaining minors, so they determine the minor vector. The case $\mathbf{x} = \mathbf{y} = 0$ has $O(X^3)$ choices. Otherwise write $\mathbf{x} = \lambda_1 \mathbf{v}$, $\mathbf{y} = \lambda_2 \mathbf{v}$ with $\mathbf{v}$ primitive. In a dyadic range $R \leq |\mathbf{v}|_\infty < 2R$, there are $O(R^3)$ choices for $\mathbf{v}$ and $O((X/R)^2)$ for (λ_1, λ_2) . The equation $\mathbf{v} \cdot \mathbf{z} = 0$ has $O(X^2/R)$ solutions in the box. Indeed, take a coordinate of $\mathbf{v}$ of absolute value $|\mathbf{v}|_\infty \in [R, 2R)$. Primitivity gives $|\mathbf{v}|_\infty$ admissible pairs of the other two coordinates of $\mathbf{z}$ modulo $|\mathbf{v}|_\infty$, and the remaining coordinate is then determined. Since $|\mathbf{v}|_\infty \ll X$, this gives $O(X^2/R)$ choices. Each range therefore contributes $O(X^4)$, and summing over $O(\log(2X))$ ranges proves the pattern count. Since $X \leq a^{O(1)}$, the logarithm is absorbed into a^η .

For the second claim the strategy is as follows: a line through a good start forces the pattern to be exceptional and the nodes to satisfy one of four special relations; each relation makes some increment at most $2X$, and then the label bound makes q_i too small. Now suppose that a line contained in the zero set of (9) passes through the free-node triple $\mathbf{s}$ of a good start, with direction w . Extend w to the five nodes linearly, as in Section 8.1, and write $\mathbf{s}(t) = \mathbf{s} + tw$. Both equations of (9) hold identically in t .

1. The resultant has degree at most ten in t , so the t^{12} coefficient of the second equation, $-a(\sum_j \gamma_j w_j^3)^4$, vanishes. Then $\Psi(\mathbf{s}(t))$ has degree at most two, so Ψ^4 has degree at most eight, and the t^{10} coefficient $[\mathcal{R}]_{10}(w)$ of the resultant, given by (18), vanishes too: two coordinates of w coincide.
2. The direction satisfies the four moment equations of Lemma 8.4: those of orders zero and one by the definition of the nodes, order two from the t^2 coefficient of f , and order three from step (1). The grouping argument in the proof of Lemma 8.4 then shows that the pattern is exceptional, with exactly two groups of equal coordinates, of sizes two and three. After scaling, and in the order of (14), w is one of

$$(0, 1, 1, 0, 1), \quad (0, 1, 1, 0, 0), \quad (0, 0, 1, 1, 0), \quad (0, 1, 0, 0, 1).$$

3. For each of these four directions the resultant has degree at most six in t . Here the Bézout matrix of (10) is not fine enough, because its degree bound ignores the coinciding nodes. Instead write the Bézout matrix of P_4 and R_4 in a Newton basis $\prod_{l < \alpha} (x - n_l)$ whose nodes n_0, n_1, n_2 are adapted to the direction. Along the line, its (α, β) entry then has t -degree at most $\varrho_\alpha + \kappa_\beta$, with weights summing to six, so every term of the determinant has degree at most six. Appendix E lists the nodes and weights; the bounds hold for all values of the γ_j and γ'_j . Hence Ψ^4 has degree at most six, $\Psi(\mathbf{s}(t))$ is affine in t , and its second difference $\Psi(\mathbf{s} + w) + \Psi(\mathbf{s} - w) - 2\Psi(\mathbf{s})$ vanishes.
4. In each case, imposing the corresponding relation (14) and expanding the cubic Ψ gives the polynomial identity

$$\Psi(\mathbf{s} + w) + \Psi(\mathbf{s} - w) - 2\Psi(\mathbf{s}) = E_w(\mathbf{s}) + c_w \nabla f(\mathbf{s}) \cdot w,$$

in the free nodes and the pattern parameters, with $c_w = 4, 2, 2, 4$ and respectively

$$E_w = -2\rho s_3, \quad 2(\rho + \varphi)(k_1 - 1), \quad 2(\chi - 1)(k_2 - 1), \quad 2((\varphi - \chi) - (s_4 - s_1)).$$

The quadratic part of f vanishes at w (the moment equation of order two), so $f(\mathbf{s} \pm w) = f(\mathbf{s}) \pm \nabla f(\mathbf{s}) \cdot w$, and $f = 0$ along the line gives $\nabla f(\mathbf{s}) \cdot w = 0$. Hence $E_w(\mathbf{s}) = 0$. The factors ρ , $\rho + \varphi$ and $\chi - 1$ are nonzero, so respectively

$$s_3 = 0, \quad k_1 = 1, \quad k_2 = 1, \quad s_4 - s_1 = \varphi - \chi.$$

The first case is impossible at a good start, where $s_3 \geq 3$. Next bound the polynomial $\mathcal{G}$ of Lemma 7.4. At a good start $1 \leq k_j \leq H$ and $s_j \leq jH$, so its three factors satisfy $k_0 + 1 \leq 2H$,

$$k_1 s_2 + s_2 + k_1 + k_2 + 1 \leq k_1(2H + 2) + 2H + 1 \leq 7Hk_1 \leq 7H^2,$$

and $k_2(k_1 + k_2)s_3 + g_3 \leq 6H^2k_2 + 28H^2 \leq 34H^2k_2 \leq 34H^3$, since the five groups of terms of g_3 are at most $6H^2$, $3H^2$, $2H^2$, $12H^2$ and $5H^2$ in turn (the last for $k_3^2 + 2k_3 + k_4 + 1$). Hence

$$\mathcal{G} \leq 476 H^5 k_1, \quad \mathcal{G} \leq 476 H^5 k_2.$$

In the last obstruction, $k_1 + k_2 + k_3 = \varphi - \chi = (u_1 - u_2)/N_0 \leq 2X$. Thus each surviving case gives $k_1 \leq 2X$ or $k_2 \leq 2X$. The label bound $N_0 q_i^4 \leq a\mathcal{G}$, with $N_0 \geq 1$, now gives

$$q_i^4 \leq 952 a H^5 X.$$

But $q_i \geq a/(4A)$ (Lemma 2.4), so $a^3 \leq 4^4 \cdot 952 A^4 H^5 X$. Condition (22) with $c_L = 4^4 \cdot 952$ gives the strict reverse inequality. (The formalisation expands $\mathcal{G}$ more carefully, as $\mathcal{G} \leq H^5(12k_1 + 464)$, and uses $c_L = 4^4 \cdot 488$.) $\square$

8.5 Orders of the reconstructed functions

Lemma 8.7 (Orders of the reconstructed functions). *Let C be a retained nonexceptional component of positive degree containing a good start. At every point above infinity, with hyperplane multiplicity e , one has*

$$\text{ord } Q = -\frac{3e}{2}, \quad \text{ord } E_0 = \frac{3e}{2}.$$

In particular, E_0 is a nonzero rational function vanishing at every point above infinity.

Proof. The good start guarantees that the reconstruction denominators are not identically zero on C . Thus Q and E_0 are rational functions on C , and retention makes both identities (13) identities in its function field.

The quotient order. Write $P_4(x) = x^4 + p_3x^3 + p_2x^2 + p_1x + p_0$. Its Newton expression gives $\deg p_j \leq 4 - j$, and hence $\text{ord } p_j \geq -(4 - j)e$. The first retained identity gives

$$\text{ord } P_4(-Q) = \text{ord}(aH_4/N_0) = -6e.$$

If $\text{ord } Q \geq -e$, every term of $P_4(-Q)$ would have order at least $-4e$, a contradiction. Put $r = \text{ord } Q < -e$. For $0 \leq j < 4$,

$$\text{ord}(p_j(-Q)^j) - 4r \geq (4 - j)(-e - r) > 0.$$

Thus Q^4 is the unique term of lowest order. It follows that $4r = -6e$, or $\text{ord } Q = -3e/2$.

The reconstructed curvature order. The coefficient $p_3 = \chi - 1 - s_1 - s_2 - s_3$ has order at least $-e$. By the proof of Lemma 8.4, H_{013} has degree at most four, with degree-four part (21). Consequently

$$\text{ord}(Q - p_3) = -3e/2, \quad \text{ord}((Q - p_3)H_3) = -9e/2 < -4e \leq \text{ord } H_{013}.$$

There is again a unique term of lowest order. Dividing the second retained identity by $P_4(-Q)$ therefore gives

$$\text{ord } E_0 = -9e/2 - (-6e) = 3e/2.$$

The argument uses the discrete valuation on the normalisation, so it also covers branches above singular points of the original curve. $\square$

8.6 The joint coordinates

Put $d_j = b_{i+j} - b_{i+j+1}$ and

$$D_{01} = d_0 k_1 - d_1 k_0, \quad D_{02} = d_0 k_2 - d_2 k_0, \quad D_{12} = d_1 k_2 - d_2 k_1.$$

The first of these, D_{01} , is the curvature $I_3(i)$. With $s_j = q_{i+j} - q_i$ and $r_j = k_j(s_j + s_{j+1}) + k_{j+1}$, the Hankel polynomial H_3 of Section 8.2 and its shift are

$$\begin{aligned} H_3 &= k_0 r_1 - k_1 r_0 = \sigma(k_0, k_1, k_2), \\ H'_3 &= k_1 r_2 - k_2 r_1 = \sigma(k_1, k_2, k_3), \end{aligned}$$

with σ as in Section 4. They are positive at every state at which they are defined. At a good start, $W = (D_{12}, -D_{02}, D_{01}) = \mathbf{d} \times \mathbf{k}$ satisfies $W \cdot \mathbf{k} = 0$ and $W \cdot \mathbf{r} = \det(\mathbf{d}, \mathbf{k}, \mathbf{r}) = p_{012} = N_0$; these two linear equations determine W from its last coordinate $D_{01} = I_3(i)$. Solving them with E_0 in place of $I_3(i)$, on a retained component, gives

$$E_{12} = \frac{H'_3 E_0 - N_0 k_1}{H_3}, \quad E_{02} = \frac{k_0 E_{12} + k_2 E_0}{k_1}, \quad W = (E_{12}, -E_{02}, E_0), \quad J = N_0 \Psi.$$

The determinant and cross-product identities give, identically,

$$W \cdot (k_0, k_1, k_2) = 0, \quad W \cdot (r_0, r_1, r_2) = N_0. \quad (23)$$

At a good start these rational functions take the integer values $W = (D_{12}, -D_{02}, D_{01})$. Changing the rank-five Gram basis to $1, x, x^2, x^3, P_4$ gives

$$\det(\nu_i(x^{r+s}))_{r,s < 5} = \det(\nu_i(x^{r+s}))_{r,s < 4} \nu_i(P_4^2).$$

Consequently $J = I_5(i)$ is a positive integer at good starts. The heights of these integers are controlled by

$$S = \left\lfloor \frac{c_S A^3 H^3}{a^2} \right\rfloor, \quad Y = \left\lfloor \frac{c_Y A^5 H^{10}}{a^4} \right\rfloor,$$

for suitable absolute constants c_S, c_Y .

Lemma 8.8 (Heights of the joint coordinates). *At good starts, all ten 2×2 minors $D_{jl} = d_j k_l - d_l k_j$ with $0 \leq j < l \leq 4$ satisfy $|D_{jl}| \leq S$; in particular $|W_j| \leq S$. Moreover $1 \leq J \leq Y$.*

Proof. Write $q = q_i$, and recall $q \geq a/(4A)$ (Lemma 2.4). The proof of Proposition 7.8 gives $d_j/a = k_j/q^2 + O(H^2/q^3)$ for $j \leq 4$, the error terms of Lemma 7.7 being absorbed because $H \leq 1 + 28a/(AT) \leq 29a/A \leq 116q$. In $D_{jl} = d_j k_l - d_l k_j$ the leading terms cancel, so $|D_{jl}|q^3 \ll aH^3$, and $q \geq a/(4A)$ gives $|D_{jl}| \leq S$ for suitable c_S , including the floor because the minors are integers.

For $J = I_5(i)$, argue as in the proof of Lemma 7.4 with five Newton polynomials instead of four: Hadamard's inequality and $\nu_i \leq \mu_i/q$ give $Jq^5 \leq a \prod_{j < 5} \mu_i(F_j^2)$. By the shift rule, $\mu_i(F_j^2) = \mu_{i+j}(F_j(x + s_j))$ is a polynomial of degree j in the increments, and the increment bound gives $\mu_i(F_j^2) \ll H^j$. Hence $Jq^5 \ll aH^{10}$, and $q \geq a/(4A)$ gives $J \leq Y$ for suitable c_Y . $\square$

Corollary 8.9 (Orders of the joint coordinates). *In the setting of Lemma 8.7, the joint coordinates satisfy*

$$\text{ord } W_0 = \text{ord } W_2 = 3e/2, \quad \text{ord } W_1 \geq 3e/2, \quad \text{ord } J = -5e/2. \quad (24)$$

Proof. The cubic leading part of H'_3 is $k_1 k_2 (k_1 + k_2)$. By (16), it has order $-3e$, whereas the quadratic remainder $k_1 k_3 - k_2^2$ has order at least $-2e$. Thus $\text{ord } H_3 = \text{ord } H'_3 = -3e$ and $\text{ord } k_1 = -e$. In the numerator defining E_{12} the two terms have different orders:

$$\text{ord}(H'_3 E_0) = -3e/2 < -e = \text{ord}(N_0 k_1).$$

Therefore $\text{ord } E_{12} = 3e/2$. Each term in $k_0 E_{12} + k_2 E_0$ has order $e/2$, so their sum has order at least $e/2$. Dividing by k_1 gives $\text{ord } E_{02} \geq 3e/2$. Cancellation here is harmless: equality for the first and third coordinates already gives $\min_j \text{ord } W_j = 3e/2$. The formula for $J = N_0 \Psi$ is Corollary 8.5. $\square$

8.7 Counting the nonexceptional conics

Proposition 8.10 (Conics). *A nonexceptional component of degree two contains the free-node triples of $O_\eta(a^\eta)$ good starts, with an implied constant independent of the pattern and the component. Consequently all nonexceptional conics together contribute $O_\eta(a^\eta X^5)$ good starts.*

In outline: a component that is not retained contributes $O(1)$ good starts (Lemma 8.3). A retained conic through three free-node triples of good starts lies in a plane with integer coefficients of size $a^{O(1)}$, and Corollary 8.5 leaves it a single rational branch at infinity, of multiplicity two. So it is a parabola, the image of a quadratic polynomial map $t \mapsto (\xi_0(t), \xi_1(t), \xi_2(t))/\Delta_c$ whose parameter t is an integer at every integer point. By Lemma 8.7, the reconstructed curvature E_0 is a rational function of t with a zero of order three at infinity, and it takes the integer value $I_3(i)$ at each good start. Its nonconstant denominator must then divide a fixed nonzero resultant at each such t , which leaves $a^{o(1)}$ possibilities. Appendix A gives the details, including the uniform bounds on the coefficients.

8.8 The extra degree and the powered graph

The conic count leaves retained nonexceptional components of degree at least four. Fix such a component C of degree d . The additional integer coordinates will increase its degree without increasing their heights too much. The key is one extra zero over the affine part of C .

All divisors in this argument are on the normalisation $\tilde{C}$ of the projective closure of C , over $\mathbb{Q}$. For a closed point P of $\tilde{C}$, write $\deg P = [\mathbb{Q}(P) : \mathbb{Q}]$, where $\mathbb{Q}(P)$ is the residue field; thus $\deg(\sum_P a_P [P]) = \sum_P a_P \deg P$. Let $E_\infty = \sum_{P|\infty} e_P [P]$ be the pullback of the original hyperplane at infinity. Then $\sum_{P|\infty} e_P \deg P = d$. We do not require the closed points, or the branches over a rational point of C , to be rational.

Let $\mathcal{F}$ be the pole envelope of the curvature coordinates:

$$\mathcal{F} = \sum_P \max\{0, -\text{ord}_P(W_0), -\text{ord}_P(W_1), -\text{ord}_P(W_2)\} [P],$$

with $\text{ord}_P(0) = +\infty$. The orders in (24) show that $\mathcal{F}$ is supported over the affine part of C , and hence is disjoint from E_∞ .

Lemma 8.11 (An additional affine zero). *There is a rational linear combination Φ of W_0, W_1, W_2 which is not the zero function and has a positive-order zero at a closed point P_0 of $\tilde{C}$ outside E_∞ .*

Proof. First, the W_j cannot all be scalar multiples, in $\mathbb{Q}(C)$, of a fixed nonzero vector $v \in \mathbb{Q}^3$. Such a relation, together with $W \cdot (k_0, k_1, k_2) = 0$, would put C in the plane $v_0 k_0 + v_1 k_1 + v_2 k_2 = 0$. This is a genuine plane: the first three increments are an invertible linear change of the three node coordinates. The pattern quadratic is irreducible, so its restriction to this plane is nonzero and has

degree at most two. Every curve contained in this plane section therefore has degree at most two, contrary to $d \geq 4$.

Choose the free-node triple $\mathbf{s}_* \in C(\mathbb{Q})$ of a good start. All the W_j are regular there, because their common reconstruction denominator does not vanish at a good start. Write $w_j = W_j(\mathbf{s}_*)$. The identity $W \cdot (r_0, r_1, r_2) = N_0 \neq 0$ gives $w \neq 0$. Choose l with $w_l \neq 0$. At least one function $w_l W_j - w_j W_l$ is nonzero: otherwise $W = (W_l/w_l)w$, contradicting the preceding paragraph. Take such a function for Φ . It has rational coefficients and satisfies $\Phi(\mathbf{s}_*) = 0$.

This zero survives normalisation even if $\mathbf{s}_*$ is singular. Indeed, write $\Phi = g/g_0$ near $\mathbf{s}_*$, with g not identically zero on C , $g(\mathbf{s}_*) = 0$ and $g_0(\mathbf{s}_*) \neq 0$. The finite normalisation has a closed point P_0 above $\mathbf{s}_*$. In its discrete valuation ring the numerator is in the maximal ideal and the denominator is a unit, so $\text{ord}_{P_0}(\Phi) > 0$. Since $\mathbf{s}_*$ is affine, P_0 does not lie in E_∞ . $\square$

Lemma 8.12 (The curvature pole degree). *The pole envelope satisfies*

$$\deg \mathcal{F} \geq \frac{3d}{2} + 1. \quad (25)$$

Proof. Use the function Φ from Lemma 8.11. At every point over infinity, each W_j has order at least $3e_P/2$. The same is true of a rational linear combination; cancellation can only increase its zero order. Consequently

$$(\Phi)_0 \geq \frac{3}{2}E_\infty + [P_0], \quad (\Phi)_\infty \leq \mathcal{F}.$$

The first divisor is integral because every e_P is even. The function Φ is nonzero and has a zero, so its principal divisor has equal zero and pole degrees. Hence

$$\deg \mathcal{F} \geq \deg(\Phi)_\infty = \deg(\Phi)_0 \geq \frac{3}{2} \sum_{P|\infty} e_P \deg P + \deg P_0 \geq \frac{3d}{2} + 1.$$

The last inequality uses only $\deg P_0 \geq 1$, not that P_0 is rational. $\square$

For positive integers m, n , let $\Gamma_{m,n} \subseteq \mathbb{A}^7$ be the closure of the graph obtained by adjoining W_0^m, W_1^m, W_2^m, J^n to the three original node coordinates. The graph is taken on the common domain of these rational functions. It is an integral curve, and projection to the first three coordinates is birational: on that domain the four remaining coordinates are uniquely determined. Thus $\Gamma_{m,n}$ and C have the same function field. Taking powers does not lose this property, because the original coordinates remain in the graph.

Lemma 8.13 (Degree of the powered graph). *If $\delta = \deg \Gamma_{m,n}$, then*

$$2\delta \geq (5n + 3m)d + 2m, \quad \delta \geq 10n + 7m =: \delta_{\min}. \quad (26)$$

Moreover $\delta \leq C_{\text{gr}}(m + n)d$ for an absolute constant C_{gr} . For $m = n = 1$ the lower bound reads $\delta \geq 4d + 1 \geq 17$.

Proof. Lower bound. At a point P over the old infinity hyperplane, the original node coordinates have pole orders at most e_P , while J^n has pole order $5ne_P/2$ and all W_j^m vanish. Since $n \geq 1$, the pole envelope of the seven graph coordinates at P is exactly $5ne_P/2$. Over the affine part, the node coordinates and J are regular: $J = N_0\Psi$ is polynomial in the nodes. There the envelope is exactly m times the curvature envelope. The complete envelope of the graph coordinates is therefore

$$\mathcal{F}_{m,n} = \frac{5n}{2}E_\infty + m\mathcal{F}.$$

There is a single rational linear combination ψ of the seven coordinates whose pole divisor is $\mathcal{F}_{m,n}$. To see this, fix a point P in the finite support of $\mathcal{F}_{m,n}$ and multiply all seven coordinates by a local uniformiser to the power $\mathcal{F}_{m,n}(P)$. Their values in $\mathbb{Q}(P)$ are not all zero. Cancellation of the leading pole is therefore a proper $\mathbb{Q}$ -linear condition on the seven coefficients. The finitely many such proper subspaces do not cover $\mathbb{Q}^7$. Choosing coefficients outside their union gives the asserted simultaneous absence of cancellation, also when $\mathbb{Q}(P) \neq \mathbb{Q}$.

As a linear function on $\Gamma_{m,n}$, ψ has degree at most δ . Indeed, homogenise it on the projective closure of the graph: its pencil with the homogenising coordinate consists of hyperplane sections of degree δ . Removing common zeros can only decrease the degree of the resulting map to $\mathbb{P}^1$. Birational projection identifies its function field with $\mathbb{Q}(C)$, so this degree is $\deg(\psi)_\infty = \deg \mathcal{F}_{m,n}$. It follows that

$$\delta \geq \frac{5n}{2}d + m \deg \mathcal{F} \geq \frac{5n+3m}{2}d + m.$$

This is the first inequality in (26); $d \geq 4$ gives the second. The auxiliary linear combination is used only to measure degree: no bound on its coefficients enters the integer-point box.

Upper bound. This is a routine count of polynomial degrees, given in Appendix B. $\square$

Every good start whose free-node triple lies on C gives an integer point on this graph. By Lemma 8.8, its coordinates have absolute values at most

$$\max(3H, S^m, Y^n).$$

The original node coordinates also show that distinct such node triples remain distinct after adjoining the powers.

8.9 The window bound and the final exponents

We use the coefficient-uniform affine-curve theorem of Castryck–Cluckers–Dittmann–Nguyen [3, Theorem 3], also stated as [4, Corollary 2.2]: an integral affine curve C over $\mathbb{Q}$ of degree δ in $\mathbb{A}^r$, with $r \geq 2$ fixed, has

$$N_{\text{aff}}(C, B) \ll_r \delta^3 B^{1/\delta} (\log B + \delta), \quad B \geq 1, \quad (27)$$

integer points in the box $[-B, B]^r$, with an implied constant independent of the coefficients of C . We use it in the form proved in the library cited in Section 12: “integral” means prime over $\mathbb{Q}$, and $B \geq 1$. These two points need no new input: a curve that is integral over $\mathbb{Q}$ but not geometrically integral has at most δ^2 rational points, which lie in the intersection of its conjugate components, and the range $1 \leq B < 2$ follows from $B = 2$.

Proposition 8.14 (Window bound). *Fix positive integers m, n and $\eta > 0$, and put $\mathcal{B} = \max(H, S^m, Y^n)$. For a window run in $[A, 2A]$ with $T \geq 36$ satisfying (22),*

$$T \ll_{m,n,\eta} a^\eta \left(X^5 + X^4 H^{1/2} + X^5 \mathcal{B}^{1/\delta_{\min}} \right). \quad (28)$$

Proof. At least $T/2$ starts are good, and for each pattern the free-node triples of its good starts are distinct (Section 8.1). By Lemma 8.1, these triples lie on at most twenty-four integral curves of degree at most twenty-four, and we bound the good starts on each curve, allowing overlaps. All the heights H, S, Y and X are at most a fixed power of a , so the logarithm in (27) is $O_\eta(a^\eta)$.

- *Exceptional patterns.* There are $O_\eta(a^\eta X^4)$ of them (Lemma 8.6). Under (22) no line contains a good start, so every relevant curve has degree $2 \leq d \leq 24$, and (27) in $\mathbb{A}^3$ with $B = 3H$ gives $O_\eta(a^\eta H^{1/2})$ good starts per pattern. This gives the term $X^4 H^{1/2}$.

- *Nonexceptional patterns: components that are not retained, and conics.* By Lemma 8.3 and Section 8.7, these contribute $O_\eta(a^\eta)$ per pattern, hence $O_\eta(a^\eta X^5)$ in all.
- *Remaining nonexceptional components.* Lines do not occur, by Lemma 8.6, and odd degrees do not occur, by Corollary 8.5, so these components have degree $d \geq 4$. Their good starts give distinct integer points of $\Gamma_{m,n} \subseteq \mathbb{A}^7$ in a box of side $\max(3H, S^m, Y^n) \leq 3\mathcal{B}$. By Lemma 8.13, the degree δ of $\Gamma_{m,n}$ satisfies $\delta_{\min} \leq \delta \leq 24C_{\text{gr}}(m+n)$, so (27) gives $O_{m,n,\eta}(a^\eta \mathcal{B}^{1/\delta_{\min}})$ good starts per component. This gives the term $X^5 \mathcal{B}^{1/\delta_{\min}}$.

The added coordinates W_j and J are functions of the free nodes on each component. They therefore introduce no further freely counted coordinates, and the pattern count is not enlarged. $\square$

Corollary 8.15. *For fixed positive integers m, n and every $\varepsilon > 0$,*

$$P(a, b) \ll_{m,n,\varepsilon} a^{E(m,n)+\varepsilon}, \quad E(m, n) = \max\left(\frac{15\delta_{\min} + m}{51\delta_{\min} + 3m}, \frac{15\delta_{\min} + 6n}{51\delta_{\min} + 20n}\right),$$

where $\delta_{\min} = 10n + 7m$. In particular $E(1, 1) = 128/435$ and $E(50, 51) = 2201/7480$.

Proof. We bound the number T of terms in each dyadic window, eliminating the window scale A . If $T < 36$ there is nothing to prove. If $AT > 28a$, then Proposition 4.3 gives $T^4 \ll a$. Otherwise $HAT \leq 56a$, and Proposition 5.1 gives $T^2 \leq 8\Delta A$. Substituting $H \leq 56a/(AT)$ and then $A \geq T^2/(8\Delta)$ into the definitions of X, S and Y gives

$$H \ll \frac{\Delta a}{T^3}, \quad X \ll \frac{\Delta^2 a^3}{T^{10}}, \quad S \ll \frac{a}{T^3}, \quad Y \ll \frac{\Delta^5 a^6}{T^{20}}. \quad (29)$$

Every power of Δ is $O(a^\eta)$, and we absorb such factors into $a^{O(\eta)}$.

- If (22) fails, then $c_L A^4 H^5 X \geq a^3$. With $Xa^3 \leq c_X A^4 H^6$ this gives $A^8 H^{11} \gg a^6$, and $H \leq 56a/(AT)$ turns it into $A^3 T^{11} \ll a^5$, so $T^{17} \ll \Delta^3 a^5$.
- Otherwise Proposition 8.14 applies. By (29), its first term gives $T \ll a^{O(\eta)} a^{15}/T^{50}$, so $T^{51} \ll a^{15+O(\eta)}$. Its second term gives $T \ll a^{O(\eta)} a^{12+1/2}/T^{40+3/2}$, so $T^{85} \ll a^{25+O(\eta)}$.
- For the graph term, $T \ll a^{O(\eta)} X^5 \mathcal{B}^{1/\delta_{\min}}$ gives $T^{51\delta_{\min}} \ll a^{O(\eta)} a^{15\delta_{\min}} \mathcal{B}$. Since $H \geq 1$, $\max(H, S^m) \leq \max(H, S)^m \ll a^{O(\eta)} (a/T^3)^m$, so if $\mathcal{B} = \max(H, S^m)$ then $T^{51\delta_{\min}+3m} \ll a^{15\delta_{\min}+m+O(\eta)}$. If $\mathcal{B} = Y^n$ then $T^{51\delta_{\min}+20n} \ll a^{15\delta_{\min}+6n+O(\eta)}$.

The first three exponents all equal $5/17$, both fractions in $E(m, n)$ exceed $5/17$, and $T^4 \ll a$ is stronger still. Choosing η small in terms of ε, m and n , every window has $T \ll a^{E(m,n)+\varepsilon/2}$, and summing over the $O(\log a)$ windows proves the bound. For $m = n = 1$ the two fractions are $256/870 = 128/435$ and $261/887$, and the first is larger. For $(m, n) = (50, 51)$ the larger fraction is $2201/7480$. $\square$

Proof of Theorem 1.2. Set $r = n/m$. The two fractions in $E(m, n)$ become

$$E_S(r) = \frac{150r + 106}{510r + 360}, \quad E_Y(r) = \frac{156r + 105}{530r + 357}.$$

The first decreases and the second increases on $r > 0$. Their crossing satisfies $30r^2 - 10r - 21 = 0$, so $r_* = (5 + \sqrt{655})/30$ and their common value is $\alpha_* = (1310 - \sqrt{655})/4365$. Positive rational ratios approximate r_* ; choose the fixed positive integers m, n to make $E(m, n) < \alpha_* + \varepsilon/2$, and apply Corollary 8.15 with loss $\varepsilon/2$. The powers and all constants depend only on ε , not on a, b . $\square$

Remark 8.16. In terms of the scales, at the critical point $T \approx a^{\alpha_*}$ one has $A \approx a^{2\alpha_*}$, $H \approx S \approx a^{1-3\alpha_*} \approx a^{0.117}$, $X \approx a^{3-10\alpha_*} \approx a^{0.057}$ and $Y \approx a^{6-20\alpha_*} \approx a^{0.115}$. The pattern count $X^5 \approx a^{0.287}$ alone would give the exponent $5/17$. The graph term multiplies it by $\mathcal{B}^{1/\delta_{\min}}$, and at $n/m \approx r_*$ the two box radii S^m and Y^n balance, at $\mathcal{B} \approx a^{0.117m}$. This gives the extra factor $a^{0.117/(10r_*+7)} \approx a^{0.0068}$. Scaling m and n together does not change it, since $\log \mathcal{B}$ and $\delta_{\min}$ grow in proportion.

9 Exterior lattice coordinates

We now prove Theorem 1.3. The idea is that the pivot parts of the three rows of the pattern matrix satisfy congruences modulo the pivot minor fixed by the pattern, so they lie in a lattice Λ_p . In a reduced basis the increment row has coordinates $O(H/\lambda)$, and the curvature vector, a cross product, has coordinates $O(S/\lambda^2)$, even though the drop row itself is large; here λ is the sup norm of a shortest nonzero vector of Λ_p . We keep the notation and the conventions on constants of Section 8, and write $\delta_{\min} = 10n + 7m$ as there. Only the count of good starts on retained nonexceptional components of degree at least four changes. Proposition 9.10 replaces the term $X^5 \mathcal{B}^{1/\delta_{\min}}$ of Proposition 8.14 by

$$X^{5-1/(3\delta_{\min})} H^{1/\delta_{\min}} + X^{5-2m/(3\delta_{\min})} S^{m/\delta_{\min}} + X^5 Y^{n/\delta_{\min}}.$$

In effect H is replaced by $H/X^{1/3}$ and S by $S/X^{2/3}$, while Y is unchanged. The node term $X^{5-1/(3\delta_{\min})} H^{1/\delta_{\min}}$ gives the exponent $45\delta_{\min}/(153\delta_{\min} - 1)$, which is below γ_* once $\delta_{\min} \geq 18$, so of the two fractions balanced in the proof of Theorem 1.2 only the curvature fraction improves, and the crossing moves from $\alpha_* = 0.2942513\dots$ to $\gamma_* = 0.2942294\dots$. Every pattern has $\lambda(p) \ll X^{1/3}$ (Lemma 9.2), and the weighted count of Lemma 9.6 shows that this saving is available on average: the patterns with an unusually short vector are few enough to pay for their larger boxes.

Section 9.1 attaches to each pattern p a rank-three lattice $\Lambda_p \subseteq \mathbb{Z}^3$ that contains the restrictions of the rows of $\widehat{\mathcal{M}}_i$ to three pivot columns, and gives it a reduced basis. Section 9.2 shows that few patterns have a short vector in Λ_p . Section 9.3 writes the nodes and the curvature coordinates in the reduced basis. Section 9.4 checks that this linear change of coordinates keeps the degree bounds of Lemma 8.13, and Section 9.5 sums over patterns and optimises the exponent.

9.1 The lattice of a pattern

Let $p \neq 0$ be the minor vector of a 3×5 integer matrix $\mathcal{M}$, with pivot π as in the proof of Proposition 7.9; so $|p_\pi|$ is maximal among the coordinates of p . Put $\kappa = |p_\pi| \geq 1$, let $j_1 < j_2 < j_3$ be the pivot columns and $j_4 < j_5$ the other two, and call $x_\pi = (x_{j_1}, x_{j_2}, x_{j_3})$ the *pivot part* of a row x of $\mathcal{M}$. Let $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^3$ be as in that proof: $\mathbf{u}_r$ ($\mathbf{v}_r$) is the minor with the r -th pivot column replaced by column j_4 (j_5), with the sign of the column permutation. So the coordinates of $\mathbf{u}$ and $\mathbf{v}$ are, up to sign, coordinates of p , and $p_\pi x_{j_4} = \mathbf{u} \cdot x_\pi$ and $p_\pi x_{j_5} = \mathbf{v} \cdot x_\pi$ for every row x (Lemma 9.1). Let $\Lambda_p \subseteq \mathbb{Z}^3$ be the subgroup of x with $\kappa \mid \mathbf{u} \cdot x$ and $\kappa \mid \mathbf{v} \cdot x$. It contains $\kappa \mathbb{Z}^3$. Write $|x|$ for the Euclidean norm and $|x|_\infty$ for the sup norm, and let the *minimum* $\lambda(p)$ be the least sup norm of a nonzero vector of Λ_p . Then $|x| \geq \lambda(p)$ for every nonzero $x \in \Lambda_p$.

Lemma 9.1 (Lifting). *For a column l , let $\mathbf{L}_l \in \mathbb{Z}^3$ be $p_\pi \mathbf{e}_r$ if $l = j_r$ is a pivot column, $\mathbf{u}$ if $l = j_4$ and $\mathbf{v}$ if $l = j_5$, where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is the standard basis. For every row x of $\mathcal{M}$ and every column l , $p_\pi x_l = \mathbf{L}_l \cdot x_\pi$. Moreover $\det(\mathbf{L}_0, \mathbf{L}_1, \mathbf{L}_2) = p_\pi^2 p_{012}$, and $|\mathbf{L}_l|_\infty \leq \kappa$.*

Proof. The vectors $\mathbf{L}_l$ are the rows of $\mathcal{M}^\top \text{adj}(\mathcal{M}_\pi^\top)$, where $\mathcal{M}_\pi$ is the pivot 3×3 block, and the first identity is Cramer's rule for the system $\mathcal{M}_\pi c = (\text{column } l)$, whose solution is $c = \mathbf{L}_l / p_\pi$. Taking the rows $l = 0, 1, 2$ gives $\det(\mathbf{L}_0, \mathbf{L}_1, \mathbf{L}_2) = p_{012} \det \text{adj}(\mathcal{M}_\pi^\top) = p_{012} p_\pi^2$. The bound holds because the coordinates of $\mathbf{u}$ and $\mathbf{v}$ are, up to sign, coordinates of p , and $|p_\pi|$ is maximal. $\square$

So the pivot parts x_π of the rows of $\mathcal{M}$ lie in Λ_p . They are the rows of $\mathcal{M}_\pi$, which span a sublattice of index $|\det \mathcal{M}_\pi| = \kappa$, so $[\mathbb{Z}^3 : \Lambda_p] \leq \kappa$.

Lemma 9.2 (Index). *If K is a nonnegative integer and $(K+1)^3 > \kappa$, then $\lambda(p) \leq K$.*

Proof. The box $[0, K]^3$ has more than $\kappa \geq [\mathbb{Z}^3 : \Lambda_p]$ points, so two of them lie in the same coset of Λ_p , and their difference is a nonzero vector of Λ_p with sup norm at most K . $\square$

The next lemma is the classical reduction of a three-dimensional lattice: any basis with $|\mathbf{f}_1||\mathbf{f}_2||\mathbf{f}_3| \ll \det \Lambda$ would do, and the constant does not affect the exponents. For completeness we give the elementary proof in Appendix C.

Lemma 9.3 (Reduction). *Let $\Lambda \subseteq \mathbb{Z}^3$ be a subgroup containing $\kappa \mathbb{Z}^3$, $\kappa \geq 1$. There is a basis $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ of Λ with $\det \Lambda = \det(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$, $0 < \det \Lambda \leq \kappa^3$ and*

$$27|\mathbf{f}_1|^2|\mathbf{f}_2|^2|\mathbf{f}_3|^2 \leq 64(\det \Lambda)^2.$$

Hence every $x \in \Lambda$ is $\sum \xi_i \mathbf{f}_i$ with $\xi \in \mathbb{Z}^3$ and $27|\mathbf{f}_i|^2 \xi_i^2 \leq 64|x|^2$.

9.2 Few patterns have a short vector

Lemma 9.4 (Divisibility count). *For $\ell \geq 1$ and $x \in \mathbb{Z}^3$, at most $27\ell^2 \gcd(\ell, x)$ vectors $y \in [-\ell, \ell]^3$ have $\ell \mid y \cdot x$. Here $\gcd(\ell, x) = \gcd(\ell, x_1, x_2, x_3)$.*

Proof. Each residue class modulo ℓ meets $[-\ell, \ell]^3$ in at most 3^3 points, so it suffices to show that the set $\mathcal{K}$ of solutions in $[0, \ell)^3$ has at most $\ell^2 e$ elements, where $e = \gcd(\ell, x)$. By Bézout choose $z \in \mathbb{Z}^3$ with $z \cdot x \equiv e \pmod{\ell}$. The map $(y, j) \mapsto (y + jz) \bmod \ell$ is injective on $\mathcal{K} \times [0, \ell/e)$: equal images give $\ell \mid (j - j')e$, so $j = j'$, and then $y = y'$. Hence $|\mathcal{K}| \ell/e \leq \ell^3$. $\square$

Proposition 9.5 (Short vectors). *Let $\mathcal{S}$ be a set of nonzero minor vectors with coordinates in $[-X, X]$. Then for every $R \geq 1$,*

$$\#\{p \in \mathcal{S} : \lambda(p) \leq R\} \ll R^3 X^4 \tau_{\max}(X)^2.$$

Proof. If $\lambda(p) \leq R$, some nonzero $x \in [-R, R]^3$ lies in Λ_p , so $\kappa \mid \mathbf{u} \cdot x$, where $\kappa = |p_\pi| \leq X$; we drop the condition $\kappa \mid \mathbf{v} \cdot x$. For fixed x , the encoding $p \mapsto (\pi, \text{sgn } p_\pi, \kappa, \mathbf{u}, \mathbf{v})$ of the proof of Proposition 7.9 gives at most $20 \sum_{\kappa \leq X} N(\kappa, x)$ such p , where $N(\kappa, x)$ counts pairs $(\mathbf{u}, \mathbf{v})$ in $[-\kappa, \kappa]^3$ with $\kappa \mid \mathbf{u} \times \mathbf{v}$ and $\kappa \mid \mathbf{u} \cdot x$. By the proof of Proposition 7.9, for each $\mathbf{u}$ there are at most $27\kappa \gcd(\kappa, \mathbf{u})^2$ vectors $\mathbf{v}$. Group the $\mathbf{u}$ by $g = \gcd(\kappa, \mathbf{u})$ and write $\mathbf{u} = g\mathbf{u}_0$ with $(\kappa/g) \mid \mathbf{u}_0 \cdot x$. Lemma 9.4 with $\ell = \kappa/g$, and $\gcd(\kappa/g, x) \leq \gcd(\kappa, x)$, give

$$\sum_{\mathbf{u}} g^2 \leq \sum_{g \mid \kappa} g^2 \cdot 27 \left(\frac{\kappa}{g} \right)^2 \gcd(\kappa, x) = 27\kappa^2 \tau(\kappa) \gcd(\kappa, x),$$

so $N(\kappa, x) \leq 729\kappa^3 \tau(\kappa) \gcd(\kappa, x)$. Next, $\gcd(\kappa, x) \leq \sum_{e \mid \kappa, e \mid x} e$, and swapping the sums,

$$\sum_{0 < |x|_\infty \leq R} \gcd(\kappa, x) \leq \sum_{e \mid \kappa} e \#\{x \neq 0 : |x|_\infty \leq R, e \mid x\} \leq 27R^3 \tau(\kappa).$$

Indeed, a divisor $e > R$ of κ divides no nonzero x in the box, and for $e \leq R$ the multiples of e number at most $(2R/e + 1)^3$, with $e(2R/e + 1)^3 \leq 27R^3$. Sum over x and then over $\kappa \leq X$, using $\sum_{\kappa \leq X} \kappa^3 \leq X^4$. $\square$

Lemma 9.6 (Fractional weights). *Let $\mathcal{S}$ be a set of nonzero minor vectors with coordinates in $[-X, X]$, $X \geq 1$. Then for $0 \leq t \leq 1$,*

$$\sum_{p \in \mathcal{S}} \lambda(p)^{-t} \ll \tau_{\max}(X)^2 X^{5-t/3}.$$

Proof. Choose an integer K with $X \leq K^3 \leq 8X$. By Lemma 9.2, $1 \leq \lambda(p) \leq K$ for $p \in \mathcal{S}$. Writing $1/\lambda = 1/K + \sum_{\lambda \leq j < K} 1/(j(j+1))$ and using Proposition 9.5 with $R = j$,

$$K \sum_{p \in \mathcal{S}} \frac{1}{\lambda(p)} \leq \#\mathcal{S} + O\left(KX^4 \tau_{\max}(X)^2 \sum_{1 \leq j < K} \frac{j^2}{j+1}\right) \leq \#\mathcal{S} + O(X^4 \tau_{\max}(X)^2 K^3).$$

The proof of Proposition 7.9 gives $\#\mathcal{S} \ll \tau_{\max}(X)X^5$. Since $K^3 \leq 8X$ and $\tau_{\max}(X) \geq 1$, the right side is $O(\tau_{\max}(X)^2 X^5)$. Dividing by $K \geq X^{1/3}$ proves the case $t = 1$. For $0 < t < 1$, Hölder's inequality with exponents $1/(1-t)$ and $1/t$ gives $\sum \lambda^{-t} \leq (\#\mathcal{S})^{1-t} (\sum \lambda^{-1})^t$, which is the claim; the cases $t = 0$ and $t = 1$ are the two bounds above. $\square$

9.3 Exterior coordinates

Lemma 9.7 (Exterior coordinates). *Let $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ be a basis of $\Lambda \subseteq \mathbb{Z}^3$ as in Lemma 9.3, with $|\mathbf{f}_i| \geq \lambda > 0$. Let $x, x' \in \Lambda$ have coordinates $\xi, \xi' \in \mathbb{Z}^3$, and put $w = x \times x'$ and $c = \xi \times \xi' \in \mathbb{Z}^3$. Then*

$$w = c_1 \mathbf{f}_2 \times \mathbf{f}_3 + c_2 \mathbf{f}_3 \times \mathbf{f}_1 + c_3 \mathbf{f}_1 \times \mathbf{f}_2, \quad \mathbf{f}_i \cdot w = (\det \Lambda) c_i, \quad 27\lambda^4 c_i^2 \leq 64|w|^2.$$

Proof. The first identity follows by expanding $(\sum \xi_i \mathbf{f}_i) \times (\sum \xi'_i \mathbf{f}_i)$ bilinearly, and taking the dot product with $\mathbf{f}_i$ gives the second. By Cauchy–Schwarz and Lemma 9.3,

$$27|\mathbf{f}_1|^2 |\mathbf{f}_2|^2 |\mathbf{f}_3|^2 c_i^2 \leq 64(\det \Lambda)^2 c_i^2 = 64(\mathbf{f}_i \cdot w)^2 \leq 64|\mathbf{f}_i|^2 |w|^2.$$

Cancelling $|\mathbf{f}_i|^2$ and using $|\mathbf{f}_j| \geq \lambda$ for the other two vectors gives the claim. $\square$

The lemma replaces the bounds $|\xi| \ll |x|/\lambda$ and $|\xi'| \ll |x'|/\lambda$, which would give $|c| \ll |x||x'|/\lambda^2$, by $|c| \ll |x \times x'|/\lambda^2$. This matters because the drop row is large while the curvature minors are small.

Now fix a pattern p that occurs at a good start, and let $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ be a basis of Λ_p as in Lemma 9.3. Write $\lambda = \lambda(p)$. At a good start i with pattern p , the pivot parts d_π and k_π of the drop and increment rows of $\widehat{\mathcal{M}}_i$ lie in Λ_p , with coordinates ξ and ξ' . Put

$$W' = \xi \times \xi', \quad w = d_\pi \times k_\pi.$$

The entries of w are, up to sign, the minors D_{jl} with j, l pivot columns. The next lemma shows that $W' = \Upsilon_p W$ for a constant matrix Υ_p . We use W' rather than $\Upsilon_p W$ because W' is visibly integral, while Υ_p has rational entries.

Lemma 9.8 (Coordinates in the reduced basis). *There are matrices $\mathcal{T}_p, \Upsilon_p \in \text{GL}_3(\mathbb{Q})$, depending only on p , such that at every good start with pattern p*

$$(s_1, s_2, s_3) = \mathcal{T}_p \xi', \quad W' = \Upsilon_p W, \quad |\xi'|_\infty \leq \frac{3H}{\lambda}, \quad |W'|_\infty \leq \frac{8S}{3\lambda^2}.$$

Proof. By Lemma 9.1, $p_\pi(k_0, k_1, k_2) = \mathcal{L}k_\pi$, where $\mathcal{L}$ has rows $\mathbf{L}_0, \mathbf{L}_1, \mathbf{L}_2$ and $\det \mathcal{L} = p_\pi^2 p_{012} \neq 0$, since $p_{012} = N_0 = I_4(i) \geq 1$ at a good start. With $k_\pi = (\mathbf{f}_1 \mathbf{f}_2 \mathbf{f}_3)\xi'$ and the cumulative sums $s_j = k_0 + \dots + k_{j-1}$, this gives $\mathcal{T}_p$. Since $|k_\pi|^2 \leq 3H^2$ and $|\mathbf{f}_i| \geq \lambda$, Lemma 9.3 gives $|\xi'_i| \leq (8/3)H/\lambda \leq 3H/\lambda$.

For the curvature, $W = (D_{12}, -D_{02}, D_{01}) = (d_0, d_1, d_2) \times (k_0, k_1, k_2)$, and $p_\pi(d_0, d_1, d_2) = \mathcal{L}d_\pi$. Let $\text{Cof}(\mathcal{L})$ have rows $\mathbf{L}_1 \times \mathbf{L}_2, \mathbf{L}_2 \times \mathbf{L}_0, \mathbf{L}_0 \times \mathbf{L}_1$; its determinant is $(\det \mathcal{L})^2 \neq 0$. The Binet–Cauchy identity $(\mathbf{L} \cdot x)(\mathbf{L}' \cdot y) - (\mathbf{L}' \cdot x)(\mathbf{L} \cdot y) = (\mathbf{L} \times \mathbf{L}') \cdot (x \times y)$, applied to pairs of rows of $\mathcal{L}$, gives $(\mathcal{L}x) \times (\mathcal{L}y) = \text{Cof}(\mathcal{L})(x \times y)$, so $p_\pi^2 W = \text{Cof}(\mathcal{L})w$. Lemma 9.7 gives $(\det \Lambda_p)W' = (\mathbf{f}_1 \mathbf{f}_2 \mathbf{f}_3)^\top w$. Together, $W' = p_\pi^2 (\det \Lambda_p)^{-1} (\mathbf{f}_1 \mathbf{f}_2 \mathbf{f}_3)^\top \text{Cof}(\mathcal{L})^{-1} W$, which defines Υ_p . Finally $|w|^2 \leq 3S^2$ by Lemma 8.8, and Lemma 9.7 gives $3\lambda^2 |W'_i| \leq 8S$. $\square$

9.4 The lattice graph

On a retained nonexceptional component C of degree $d \geq 4$ of a pattern p , regard $W' = \Upsilon_p W$ as a vector of rational functions on C , and let $\Gamma'_{m,n}$ be the image of the graph of

$$(s_1, s_2, s_3, (W'_0)^m, (W'_1)^m, (W'_2)^m, J^m)$$

under the linear map that replaces the nodes s by $\xi' = \mathcal{T}_p^{-1}s$.

Lemma 9.9 (Degree of the lattice graph). *The degree δ' of $\Gamma'_{m,n}$ satisfies $\delta_{\min} \leq \delta' \leq C_{\text{gr}}(m+n)d$.*

Proof. An invertible linear change of the node coordinates is a linear automorphism of $\mathbb{A}^7$, so it preserves the affine Hilbert function, and we may keep the nodes s . The proof of Lemma 8.13, with Appendix B, uses the curvature coordinates only through the following facts.

1. At every point P over infinity, $\text{ord}_P W_j \geq 3e_P/2 > 0$ for each j (Corollary 8.9).
2. The pole envelope $\mathcal{F}$ is supported over the affine part of C , and $\deg \mathcal{F} \geq 3d/2 + 1$ (Lemma 8.12).
3. Some rational linear combination of the seven graph coordinates has pole divisor $\frac{5n}{2}E_\infty + m\mathcal{F}$; its construction uses only this envelope.
4. The W_j have a common denominator Ω and numerators of degree at most C_W (Appendix B).

Each of these is unchanged when W is replaced by $W' = \Upsilon_p W$ for a constant invertible matrix Υ_p . Each W'_j is a constant combination of the W_l , and each W_l is a constant combination of the W'_j , so for every point P

$$\min_j \text{ord}_P W'_j = \min_l \text{ord}_P W_l.$$

This gives (1) for W' , and it shows that the pole envelope of W' is $\mathcal{F}$, which gives (2) and (3). The numerators of W' over Ω are constant combinations of those of W , which gives (4). The proof of Lemma 8.13 therefore applies unchanged. $\square$

9.5 The window bound and the optimised exponent

Proposition 9.10 (Lattice window bound). *Fix positive integers m, n and $\eta > 0$. For a window run in $[A, 2A]$ with $T \geq 36$ satisfying (22),*

$$T \ll_{m,n,\eta} a^\eta \left(X^5 + X^4 H^{1/2} + X^{5-1/(3\delta_{\min})} H^{1/\delta_{\min}} + X^{5-2m/(3\delta_{\min})} S^{m/\delta_{\min}} + X^5 Y^{n/\delta_{\min}} \right).$$

Proof. The proof of Proposition 8.14 applies, except for the remaining components of the nonexceptional patterns. For such a component of a pattern p , the good starts give integer points $(\xi', (W'_0)^m, (W'_1)^m, (W'_2)^m, J^n)$ of $\Gamma'_{m,n}$, and distinct node triples give distinct points because ξ' determines s . By Lemma 9.8 they lie in the box of side

$$\mathcal{B}_p = \max\left(1, \frac{3H}{\lambda(p)}, \left(\frac{8S}{3\lambda(p)^2}\right)^m, Y^n\right).$$

By Lemma 9.9 and (27), there are $O_{m,n,\eta}(a^\eta \mathcal{B}_p^{1/\delta_{\min}})$ of them, and

$$\mathcal{B}_p^{1/\delta_{\min}} \leq 1 + 3H^{1/\delta_{\min}} \lambda(p)^{-1/\delta_{\min}} + 8S^{m/\delta_{\min}} \lambda(p)^{-2m/\delta_{\min}} + Y^{n/\delta_{\min}}.$$

Sum over the $O_\eta(a^\eta X^5)$ patterns. Since $\delta_{\min} \geq 2m$, both $1/\delta_{\min}$ and $2m/\delta_{\min}$ lie in $[0, 1]$, and Lemma 9.6, with $\tau_{\max}(X) = a^{o(1)}$, gives $\sum_p \lambda(p)^{-1/\delta_{\min}} \ll a^\eta X^{5-1/(3\delta_{\min})}$ and $\sum_p \lambda(p)^{-2m/\delta_{\min}} \ll a^\eta X^{5-2m/(3\delta_{\min})}$. $\square$

Corollary 9.11. *For fixed positive integers m, n and every $\varepsilon > 0$,*

$$P(a, b) \ll_{m,n,\varepsilon} a^{E'(m,n)+\varepsilon},$$

$$E'(m, n) = \max\left(\frac{5}{17}, \frac{45\delta_{\min}}{153\delta_{\min} - 1}, \frac{45\delta_{\min} - 3m}{153\delta_{\min} - 11m}, \frac{15\delta_{\min} + 6n}{51\delta_{\min} + 20n}\right),$$

where $\delta_{\min} = 10n + 7m$. In particular $E'(3, 2) = 153/520$.

Proof. Argue as in the proof of Corollary 8.15. The first two terms and the windows in which (22) fails give the exponent $5/17$. For the third term, $T^{3\delta_{\min}} \ll a^{O(\eta)} X^{15\delta_{\min}-1} H^3$, and (29) gives $T^{153\delta_{\min}-1} \ll a^{45\delta_{\min}+O(\eta)}$. For the fourth, $T^{3\delta_{\min}} \ll a^{O(\eta)} X^{15\delta_{\min}-2m} S^{3m}$ gives $T^{153\delta_{\min}-11m} \ll a^{45\delta_{\min}-3m+O(\eta)}$. For the fifth, $T^{\delta_{\min}} \ll a^{O(\eta)} X^{5\delta_{\min}} Y^n$ gives $T^{51\delta_{\min}+20n} \ll a^{15\delta_{\min}+6n+O(\eta)}$. For $(m, n) = (3, 2)$, $\delta_{\min} = 41$ and the last three fractions are $1845/6272$, $153/520$ and $627/2131$. $\square$

Proof of Theorem 1.3. Set $r = n/m$. The last two fractions in $E'(m, n)$ become

$$E'_S(r) = \frac{450r + 312}{1530r + 1060}, \quad E'_Y(r) = \frac{156r + 105}{530r + 357},$$

where E_Y is the same function as in the proof of Theorem 1.2. The first decreases and the second increases on $r > 0$. Their crossing satisfies $15r^2 = 7$, so $r_* = \sqrt{105}/15$, and their common value is $\gamma_* = 471/1559 - 6\sqrt{105}/7795$. The fraction $45\delta_{\min}/(153\delta_{\min} - 1)$ decreases in $\delta_{\min}$, and it is below γ_* once $\delta_{\min} > \gamma_*/(153\gamma_* - 45) \approx 17.2$; also $5/17 < \gamma_*$. Choose a positive rational r close to r_* with $\max(E'_S(r), E'_Y(r)) < \gamma_* + \varepsilon/2$, and write $r = n/m$ with $\delta_{\min} = 10n + 7m \geq 18$, scaling m and n by a common factor if necessary. Then $45\delta_{\min}/(153\delta_{\min} - 1) < \gamma_*$, so $E'(m, n) < \gamma_* + \varepsilon/2$, and Corollary 9.11 with loss $\varepsilon/2$ gives the theorem. $\square$

10 Exact polar projection and the Hilbert sieve

The curvature vector W is a multiple of an affine-linear vector $G\mathbf{k} + \ell$ in the increments (Proposition 10.1). Away from the single common zero of $G\mathbf{k} + \ell$, every zero of J is therefore a pole of W , and this raises the graph degree provided that (i) little cancellation occurs at that base point and (ii) the pattern curve is irreducible, of the full degree 24. A local intersection calculation gives (i), and a sieve for branches at infinity gives (ii). We then count the branch-bad and degenerate patterns that these steps exclude, and obtain Theorem 1.1.

10.1 An exact identity at orbit states

Recall the normalised parameters (8), with $\rho, \varphi > 0 > \chi$ and $\gamma = (\rho, -\rho - \varphi, \varphi - \chi, \chi - 1, 1)$; primes indicate the next state. We retain the notation N_0, Ψ, P_4, R_4 and D_{jl} from Section 8. Write $\mathbf{k} = (k_0, k_1, k_2)^\top$ and $\mathbf{d} = (d_0, d_1, d_2)^\top$, so that $W = (D_{12}, -D_{02}, D_{01})^\top = \mathbf{d} \times \mathbf{k}$.

The increment and drop rows of the pattern matrix satisfy the same linear relations (Lemma 7.1): the cofactor relations at i and $i + 1$ give

$$k_3 = \rho k_0 - \varphi k_1 - \chi k_2, \quad k_4 = \rho' k_1 - \varphi' k_2 - \chi' k_3,$$

and the same formulas express d_3, d_4 in terms of $\mathbf{d}$. For $\mathbf{y} \in \mathbb{Q}^3$ extend $\mathbf{y}$ to $(y_0, \dots, y_4)$ by these formulas, and let $\langle \mathbf{y} \rangle_j = y_0 + \dots + y_{j-1}$ be its partial sums, so that $\langle \mathbf{k} \rangle_j = s_j$ and $\langle \mathbf{d} \rangle_j = b_i - b_{i+j}$. Define G and ℓ by

$$\sum_{j=0}^4 \gamma_j \langle \mathbf{y} \rangle_j^2 = \mathbf{y}^\top G \mathbf{y}, \quad \sum_{j=0}^4 \gamma_j y_j = \ell^\top \mathbf{y}.$$

Explicitly,

$$G = \begin{pmatrix} \rho(\rho+1) & -\rho(\varphi-1) & -\rho(\chi-1) \\ -\rho(\varphi-1) & \varphi(\varphi-1) & \varphi(\chi-1) \\ -\rho(\chi-1) & \varphi(\chi-1) & \chi(\chi-1) \end{pmatrix}, \quad \ell = \begin{pmatrix} \rho(\chi-\chi') \\ \rho' - \rho - \varphi(\chi-\chi') \\ \varphi - \varphi' - \chi(\chi-\chi') \end{pmatrix},$$

and $\det G = \rho(\rho+\varphi)(\varphi-\chi)(\chi-1) \neq 0$. The pattern quadratic is $f = \sum_j \gamma_j (s_j^2 + k_j) = \mathbf{k}^\top G \mathbf{k} + \ell^\top \mathbf{k}$. Put

$$V = G\mathbf{k} + \ell, \quad \text{so that} \quad f = \mathbf{k} \cdot V.$$

The linear form $\mathbf{y} \mapsto \mathbf{y} \cdot V$ is the polarisation of f at $\mathbf{k}$: it is $\mathbf{y}^\top G \mathbf{k} + \ell^\top \mathbf{y}$, and it recovers f at $\mathbf{y} = \mathbf{k}$. We call V the *polar vector*.

Proposition 10.1 (Polar identity). *At every state i with $i + 7 < N$,*

$$\Psi W = N_0 V, \quad JW = N_0^2 V.$$

In particular, W is a positive multiple of V .

Proof. The vector V is orthogonal to $\mathbf{k}$, since $\mathbf{k} \cdot V = f = 0$. It is also orthogonal to $\mathbf{d}$; this is where the division equations enter. By polarisation,

$$\mathbf{d} \cdot V = \sum_{j=0}^4 \gamma_j s_j \langle \mathbf{d} \rangle_j + \sum_{j=0}^4 \gamma_j d_j.$$

By the definition of γ , $\sum_j \gamma_j = 0$. The relations of Lemma 7.1 give $\sum_j \gamma_j \langle \mathbf{y} \rangle_j = \sum_l y_l (\gamma_{l+1} + \dots + \gamma_4) = 0$ for $\mathbf{y} = \mathbf{k}, \mathbf{d}$ (Lemma 7.2), that is, $\sum_j \gamma_j s_j = \sum_j \gamma_j b_{i+j} = 0$. Multiply the division equations $(q_i + s_j)b_{i+j} + b_{i+j+1} = a$ by γ_j and sum over $j \leq 4$. The terms in a and in q_i cancel, and writing $b_{i+j} = b_i - \langle \mathbf{d} \rangle_j$ and $b_{i+j+1} = b_{i+j} - d_j$ leaves $-\sum_j \gamma_j s_j \langle \mathbf{d} \rangle_j - \sum_j \gamma_j d_j = 0$. Hence $\mathbf{d} \cdot V = 0$.

Since $W \cdot (r_0, r_1, r_2) = \det(\mathbf{d}, \mathbf{k}, \mathbf{r}) = p_{012} = N_0 > 0$, $W \neq 0$, so $\mathbf{d}$ and $\mathbf{k}$ are independent and V is a multiple of $W = \mathbf{d} \times \mathbf{k}$. To find the multiple, put $r_j = k_j(s_j + s_{j+1}) + k_{j+1}$. Substituting the definitions of $s_4, k_0, \dots, k_4, r_0, r_1, r_2, G, \ell$ and Ψ and expanding gives the polynomial identity

$$V \cdot (r_0, r_1, r_2) - \Psi = -(s_4 + 1 + \chi - \chi')f$$

in s_1, s_2, s_3 and the six normalised parameters; both sides are cubic in the nodes. At an orbit state $f = 0$, so $V \cdot r = \Psi$ while $W \cdot r = N_0$. Thus $\Psi W = N_0 V$. Now use $J = N_0 \Psi$ and $\Psi > 0$. $\square$

There is one point where this fails: at the common zero of the three forms, a zero of J need not be a pole of W , and every pattern curve passes through that point. In projective coordinates $[\mathbf{k} : k_h]$, the three homogeneous forms $G\mathbf{k} + \ell k_h$ have a unique common zero, the finite point $\mathbf{k}_* = -G^{-1}\ell$. Since $V(\mathbf{k}_*) = 0$, the identities $f = \mathbf{k} \cdot V$ and $V \cdot r - \Psi = -(s_4 + 1 + \chi - \chi')f$ from the proof of Proposition 10.1 give $f(\mathbf{k}_*) = \Psi(\mathbf{k}_*) = 0$. The point also lies on the fixed-dividend curve for every a . Indeed, its nodes are

$$s_0^* = 0, \quad s_j^* = \frac{\gamma'_{j-1} - \gamma_{j-1}}{\gamma_j} \quad (1 \leq j \leq 3), \quad s_4^* = \chi' - \chi,$$

as one checks by substituting them in V . Divide $Z(x) = xR_4(x)$ by $P_4(x)$ in the Newton basis $F_0, \dots, F_4$ at $s_0, \dots, s_3$, with $F_5 = F_4(x - s_4)$. Since $s_0 = 0$, we have $xN_j = F_{j+1}$ for the Newton polynomials N_j of R_4 , so $Z = \sum_j \gamma'_j F_{j+1}$, while $xF_j = F_{j+1} + s_j F_j$. Hence, with $c_Z = \chi' - \chi - s_4$, the quotient is $x + c_Z$ and the remainder $Z - (x + c_Z)P_4$ has the Newton coefficients

$$-c_Z \gamma_0, \quad \gamma'_{j-1} - \gamma_{j-1} - \gamma_j(s_j + c_Z) \quad (1 \leq j \leq 3);$$

the coefficient of F_4 is $\gamma'_3 - \gamma_3 - s_4 - c_Z = 0$. They all vanish at $\mathbf{k}_*$, so $Z_* = xP_{4,*}$ and $N_0 \text{Res}(P_{4,*}, Z_*) - a\Psi(\mathbf{k}_*)^4 = 0$. Hence $\mathbf{k}_*$ lies on every pattern curve, and Lemma 10.2 below bounds the loss there.

We need $JW = N_0^2 V$ as an identity in the function field, not only at good starts, so we add the three cleared polar identities to the cuts $\mathcal{C}_N, \mathcal{C}_R$ of Section 8.2. They have bounded degree, so, as in Lemma 8.3, a component on which one of them does not vanish identically contains $O(1)$ good starts. Throughout this section, *retained* means that all five cuts vanish identically; on a retained component,

$$JW = N_0^2 V$$

holds in the function field.

10.2 The base point and the graph degree

For an integral component C of a pattern curve, the *base divisor* of V is

$$\sum_P \left(\min_j \text{ord}_P V_j \right) P,$$

summed over the affine points P of the normalisation of C ; its degree weights P by $[\mathbb{Q}(P) : \mathbb{Q}]$. Since $V = G(\mathbf{k} - \mathbf{k}_*)$, it is supported over $\mathbf{k}_*$, the only common zero of the homogenised forms.

Lemma 10.2 (The base divisor). *There is a fixed rational polynomial Θ in the ten minor coordinates, independent of a and not identically zero on the Plücker cone, with the following property. For every pattern with $\Theta \neq 0$ and every integral component C of the pattern curve, the base divisor of V on C has degree at most four.*

Proof. Step 1: the quartic initial form. We work on the chart where $N_0 N_1 \neq 0$, so that the normalised parameters are defined, and

$$\gamma_1 \gamma_2 \gamma_3 \cdot \det(G) \cdot G_{00} \cdot \ell_0 \neq 0.$$

Here $\det G \neq 0$ defines the base point $\mathbf{k}_*$, and $\gamma_1 \gamma_2 \gamma_3 \neq 0$ defines its nodes below. Since $\nabla f(\mathbf{k}_*) = -\ell$ and $\ell_0 \neq 0$, the quadric is smooth at $\mathbf{k}_*$. Take $\xi_1 = k_1 - k_{*,1}$ and $\xi_2 = k_2 - k_{*,2}$ as local coordinates. Since $G_{00} \neq 0$, f is a genuine quadratic in k_0 , and $\ell_0 \neq 0$ makes $\mathbf{k}_*$ a simple root of

it; so we can solve $f = 0$ for k_0 by formal power series. Let $Z_1(x; \xi_1, \xi_2)$ be the linear part of the remainder of Z modulo P_4 computed after Proposition 10.1, and let $\Psi_1(\xi_1, \xi_2)$ be the linear part of Ψ on this tangent plane. Let $F = N_0 \operatorname{Res}(P_4, Z) - a\Psi^4$ be the fixed-dividend equation, restricted to the quadric $f = 0$. It has order at least four at $\mathbf{k}_*$, with quartic initial form

$$N_0 \mathcal{R}_*(\xi_1, \xi_2) - a\Psi_1(\xi_1, \xi_2)^4, \quad \mathcal{R}_*(\xi_1, \xi_2) = \operatorname{Res}_x(P_{4,*}, Z_1). \quad (30)$$

To see that only first derivatives enter, compute the resultant as a 4×4 Bézout determinant. Each row vanishes at $\mathbf{k}_*$, so its term of degree four uses exactly the linear part of every row. Higher terms in the formal solution of $f = 0$ cannot change (30).

Step 2: the polynomial Θ . Consider the rational function of the pattern parameters

$$M = \mathcal{R}_*(1, 0)\Psi_1(0, 1)^4 - \mathcal{R}_*(0, 1)\Psi_1(1, 0)^4.$$

If $M \neq 0$ and $N_0 \neq 0$, the two axis coefficients in (30) cannot both vanish, for any a : if they did, multiplying the two equations crosswise would give $N_0 M = 0$. Now write M , the chart conditions and $N_0 N_1$ as rational functions of the ten minor coordinates, by (8), each in lowest terms, and let Θ be the product of all their numerators and denominators. Then $\Theta \neq 0$ implies that each of these functions is defined and nonzero. To see that Θ is not identically zero on the Plücker cone, take the maximal minors of

$$\begin{pmatrix} 1 & 0 & 0 & 2 & 10 \\ 0 & 1 & 0 & -5 & -22 \\ 0 & 0 & 1 & 3 & 9 \end{pmatrix},$$

which give $(\rho, \varphi, \chi, \rho', \varphi', \chi') = (2, 5, -3, 3, 6, -5)$ and $\mathbf{k}_* = (-1/7, -3/28, -1/2)$. There the chart conditions hold, and Appendix F computes the tangent frame, the quartic and the first derivatives, giving

$$M = \frac{26291112152499}{60236288} \neq 0.$$

Step 3: from the initial form to the base divisor. Let $\mathcal{O}$ be the local ring of the smooth quadric at $\mathbf{k}_*$, with F as in Step 1. Choose $g = \xi_1$ or $g = \xi_2$ so that the complementary axis in (30) has nonzero coefficient. The completion of $\mathcal{O}/(g)$ is a power-series ring in one variable, in which F has order exactly four. Consequently $\dim_{\mathbb{Q}} \mathcal{O}/(F, g) = 4$. If C passes through $\mathbf{k}_*$, its local ring $\mathcal{O}_C = \mathcal{O}/I_C$ is a quotient with $F \in I_C$, so

$$\dim_{\mathbb{Q}} \mathcal{O}_C/g\mathcal{O}_C \leq 4.$$

In particular, g is nonzero on this one-dimensional integral local ring.

Let $\overline{\mathcal{O}}_C$ be its finite normalisation. The quotient $\overline{\mathcal{O}}_C/\mathcal{O}_C$ is finite-dimensional over $\mathbb{Q}$. Apply multiplication by g to $0 \rightarrow \mathcal{O}_C \rightarrow \overline{\mathcal{O}}_C \rightarrow \overline{\mathcal{O}}_C/\mathcal{O}_C \rightarrow 0$. Multiplication is injective on the first two rings, and its kernel and cokernel on the finite-dimensional quotient have equal dimensions. The resulting exact sequence therefore gives

$$\dim_{\mathbb{Q}} \overline{\mathcal{O}}_C/g\overline{\mathcal{O}}_C = \dim_{\mathbb{Q}} \mathcal{O}_C/g\mathcal{O}_C = \sum_{P|\mathbf{k}_*} [\mathbb{Q}(P) : \mathbb{Q}] \operatorname{ord}_P g.$$

The final equality follows by factoring the ideal (g) in the semilocal Dedekind ring $\overline{\mathcal{O}}_C$. Since G is invertible, the coordinates of $V = G(\mathbf{k} - \mathbf{k}_*)$ generate the centred maximal ideal. Their minimum order at each P is at most $\operatorname{ord}_P g$. The weighted sum of these minimum orders, which is the degree of the base divisor, is therefore at most four. A component missing $\mathbf{k}_*$ has zero base divisor. $\square$

Call a pattern *degenerate* if $\Theta = 0$ at its minor vector.

Lemma 10.3 (The polar graph degree). *Let $\Theta \neq 0$, and let C be a retained nonexceptional component of degree 24. Then the powered graph $\Gamma_{m,n}$ of C , and the lattice graph $\Gamma'_{m,n}$ of Section 9, have degree at least $\delta_{\text{pol}} := 60n + 56m$.*

Proof. At each point at infinity, $\text{ord}_P J = -5e_P/2$, by Corollary 8.5, and $\sum e_P[\mathbb{Q}(P) : \mathbb{Q}] = 24$. Thus J has pole divisor, and hence zero divisor, of degree 60; all its zeros are affine. At an affine point P the coordinates of V are regular, and $W = N_0^2 V/J$ gives

$$\max_j (-\text{ord}_P W_j) = \text{ord}_P J - \min_j \text{ord}_P V_j,$$

where the subtracted term is the multiplicity of P in the base divisor. Summing over affine points, Lemma 10.2 gives an affine curvature pole envelope of degree at least $60 - 4 = 56$, in place of the bound $3 \cdot 24/2 + 1 = 37$ of Lemma 8.12. These poles are disjoint from the poles of J at infinity, so the pole formula of Lemma 8.13 gives

$$\deg \Gamma_{m,n} \geq \frac{5n}{2} \cdot 24 + 56m = 60n + 56m.$$

An invertible constant change of curvature coordinates preserves the minimum valuation, so the same bound holds for $\Gamma'_{m,n}$. $\square$

It remains to show that components of degree less than 24 occur for relatively few patterns. Since the pattern curve has total degree at most 24 (Lemma 8.1), a component of degree 24 is the whole curve: we must show that the pattern curve is usually irreducible over $\mathbb{Q}$.

10.3 An explicit label for a branch at infinity

The plan is as follows. At a branch at infinity of a nonexceptional component, the residue field contains the ratio t of two leading nodes, which is a root of a fixed sextic, and a square root β coming from the fixed-dividend equation. If the minimal polynomial of $t + c\beta$ has degree twelve for some slope c , every branch at infinity has residue degree at least twelve; since its multiplicity e is even, the component then has the maximal degree 24. The *label* below is a polynomial whose irreducibility certifies this, and which can be sieved over patterns.

Let P be a point above infinity on the normalisation of a nonexceptional component. The five leading nodes are distinct by Lemma 8.4. Scale them to $(0, 1, t, y, z)$. The moment equations (15) of orders zero and one give $\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3 + 1 = 0$ and

$$z = -\gamma_1 - \gamma_2 t - \gamma_3 y.$$

Those of orders two and three become

$$Q_2(t, y) = \gamma_1 + \gamma_2 t^2 + \gamma_3 y^2 + z^2 = 0, \quad Q_3(t, y) = \gamma_1 + \gamma_2 t^3 + \gamma_3 y^3 + z^3 = 0.$$

Define the universal sextic $E_\gamma(t)$ by the polynomial identity

$$\text{Res}_y(Q_2, Q_3) = \gamma_3^2(\gamma_3 + 1)E_\gamma(t). \quad (31)$$

The quotient on the right is a polynomial over $\mathbb{Z}[\gamma_1, \gamma_2, \gamma_3]$, of degree six in t . At an orbit pattern $\gamma_3 = \chi - 1 < -1$, so $\gamma_3^2(\gamma_3 + 1) \neq 0$. Cancelling this factor in the resultant and linear subresultant

identities gives $E_\gamma(t) = 0$ and $\mathfrak{d}(t)y + \mathfrak{n}(t) = 0$, where

$$\begin{aligned}\mathfrak{d}(t) &= (\gamma_3 - 1)(\gamma_2(\gamma_2 + \gamma_3 + 1)t^2 + 2\gamma_1\gamma_2t + \gamma_1(\gamma_1 + \gamma_3 + 1)), \\ \mathfrak{n}(t) &= \gamma_2(\gamma_2 + 1)(\gamma_2 + \gamma_3 + 1)t^3 + \gamma_1\gamma_2(3\gamma_2 + \gamma_3 + 2)t^2 \\ &\quad + \gamma_1\gamma_2(3\gamma_1 + \gamma_3 + 2)t + \gamma_1(\gamma_1 + 1)(\gamma_1 + \gamma_3 + 1).\end{aligned}$$

These are polynomial identities, so the cleared relation remains valid when a specialised denominator vanishes.

Write

$$\mathcal{V}_4 = ty(t-1)(y-1)(y-t), \quad \mathcal{V}_5 = z(z-1)(z-t)(z-y).$$

By (18), the value of the leading part $[\mathcal{R}]_{10}$ at these nodes is $\mathcal{R}_0 = \gamma_0\gamma_1\gamma_2\gamma_3\mathcal{V}_4\mathcal{V}_5$. The four moment equations give the barycentric identity

$$\mathcal{R}_0\mathcal{V}_4 = \mathcal{V}_5^4. \quad (32)$$

Indeed, write the nodes as $z_0, \dots, z_3, z_4 = z$, with $\gamma_4 = 1$. Taking $g = \prod_{l < 4, l \neq j}(x - z_l)$ in $\sum_k \gamma_k g(z_k) = 0$ (valid since $\deg g \leq 3$) gives $\gamma_j \prod_{l < 4, l \neq j}(z_j - z_l) = -\prod_{l < 4, l \neq j}(z - z_l)$, and multiplying by $z_j - z$ gives $\gamma_j \prod_{l \leq 4, l \neq j}(z_j - z_l) = \prod_{l < 4}(z - z_l) = \mathcal{V}_5$ for $0 \leq j < 4$. In the product of these four identities each pair $l < m < 4$ contributes $-(z_m - z_l)^2$, and $\prod_{j < 4}(z_j - z) = \mathcal{V}_5$, so the product is $\gamma_0\gamma_1\gamma_2\gamma_3\mathcal{V}_4^2\mathcal{V}_5 = \mathcal{V}_5^4$, which is (32). The value $\psi_P = (\Psi^2/s_1^5)(P) \in \mathbb{Q}(P)$ is defined and nonzero, because $\text{ord}_P \Psi = -5e_P/2$. Taking leading terms in the curve equation gives $N_0\mathcal{R}_0 = a\psi_P^2$. Thus the residue field $\mathbb{Q}(P)$ contains

$$\beta = \mathcal{V}_5^2/\psi_P, \quad N_0\beta^2 = a\mathcal{V}_4. \quad (33)$$

So $\mathbb{Q}(P)$ contains $\mathbb{Q}(t, \beta)$, which generically has degree twelve, not just six.

For a slope c define

$$D_\gamma(t) = -t(t-1)\mathfrak{n}(t)(\mathfrak{n}(t) + \mathfrak{d}(t))(\mathfrak{n}(t) + t\mathfrak{d}(t))$$

and the label polynomial

$$\mathfrak{L}_{a, N_0, c, \gamma}(\zeta) = \text{Res}_t(E_\gamma(t), N_0\mathfrak{d}(t)^3(\zeta - t)^2 - c^2aD_\gamma(t)). \quad (34)$$

Here the resultant means the Sylvester determinant with degree bounds 6 and 11, padding missing coefficients with zero. This convention makes (34) a polynomial family under every specialisation. The relation $\mathfrak{d}(t)y + \mathfrak{n}(t) = 0$ gives $D_\gamma(t) = \mathfrak{d}(t)^3\mathcal{V}_4$, so (33) implies

$$\mathfrak{L}_{a, N_0, c, \gamma}(t + c\beta) = 0$$

in every such residue field, including when $\mathfrak{d}(t) = 0$.

Lemma 10.4 (The degree-twelve label). *Fix $a \in \mathbb{Q}^\times$, and regard $N_0, \gamma_1, \gamma_2, \gamma_3$ as independent parameters. For some integer c with $0 \leq c \leq 144$, $\mathfrak{L}_{a, N_0, c, \gamma}$ is irreducible of degree twelve over $\mathbb{Q}(N_0, \gamma_1, \gamma_2, \gamma_3)$. For a specialised nonexceptional pattern, if one such label polynomial is irreducible of degree twelve over $\mathbb{Q}$, every integral component of its pattern curve has degree twenty-four; that is, the pattern curve is irreducible over $\mathbb{Q}$.*

Proof. At $(\gamma_1, \gamma_2, \gamma_3) = (-7, 8, -4)$, the sextic in (31) equals $24E_\star(t)$, where

$$E_\star(t) = 2400t^6 - 20160t^5 + 69776t^4 - 126448t^3 + 125636t^2 - 64680t + 13475.$$

Appendix F proves $E_\star$ irreducible modulo 71. This specialisation preserves degree six. Gauss's lemma therefore shows that E_γ is irreducible over $K = \mathbb{Q}(\gamma_1, \gamma_2, \gamma_3)$: a factorisation into two factors of positive t -degree over the parameter ring would specialise to two factors of positive degree. Choose a root t of E_γ , so $[K(t) : K] = 6$. In this field $\mathfrak{d}(t) \neq 0$ since its degree is less than six, and $y = -\mathfrak{n}(t)/\mathfrak{d}(t)$. Each factor $t, t-1, \mathfrak{n}, \mathfrak{n}+\mathfrak{d}, \mathfrak{n}+t\mathfrak{d}$ of $D_\gamma = \mathfrak{d}(t)^3\mathcal{V}_4$ is a nonzero polynomial of degree at most three, so none vanishes at t , and $\mathcal{V}_4 \neq 0$.

Adjoin the independent parameter N_0 . The element $a\mathcal{V}_4/N_0$ has odd valuation at $N_0 = 0$, so it is not a square in $K(t)(N_0)$. Thus (t, β) generates an extension of degree twelve over $K(N_0)$. Among its twelve embeddings, equality of two values $t + c\beta$ excludes at most one slope per pair, so at most $12^2 = 144$ slopes are excluded and one of $0, \dots, 144$ gives a primitive element. (Counting only unordered pairs of distinct embeddings would give the smaller range $0, \dots, 66$; we keep the cruder bound, which is the one formalised, since only the finiteness of the range matters.) Its minimal polynomial divides (34), whose degree is at most twelve; hence that label polynomial is irreducible of degree twelve. This argument holds for each fixed nonzero a ; the chosen slope may depend on a .

For an actual pattern, the residue field at any branch at infinity contains the root $t + c\beta$. Irreducibility of the specialised degree-twelve label forces $[\mathbb{Q}(P) : \mathbb{Q}] \geq 12$. Corollary 8.5 gives $e_P \geq 2$, so

$$\deg C = \sum_{P \text{ at infinity}} e_P [\mathbb{Q}(P) : \mathbb{Q}] \geq 24.$$

Lemma 8.1 gives the reverse inequality. □

For each a , fix a slope $c(a) \in \{0, \dots, 144\}$ given by Lemma 10.4. Call a pattern *branch-good* if $\mathcal{L}_{a, N_0, c(a), \gamma}$ is irreducible of degree twelve over $\mathbb{Q}$, and *branch-bad* otherwise.

10.4 The Hilbert sieve

The following monic variant of [4, Theorem 1.10], with explicit logarithmic powers, is proved along the same lines in Appendix D.

Theorem 10.5 (Monic specialisation estimate). *For fixed integers $r, d \geq 1$ there is a constant $C_{r,d}$ with the following property. Let $F \in \mathbb{Z}[u_1, \dots, u_r][\zeta]$ be monic in ζ and irreducible over $\mathbb{Q}(u_1, \dots, u_r)$. Suppose $\deg_\zeta F \leq d$, each coefficient of F has total parameter degree at most d , and every integer coefficient has absolute value at most $\mathcal{H} \geq 1$. For every integer $B \geq 1$,*

$$\#\{v \in [-B, B]^r \cap \mathbb{Z}^r : F(v, \zeta) \text{ is reducible over } \mathbb{Q}\} \leq C_{r,d} (1 + \log(\mathcal{H} + 1))^4 B^{r-1/2} (1 + \log B)^{4r+1}.$$

The label polynomial (34) depends on the pattern only through N_0 and $\gamma_1, \gamma_2, \gamma_3$, which are functions of the four minors on the first four columns. We therefore sieve over these minors, and then count completions to full patterns (Lemma 10.6). To apply the theorem with a fixed, call the four minors on the first four columns,

$$\mathbf{p}_\square = (p_{012}, p_{013}, p_{023}, p_{123}),$$

the *vertex* of a minor vector p . For $p_{012} = N_0 \neq 0$ the vertex determines the independent parameters above:

$$\gamma_1 = -\frac{p_{123} + p_{023}}{p_{012}}, \quad \gamma_2 = \frac{p_{023} + p_{013}}{p_{012}}, \quad \gamma_3 = -\frac{p_{013} + p_{012}}{p_{012}}. \quad (35)$$

This change is birational. Substitute it in (34) and clear denominators by a fixed power of p_{012} . To make the resulting degree-twelve polynomial monic without changing its generic irreducibility, use the following polynomial operation. For $F(\zeta) = \sum_{i=0}^e F_i \zeta^i$ put

$$\text{mon}_e(F)(\zeta) = \zeta^e + \sum_{i < e} F_i F_e^{e-1-i} \zeta^i.$$

When $F_e \neq 0$, this is $F_e^{e-1} F(\zeta/F_e)$. When $F_e = 0$ and $e \geq 2$, it is $\zeta^{e-1}(\zeta + F_{e-1})$, hence reducible. Thus irreducibility of a specialised monic normalisation implies both preservation of degree and irreducibility of the original label.

For each fixed positive integer a , use the slope $c(a)$ and set $e = 12$. We obtain a generically irreducible monic integer family in the vertex, with uniformly bounded degree and coefficient height at most $C(a+1)^{C'}$, for absolute constants C, C' . The same is true after dilation $\mathbf{p}_\square = g\mathbf{z}$, with height at most $C(a+1)^{C'} g^{C'}$. The slope may depend on a , but the bounds are uniform because there are only 145 choices. Counting vertices with $p_{012} = 0$ as bad, which adds the $O(B^3)$ points of that slice, Theorem 10.5 gives

$$\#\{\mathbf{z} \in [-B, B]^4 \cap \mathbb{Z}^4 : g\mathbf{z} \text{ is branch-bad}\} \ll_\eta a^\eta B^{7/2} \quad (36)$$

whenever $1 \leq g, B \leq a^7$. Any fixed power of a would do here: the range only has to contain the dilations and boxes used for orbit patterns below, which are at most the pattern height $X \ll a^3$ (by (29), or directly from $X \ll A^4 H^6 / a^3$ and $H \ll a/A$).

Lemma 10.6 (Branch-bad full patterns). *Among the minor vectors of coordinate height at most X with $p_{012} \neq 0$, the number that are branch-bad is $O_\eta(a^\eta X^{9/2})$.*

Proof. Since $p_{012} \neq 0$, the vertex is nonzero. Partition the vertices into dyadic bands with $|\mathbf{p}_\square|_\infty \asymp R$, $1 \leq R \leq X$, and choose a maximal coordinate as signed pivot κ . The other three vertex coordinates form $\mathbf{u}$; put $g = \gcd(\kappa, \mathbf{u})$. Completing the fifth column amounts to choosing $\mathbf{v} \in [-X, X]^3$ with $\kappa \mid \mathbf{u} \times \mathbf{v}$, as in the proof of Proposition 7.9. Since κ is maximal only among the first four columns, $\mathbf{v}$ ranges over the full X -box rather than $[-\kappa, \kappa]^3$, and the congruence count is

$$O(X^3 g^2 / R^2).$$

Apply (36) after dilation by g . There are $O_\eta(a^\eta (R/g)^{7/2})$ bad vertices in this common-factor class; imposing the exact gcd only reduces this count. Their total contribution is at most

$$O_\eta \left(a^\eta X^3 R^{3/2} \sum_{g \geq 1} g^{-3/2} \right).$$

The series converges, and summing the bands $R \leq X$ proves the claim. The finitely many pivot choices only change the constant. $\square$

10.5 Counting degenerate patterns

The degenerate patterns, on the hypersurface $\Theta = 0$, must be counted with a saving over the $X^{5+o(1)}$ patterns in total. The following estimate saves $X^{1/3}$.

Proposition 10.7 (Hypersurface pattern count). *Let Θ be a rational polynomial in the ten minor coordinates, with $\Theta \neq 0$ at the minor vector of some rational 3×5 matrix $\mathcal{M}_0$. For every $\eta > 0$, the number of integral points of the Plücker cone of coordinate height at most $X \geq 1$ satisfying $\Theta = 0$ is*

$$O_{\Theta, \eta}(X^{14/3+\eta}).$$

The count includes nonprimitive vectors and the zero vector.

Proof. Choose a minor of maximal absolute value $\kappa > 0$ and record its sign. The six one-column replacements form $\mathbf{u}, \mathbf{v} \in [-\kappa, \kappa]^3$, and the other coordinates are signed entries of $(\mathbf{u} \times \mathbf{v})/\kappa$. There are twenty signed pivot charts. In each chart, clearing denominators in Θ gives a nonzero polynomial $\Theta_\pi(\kappa, \mathbf{u}, \mathbf{v})$. Indeed, the product of the pivot minor and Θ at the minor vector of $\mathcal{M}_0$ is a nonzero polynomial in the matrix entries, so it is nonzero at some rational matrix; every pivot chart meets the nonvanishing locus.

Put $g = \gcd(\kappa, \mathbf{u})$, $\bar{\kappa} = \kappa/g$ and $\mathbf{u}_0 = \mathbf{u}/g$. Since $\gcd(\bar{\kappa}, \mathbf{u}_0) = 1$, the congruence $\mathbf{u}_0 \times \mathbf{v} \equiv 0 \pmod{\bar{\kappa}}$ is equivalent to

$$\mathbf{v} \in \Lambda := \mathbb{Z}\mathbf{u}_0 + \bar{\kappa}\mathbb{Z}^3.$$

This lattice contains $\bar{\kappa}\mathbb{Z}^3$, so Lemma 9.3 gives a basis $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ of Λ . We need two bounds on their lengths.

- (a) $|\mathbf{f}_i| \ll \bar{\kappa}$. The vectors $\bar{\kappa}\mathbf{e}_1, \bar{\kappa}\mathbf{e}_2, \bar{\kappa}\mathbf{e}_3$ lie in Λ and span a sublattice of full rank, so for each i one of them has a nonzero i th coordinate in the basis, and the coordinate bound of Lemma 9.3 gives $27|\mathbf{f}_i|^2 \leq 64\bar{\kappa}^2$.
- (b) $|\mathbf{f}_i||\mathbf{f}_j| \geq \bar{\kappa}\lambda^\perp(\mathbf{u}_0)$ for $i \neq j$, where $\lambda^\perp(\mathbf{u}_0)$ is the least sup norm of a nonzero vector of the normal lattice $\Lambda_{\bar{\kappa}}^\perp = \{z \in \mathbb{Z}^3 : \bar{\kappa} \mid z \cdot \mathbf{u}_0\}$. Indeed, if $x, y \in \Lambda$ are independent, then $x \times y = \bar{\kappa}z$ for a nonzero $z \in \Lambda_{\bar{\kappa}}^\perp$, so $|x||y| \geq |x \times y| \geq \bar{\kappa}\lambda^\perp(\mathbf{u}_0)$.

By Lemma 9.3, the vectors $\mathbf{v} \in \Lambda$ with $|\mathbf{v}|_\infty \leq \kappa$ have basis coordinates in a box with sides $B_i = O(\kappa/|\mathbf{f}_i|)$, and by (a) and $\bar{\kappa} \leq \kappa$ these sides are $\gg 1$. A nonzero polynomial of bounded degree vanishes at $\ll B_1 B_2 B_3 / \min_i B_i$ integer points of such a box (the polynomial grid-zero bound). Applied to a nonzero fibre $\Theta_\pi(\kappa, \mathbf{u}, \cdot)$ in the basis coordinates, this and (b) give

$$\#\{\mathbf{v} \in \Lambda \cap [-\kappa, \kappa]^3 : \Theta_\pi(\kappa, \mathbf{u}, \mathbf{v}) = 0\} \ll_\Theta \sum_{i < j} \frac{\kappa^2}{|\mathbf{f}_i||\mathbf{f}_j|} \ll_\Theta \frac{\kappa^2}{\bar{\kappa}\lambda^\perp(\mathbf{u}_0)}. \quad (37)$$

Here is the reciprocal estimate that produces the exponent 14/3:

$$\sum_{\mathbf{u}_0 \in [-\bar{\kappa}, \bar{\kappa}]^3 \cap \mathbb{Z}^3} \frac{1}{\lambda^\perp(\mathbf{u}_0)} \ll \tau(\bar{\kappa})\bar{\kappa}^{8/3}. \quad (38)$$

It holds for all $\mathbf{u}_0$, without a primitivity restriction. Pigeonholing $z \mapsto \mathbf{u}_0 \cdot z \pmod{\bar{\kappa}}$ gives $\lambda^\perp(\mathbf{u}_0) \ll \bar{\kappa}^{1/3}$. For $R \geq 1$, Lemma 9.4 and a divisor sum give

$$\begin{aligned} \#\{\mathbf{u}_0 : \lambda^\perp(\mathbf{u}_0) \leq R\} &\leq \sum_{0 < |z|_\infty \leq R} \#\{\mathbf{u}_0 \in [-\bar{\kappa}, \bar{\kappa}]^3 : \bar{\kappa} \mid \mathbf{u}_0 \cdot z\} \\ &\ll \bar{\kappa}^2 \sum_{0 < |z|_\infty \leq R} \gcd(\bar{\kappa}, z_1, z_2, z_3) \ll \tau(\bar{\kappa})\bar{\kappa}^2 R^3. \end{aligned}$$

For the last bound, expand the gcd as a divisor sum and note that a nonzero z of size at most R has no common divisor greater than R . Partial summation up to $K \asymp \bar{\kappa}^{1/3}$ now bounds the reciprocal sum by $O(\bar{\kappa}^3/K + \tau(\bar{\kappa})\bar{\kappa}^2 K^2)$, proving (38).

For fixed κ, g , equations (37) and (38) bound the nonzero fibres by $O_\Theta(\tau(\bar{\kappa})\kappa^2\bar{\kappa}^{5/3})$. Sum over $g \mid \kappa$ and $\kappa \leq X$ to obtain

$$O_\Theta(\tau_{\max}(X)^2 X^{14/3}).$$

Finally, an identically zero fibre forces one fixed nonzero coefficient polynomial $\Theta_0(\kappa, \mathbf{u})$ to vanish. For fixed g , the polynomial $\Theta_0(g\bar{\kappa}, g\mathbf{u}_0)$ is still nonzero. The four-variable grid-zero bound

supplies $O_{\Theta}((X/g)^3)$ possible pairs $(\bar{\kappa}, \mathbf{u}_0)$. Each has $O(\kappa g^2) \leq O(Xg^2)$ completions in its κ -box, so these fibres total

$$O_{\Theta} \left(X^4 \sum_{g \leq X} \frac{1}{g} \right) = O_{\Theta}(X^4 \log(2X)).$$

The divisor estimate absorbs $\tau_{\max}(X)^2$ into X^{η} . Summing the charts and adding the zero vector completes the proof. $\square$

10.6 The window estimate and the polar exponent

Retain the window notation $[A, 2A)$, $T \geq 36$, and the heights H , X , S and Y of Sections 8 and 9. Only windows satisfying (22) need the new estimate.

Proposition 10.8 (Polar window estimate). *Fix positive integers m, n and $\eta > 0$, and let $\delta_{\text{pol}} = 60n + 56m$. Every window above satisfying (22) obeys*

$$T \ll_{m,n,\eta} a^{\eta} (X^5 + X^4 H^{1/2} + X^{14/3} \max(H, S, Y)^{1/17} + X^{5-1/(3\delta_{\text{pol}})} H^{1/\delta_{\text{pol}}} + X^{5-2m/(3\delta_{\text{pol}})} S^{m/\delta_{\text{pol}}} + X^5 Y^{n/\delta_{\text{pol}}}). \quad (39)$$

Proof. The earlier exceptional patterns, conics and non-retained components contribute $X^5 + X^4 H^{1/2}$, as in Proposition 9.10. Among the remaining patterns, Lemma 10.6 and Proposition 10.7 bound those that are branch-bad or degenerate by $O_{\eta}(a^{\eta} X^{14/3})$; orbit patterns have $p_{012} = N_0 \geq 1$, as Lemma 10.6 requires. Here $X \ll a^3$, so the extra X^{η} loss is absorbed into a^{η} . For these patterns use the unpowered graph of Section 8 ($m = n = 1$, so $\delta_{\min} = 17$), whose box has size $\max(H, S, Y)$; this gives the third term of (39).

For each remaining retained component, Lemma 10.4 gives degree 24, and Lemma 10.3 gives graph degree at least δ_{pol} . Its upper degree is bounded in terms of m, n by Lemma 9.9. The coefficient-uniform curve bound and Lemma 9.8 therefore give, for a pattern with shortest lattice length λ , a contribution

$$\ll_{m,n,\eta} a^{\eta} ((H/\lambda)^{1/\delta_{\text{pol}}} + (S/\lambda^2)^{m/\delta_{\text{pol}}} + Y^{n/\delta_{\text{pol}}}).$$

Summing with Lemma 9.6, at exponents $1/\delta_{\text{pol}}$ and $2m/\delta_{\text{pol}}$, produces the last three terms. $\square$

Theorem 10.9 (Polar bound). *For every fixed pair of positive integers m, n and every $\varepsilon > 0$,*

$$P(a, b) \ll_{m,n,\varepsilon} a^{E_{\text{pol}}(m,n)+\varepsilon} \quad (0 < b < a),$$

where, with $\delta_{\text{pol}} = 60n + 56m$,

$$E_{\text{pol}}(m, n) = \max \left\{ \frac{5}{17}, \frac{45\delta_{\text{pol}}}{153\delta_{\text{pol}} - 1}, \frac{45\delta_{\text{pol}} - 3m}{153\delta_{\text{pol}} - 11m}, \frac{15\delta_{\text{pol}} + 6n}{51\delta_{\text{pol}} + 20n} \right\}.$$

For $(m, n) = (3, 2)$ this exponent is 4317/14677. Optimising over the fixed powers gives

$$P(a, b) \ll_{\varepsilon} a^{\beta_{*}+\varepsilon}, \quad \beta_{*} = \frac{41301 - 3\sqrt{82849}}{137480} = 0.2941336553 \dots$$

Proof of Theorems 10.9 and 1.1. Apply the window argument of Section 9 to (39). The three terms from the contribution $X^{14/3} \max(H, S, Y)^{1/17}$ of branch-bad and degenerate patterns give exponents 717/2440 (twice, since H and S satisfy the same bound $\ll a^{1+o(1)}/T^3$) and 732/2491, all below

5/17. The other terms give the displayed maximum. Windows with $AT > 28a$, or failing (22), are handled as in the proof of Corollary 8.15 and give at most 5/17; dyadic summation and the factor a^η absorb the remaining losses.

For $(m, n) = (3, 2)$, $\delta_{\text{pol}} = 288$ and the last three fractions are 12960/44063, 4317/14677 and 1083/3682. For optimisation put $r = n/m$. The curvature and rank-five fractions become

$$E_S^{\text{pol}}(r) = \frac{2700r + 2517}{9180r + 8557}, \quad E_Y^{\text{pol}}(r) = \frac{906r + 840}{3080r + 2856}.$$

The first decreases and the second increases. They cross at $r_* = (\sqrt{82849} - 47)/360$, with common value β_* . Approximate r_* by positive rationals and scale the corresponding numerator and denominator until $45\delta_{\text{pol}}/(153\delta_{\text{pol}} - 1) < \beta_*$. The arbitrarily small exponent loss absorbs the approximation, proving Theorem 1.1. $\square$

11 Rises in the quotient increments

Throughout this section $k_i = q_{i+1} - q_i \geq 1$, $d_i = b_i - b_{i+1} \geq 1$, and a *rise* is an index $0 \leq i \leq N-3$ with $k_i < k_{i+1}$.

Lemma 11.1 (Six quotients in arithmetic progression). *Let $q, h \geq 1$ and $x_0 > x_1 > \dots > x_6 \geq 0$ satisfy $x_j(q + jh) + x_{j+1} = a$ for $0 \leq j \leq 5$. Then $q^4 \leq 324h^3a$.*

Proof. Let $f_j = x_j - x_{j+1}$, and let $E = (f_1 - f_2)h$ and $F = (f_2 - f_3)h$ be the curvatures of the blocks starting at x_1 and x_2 (all their increments equal h). With $U = f_0 - 2f_1 + f_2$, the division equations $\mathcal{E}_j: x_j(q + jh) + x_{j+1} - a = 0$ give the exact identity

$$qhU = (3h - 1)E + F, \quad \text{namely} \quad qhU - (3h - 1)E - F = h(\mathcal{E}_0 - 3\mathcal{E}_1 + 3\mathcal{E}_2 - \mathcal{E}_3).$$

This is the case $k_0 = k_1 = k_2 = k_3 = h$ of Lemma 4.1, where $L = 2h^3U$. By Lemma 3.2, $E, F \geq 1$, so the integer U is at least 1.

Now bound E and F . Subtracting consecutive division equations gives $q_j f_j \leq hx_{j+1}$, where $q_j = q + jh$. Since $x_{j+1}q_{j+1} \leq a$, this yields $q^2 f_j \leq ha$ for $1 \leq j \leq 4$. By Lemma 3.3, $Ex_3 \leq (h+1)(f_1 + f_2 + f_3)^2 \leq 9h^2(h+1)a^2/q^4$. Also $a < x_3(q + 3h + 1)$, so $Eq^4 \leq 9h^2(h+1)(q + 3h + 1)a$, and likewise $Fq^4 \leq 9h^2(h+1)(q + 4h + 1)a$. Hence

$$q^5 h \leq q^4((3h - 1)E + F) \leq 27h^3(h+1)(q + 4h + 1)a.$$

If $q \geq h$ this gives $q^4 \leq 324h^3a$. If $q < h$ then $q^4 < h^4 \leq h^3a$, since $h \leq q_1 \leq a$. $\square$

Lemma 11.2 (Drops decrease between rises). *Whenever $b_{i+2} > 0$, $d_{i+1}k_i < d_ik_{i+1}$. In particular, if $k_{i+1} \leq k_i$ then $d_{i+1} < d_i$.*

Proof. This is $D = d_ik_{i+1} - d_{i+1}k_i \geq 1$ (Lemma 3.2). If $k_{i+1} \leq k_i$ and $d_{i+1} \geq d_i$, then $d_ik_{i+1} \leq d_{i+1}k_i$, which is a contradiction. $\square$

Lemma 11.3 (Bounded multiplicity). *Let $c_j \geq 1$ be integers, $j \in J$, such that no value occurs more than M times. Then $|J|^2 \leq 2M \sum_{j \in J} c_j$.*

Proof. Induct on $|J|$. Let c_{j_0} be the largest value. The values lie in $[1, c_{j_0}]$, so $|J| \leq Mc_{j_0}$. Removing j_0 and using $|J|^2 = (|J| - 1)^2 + 2|J| - 1$ completes the induction. $\square$

Proposition 11.4 (Rises in a window). *Let $x_0, \dots, x_{T-1}$ be a window run in $[A, 2A)$ with $T \geq 2$, and let r be the number of rises among its $T - 1$ increments. Then:*

1. $(T - 1)^2 \leq 2(r + 1)(A - 1)$;
2. $T - 1 \leq \lfloor a/A \rfloor$;
3. if $6A \leq a$, then $T - 1 = \ell + s$ with $\ell^2 \leq 8(r + 1)\lfloor a/A \rfloor$ and $s^3 \leq c_{\text{low}}A$, where $c_{\text{low}} = 324 \cdot 3^4 = 26244$.

Proof. Split the increments into *stretches*: maximal blocks containing no rise. There are at most $r + 1$ stretches.

(1) By Lemma 11.2, the drops strictly decrease along a stretch, so each value of d_j occurs at most once per stretch, hence at most $r + 1$ times. Lemma 11.3 and $\sum_j d_j = x_0 - x_{T-1} \leq A - 1$ give the claim.

(2) The increments are ≥ 1 , and they sum to $q_{T-1} - q_0 \leq \lfloor a/A \rfloor$.

(3) In $c_{\text{low}} = 324 \cdot 3^4$, the factor 324 comes from Lemma 11.1 and 3^4 from the bound $a < 3Aq$ below. Call an increment h *low* if $c_{\text{low}}h^3A^4 < a^3$. Five consecutive equal increments h inside a stretch give six quotients in arithmetic progression. Lemma 11.1 then gives $q^4 \leq 324h^3a$. On the other hand, $a < 3Aq$, which follows from $a = qx_j + x_{j+1}$, where x_j is the first term of the progression, $x_{j+1} < x_j < 2A$ and $6A \leq a$. Together these give $a^3 \leq c_{\text{low}}h^3A^4$, so h is not low. Within a stretch the increments are nonincreasing, so equal values are consecutive. Hence each low value occurs at most 4 times per stretch, and at most $4(r + 1)$ times in all. Lemma 11.3 bounds the number ℓ of low increments by $\ell^2 \leq 8(r + 1)\sum_j k_j \leq 8(r + 1)\lfloor a/A \rfloor$. Let s be the number of high increments. If $s = 0$, its bound is immediate. Otherwise let h_* be their smallest value. Then $sh_* \leq \lfloor a/A \rfloor$ and $a^3 \leq c_{\text{low}}h_*^3A^4$, so $s^3a^3 \leq c_{\text{low}}A^4(a/A)^3 = c_{\text{low}}Aa^3$. $\square$

Lemma 11.5 (Geometric sums). *Let $\alpha, \beta, s, t > 0$ and $K \geq 0$. Then*

$$\sum_{k < K} \min(\alpha 2^{-sk}, \beta 2^{tk}) \leq \left(\frac{1}{1 - 2^{-s}} + \frac{2^t}{2^t - 1} \right) \alpha^{t/(s+t)} \beta^{s/(s+t)}.$$

Proof. Let $M = \alpha^{t/(s+t)} \beta^{s/(s+t)}$. Then $(\alpha 2^{-sk})^t (\beta 2^{tk})^s = M^{s+t}$, so for each k either $\beta 2^{tk} \leq M$ or $\alpha 2^{-sk} < M$. The terms $\beta 2^{tk} \leq M$ grow with ratio 2^t , so they sum to at most $M 2^t / (2^t - 1)$. The terms $\alpha 2^{-sk} < M$ decay with ratio 2^{-s} , so they sum to at most $M / (1 - 2^{-s})$. $\square$

Proof of Theorem 1.5. The rises inside a window are rises of the orbit, so $r + 1 \leq R$. Combining the three parts of Proposition 11.4 gives, for the window $[2^k, 2^{k+1})$,

$$T_k \leq 1 + \sqrt{R} \min(\sqrt{8a} 2^{-k/2}, \sqrt{2} 2^{k/2}) + \min(a 2^{-k}, 30 \cdot 2^{k/3});$$

when $6A > a$, part (2) gives $T - 1 \leq a/A < 6 \leq 30A^{1/3}$ instead of part (3). Lemma 11.5 gives

$$\sum_k \min(\sqrt{8a} 2^{-k/2}, \sqrt{2} 2^{k/2}) \ll (\sqrt{8a}\sqrt{2})^{1/2} \ll a^{1/4}.$$

It also gives $\sum_k \min(a 2^{-k}, 30 \cdot 2^{k/3}) \ll a^{1/4}$. Since $\log_2 a + 1 \ll a^{1/4}$, summing over k gives the theorem. $\square$

Remark 11.6. By Theorem 1.5, an orbit with $N \geq a^\theta$, $\theta > 1/4$, must have $R \gg a^{2\theta-1/2}$; for θ near the exponent of Theorem 1.1 this is about $a^{0.088}$. Proposition 11.4(3) also localises the rises. There is an absolute constant $c > 0$ such that if $6A \leq a$, $A \leq ca^{3\theta}$ and a window run in $[A, 2A)$ has $T \geq a^\theta$ terms, then $s \leq T/4$, so $\ell \geq T/4$ and the number r of rises in the run satisfies $r + 1 \gg Aa^{2\theta-1}$.

12 The Lean formalisation

The source repository is <https://github.com/jasondavies/pierce-length-bound>. The proof baseline is commit 6da6511cea9f8526054dd5ad60d77164f45de52e, which is registered in the Palomar registry as [PALOMAR-2026-10-27-000017](#), version 1. The development uses Lean 4 v4.35.0 and Mathlib v4.35.0. Every theorem stated in this paper is proved there in full; the one conditional result in the repository is described below.

The strongest bound is exported as

```
theorem P_le_polar_optimal :
  ∀ ε > (0 : ℝ), ∃ C > 0, ∀ a b : ℕ,
    0 < b → b < a → (P a b : ℝ) ≤
      C * (a : ℝ) ^ (((41301 : ℝ) - 3 * Real.sqrt 82849) / 137480 + ε)
```

The quantifier order makes uniformity explicit: C may depend on ε , but not on a or b . The function $P\ a\ b$ is defined by well-founded recursion,

$$P(a, 0) = 0, \quad P(a, b) = 1 + P(a, a \bmod b) \quad (b > 0),$$

and counts the positive terms of the zero-indexed orbit, as in Section 1. The rational bound 4317/14677 is exported separately as `P_le_4317_14677`. Both results are assembled in `PolarProjectionMain` from proved arithmetic, geometric and counting estimates. The repository README gives the correspondence between the other results and their Lean declarations.

The affine-curve theorem (27) is imported from the separate [ccdn-dimension-growth library at revision bb025e7](#), which is registered in the Palomar registry as [PALOMAR-2026-10-22-000015](#), version 1. That development proves the plane estimate, degree-preserving projection and box rescaling, using the same affine Hilbert degree as this paper. The present repository proves the monic Hilbert irreducibility estimate (Theorem 10.5) and its application to the branch-label family, so neither counting input is assumed.

The repository also contains a conditional refinement of Theorem 1.4, `P_le_loglog`, which proves $P(a, b) \ll a^{3/10} \sqrt{\log \log 3a}$ from two explicit hypotheses: the short-sum estimate for the Erdős–Hooley function from [10], over intervals $(x, x + y]$ with $y \geq x^{1/4}$, and the mean-value estimate from [2]. Their statements include the constants and thresholds; their proofs are not formalised here. This result is weaker than Theorem 1.1 and is not used elsewhere.

The proof library contains no `sorry` and uses no `native_decide`. Its axiom audit permits only `propext`, `Classical.choice` and `Quot.sound`, the standard axioms used by Mathlib. Sixteen public statements are compared with independent Mathlib-only specifications, and their proofs are replayed by Lean, NanoDa and con-ron. The specification file `PierceLengthBoundPalomar/Challenge.lean` deliberately contains theorem holes; it is separate from the proved solution. Build instructions, the full theorem map, and the verification commands are recorded in the README.

The divisor estimate uses `DivisorPowerBound.lean` from [openai/math \[11\]](#) under Apache-2.0, with unchanged proof bodies and adapted imports and module visibility. The repository records the source revision and the provenance of its other dependencies.

13 Limits of the method

The polar refinement brings the exponent close to 5/17, but does not remove that threshold. As explained in Section 1, 5/17 is the exponent obtained if every pattern carries $O(1)$ starts, and the estimates here never do better than that: the non-retained components and conics contribute $a^{o(1)}$

starts for each of the $X^{5+o(1)}$ patterns, and the exceptional patterns and the windows where the line-exclusion condition (22) fails also lead to exponent $5/17$, independently of the coordinate powers m, n . The branch-bad and degenerate patterns give only the exponents $717/2440$ and $732/2491$, below this threshold.

The improved graph degree gives

$$\beta_* - \frac{5}{17} = 0.00001600826\dots$$

Increasing the graph degree further can reduce this gap. Going below $5/17$ requires an improvement in the other counting contributions, or a different way to combine the information from successive orbit states. This is a limitation of the present estimates, not a lower bound for the Pierce iteration.

The elementary $3/10$ argument has a similar barrier: even with no divisor loss, its two window bounds balance at $A \asymp a^{3/5}$ (Section 5), so better constants or short-interval divisor estimates alone do not change that exponent.

Finally, there is no obvious monotone quantity along an orbit that could replace the window-by-window counting; in particular the curvatures need not decrease. For $a = 13242$, the orbit of 255 begins

$$255, 237, 207, 201, 177, 144, 138, 132, \dots$$

with quotients $51, 55, 63, 65, 74, 91, 95, 100, \dots$ and first five curvatures $24, 12, 6, 111, 30$. All these remainders lie in the single window $[128, 256)$. The rise bound gives one constraint on the order of the increments, but the present argument does not capture all compatibility conditions along an entire orbit. Shallit's conjectured $O((\log a)^2)$ bound remains open.

A Conic components

In this appendix we prove Proposition 8.10. Let C be a nonexceptional component of degree two. A component that is not retained already contributes $O(1)$ starts, by Lemma 8.3. We may therefore assume that the component C is retained. If it contains at most two free-node triples of good starts, there is nothing to prove. In the remaining case we first construct a parameter whose value is an integer at every such triple.

Lemma A.1 (A bounded integral parameter). *Let C be a retained nonexceptional degree-two component containing the free-node triples of at least three good starts. After permuting the three free-node coordinates to (x, y, z) , there are an integer $\Delta_c \neq 0$ and polynomials $\xi_0, \xi_1, \xi_2 \in \mathbb{Z}[t]$ such that*

$$\deg \xi_j \leq 2, \quad |\Delta_c|, \max_{j,k} |[t^k]\xi_j| \leq (aX)^{O(1)}.$$

The map

$$t \mapsto (\xi_0(t), \xi_1(t), \xi_2(t))/\Delta_c$$

parametrises C over $\mathbb{Q}$. Its inverse takes an integer value at every integer point of C .

Proof. Every free node of a good start has absolute value at most a . Three distinct such triples lie on an integer plane

$$\lambda_z z + \lambda_x x + \lambda_y y + \lambda_1 = 0$$

with coefficients of size $a^{O(1)}$: take a cross product of two differences as the normal, or, if it vanishes, a nonzero perpendicular vector of the same size. A plane not containing an integral curve of degree

two meets it in at most two distinct points. Thus this plane contains C . Choose the eliminated coordinate z with $\lambda_z \neq 0$.

The six rational pattern coordinates (8) have common denominator $N_0 N_1 \neq 0$ and integer numerators of size at most X^2 , so the pattern quadratic f has rational coefficients with numerators and denominators of size $X^{O(1)}$. Substitute $z = -(\lambda_x x + \lambda_y y + \lambda_1)/\lambda_z$ and clear denominators. The resulting plane equation is

$$\alpha_c x^2 + \beta_c xy + \gamma_c y^2 + \ell_x x + \ell_y y + \ell_0 = 0, \quad (40)$$

with integer coefficients of size $(aX)^{O(1)}$. The restriction of f to the plane is nonzero: otherwise the plane would be a component of the irreducible pattern quadric. It has degree two because it contains C , and hence defines C in that plane.

The pullback of the hyperplane at infinity to the normalisation of C has degree two. Corollary 8.5 says that each of its positive multiplicities is even. Counting residue degrees, there is therefore just one branch over infinity, of multiplicity two and residue degree one. The normalisation map is finite and surjective, so the projective conic meets the line at infinity only at the image of this rational point. Hence the leading binary quadratic in (40) has a single rational zero, of multiplicity two: it is a nonzero rational scalar times the square of a rational linear form. Interchanging x and y if necessary gives

$$\alpha_c \neq 0, \quad \beta_c^2 - 4\alpha_c \gamma_c = 0.$$

Set

$$\kappa_c = 2\beta_c \ell_x - 4\alpha_c \ell_y, \quad \epsilon_c = \ell_x^2 - 4\alpha_c \ell_0, \quad t = 2\alpha_c x + \beta_c y + \ell_x.$$

Completing the square in (40) gives

$$t^2 = \kappa_c y + \epsilon_c. \quad (41)$$

Here $\kappa_c \neq 0$. If it were zero, a rational point of C with parameter t_0 would give $\epsilon_c = t_0^2$, and the equation would factor as $(t - t_0)(t + t_0) = 0$. Since C is irreducible, it would lie in one of these two lines, contradicting its degree.

We can therefore solve explicitly:

$$y = \frac{t^2 - \epsilon_c}{\kappa_c}, \quad x = \frac{t - \beta_c y - \ell_x}{2\alpha_c}, \quad z = -\frac{\lambda_x x + \lambda_y y + \lambda_1}{\lambda_z}.$$

Substitution verifies the conic and plane equations for every rational t , and the displayed linear expression in x, y recovers t . In particular, t is an integer whenever (x, y, z) is an integer point. Distinct points have distinct parameters. Taking $\Delta_c = 2\alpha_c \kappa_c \lambda_z$ clears all three denominators, and all the quantities involved are polynomials of bounded degree in coefficients of size $(aX)^{O(1)}$, which gives the stated bounds. $\square$

Lemma A.2 (Properness and coefficient height). *In the parameter of Lemma A.1, write*

$$E_0(t) = \frac{F_E(t)}{G_E(t)}, \quad F_E, G_E \in \mathbb{Z}[t], \quad \gcd_{\mathbb{Q}[t]}(F_E, G_E) = 1.$$

The polynomials may be chosen so that

$$\deg G_E - \deg F_E = 3, \quad \max(\deg F_E, \deg G_E) \leq C_{\deg}, \quad \max_k (|[t^k]F_E|, |[t^k]G_E|) \leq C_{\text{ht}} a^{C_{\text{exp}}},$$

for absolute constants $C_{\deg}, C_{\text{ht}}, C_{\text{exp}}$. In particular, G_E is nonconstant. At the parameter t of every good start, $G_E(t) \neq 0$ and $F_E(t)/G_E(t) = I_3(i)$ is an integer.

Proof. The explicit parametrisation identifies C with the affine line and its smooth projective model with $\mathbb{P}^1$. Its point at infinity has hyperplane multiplicity two. Lemma 8.7 therefore gives

$$\text{ord}_{t=\infty} E_0 = 3.$$

The function is nonzero, since it is positive at a good start. For a reduced rational function on $\mathbb{P}^1$, its order at infinity is $\deg G_E - \deg F_E$, proving properness.

We next bound the coefficients uniformly in the component, including after cancellation. The six rational pattern coordinates, the three coordinate polynomials ξ_j/Δ_c , and the integer a can be written over the common nonzero denominator $N_0 N_1 \Delta_c$, and their numerators are polynomials in t of degree at most two with coefficients of size $(aX)^{O(1)}$. The cofactor formula defining E_0 is a quotient of two fixed polynomials in these inputs. Substituting and clearing a fixed power of the common denominator gives nonzero integer polynomials $\tilde{n}, \tilde{d}$ of bounded degree with coefficients of size $(aX)^{O(1)}$.

Choose a primitive $\Pi \in \mathbb{Z}[t]$ representing $\gcd_{\mathbb{Q}[t]}(\tilde{n}, \tilde{d})$. By Gauss's lemma, Π divides both $\tilde{n}$ and $\tilde{d}$ in $\mathbb{Z}[t]$, so $F_E = \tilde{n}/\Pi$ and $G_E = \tilde{d}/\Pi$ are coprime integer polynomials with the same ratio. Every integer factor of a nonzero integer polynomial of degree at most $C_{\deg}$ and height $\mathcal{H}$ has height at most $2^{C_{\deg}}(C_{\deg} + 1)\mathcal{H}$, by comparing Mahler measures; so cancellation preserves the bound $(aX)^{O(1)}$. Since $X \leq a^{O(1)}$, this is the asserted $C_{\text{ht}} a^{C_{\text{exp}}}$. Finally, the reconstruction denominator is nonzero at every good start, hence so are $\tilde{d}$ and its factor G_E at the corresponding parameter, and the reconstruction identity gives $F_E(t)/G_E(t) = I_3(i) \in \mathbb{Z}$. $\square$

Lemma A.3 (Integer values of a rational function). *Let $F_E, G_E \in \mathbb{Z}[t]$ be coprime over $\mathbb{Q}$, of degrees at most $C_{\deg}$, with G_E nonconstant and coefficient heights at most $C_{\text{ht}} a^{C_{\text{exp}}}$. For every $\eta > 0$, the number of integers t such that $G_E(t) \neq 0$ and $F_E(t)/G_E(t) \in \mathbb{Z}$ is $O_\eta(a^\eta)$. The implied constant depends only on $\eta, C_{\deg}, C_{\text{ht}}, C_{\text{exp}}$.*

Proof. The integer resultant $R = \text{Res}(F_E, G_E)$ is nonzero. The adjugate of the Sylvester matrix gives $U, V \in \mathbb{Z}[t]$ with $UF_E + VG_E = R$. At any parameter in question, $G_E(t) \mid F_E(t)$, and hence $G_E(t) \mid R$. There are $2\tau(|R|)$ possible signed values of $G_E(t)$. For each such value v , the nonzero polynomial $G_E - v$ has at most $\deg G_E \leq C_{\deg}$ roots. Thus the number of parameters is at most

$$2C_{\deg}\tau(|R|).$$

The Sylvester matrix has size at most $2C_{\deg}$, and its entries are bounded by $C_{\text{ht}} a^{C_{\text{exp}}}$. Expanding its determinant gives

$$1 \leq |R| \leq (2C_{\deg})! (C_{\text{ht}} a^{C_{\text{exp}}})^{2C_{\deg}}.$$

The divisor estimate $\tau(n) \ll_\varepsilon n^\varepsilon$, with ε chosen in terms of η and the fixed degree and height exponents, proves the claim. $\square$

Proof of Proposition 8.10. Apply the last lemma to the reconstructed invariant E_0 . The parameter is injective on free-node triples, and reconstruction of the strictly increasing q_i makes those triples injective in the good starts. Each retained nonexceptional conic therefore contributes $O_\eta(a^\eta)$ starts. The components with at most two free-node triples of good starts and the nonretained components were already bounded absolutely. Apply the conic estimate with $\eta/2$ and the pattern estimate $O_{\eta/2}(a^{\eta/2} X^5)$. Since there are at most twenty-four components per pattern, all nonexceptional conics together contribute $O_\eta(a^\eta X^5)$. $\square$

B The upper bound for graph degrees

Proof of the upper bound in Lemma 8.13. Only total polynomial degrees in the nodes matter here, not the heights of the pattern-dependent coefficients. The cofactors c_j of Section 8.2 have bounded degree, hence so do the numerator and denominator of $E_0 = E_n/E_d$. Since H_3, H'_3 have degree at most three and the increments have degree one, the three curvature coordinates have the common denominator $\Omega = E_d H_3 k_1$, and Ω and the numerators $\mathcal{N}_0, \mathcal{N}_1, \mathcal{N}_2$ have degree at most an absolute constant C_W . The numerator for J with this denominator is $\mathcal{N}_3 = J\Omega$, of degree at most $C_W + 3$. The denominator is nonzero on C , since it is nonzero at a good start.

Put $\bar{m} = m + n$, and represent the powered coordinates with denominator $\Omega^{\bar{m}}$ and numerators

$$\mathcal{N}_j^m \Omega^{\bar{m}-m} \quad (0 \leq j \leq 2), \quad \mathcal{N}_3^n \Omega^{\bar{m}-n}.$$

All these numerators, and those of the node coordinates written with the same denominator, have degree at most $(C_W + 3)\bar{m} + 1$. So multiplying a polynomial of degree at most t in the graph coordinates by $\Omega^{\bar{m}t}$ gives a polynomial in the original nodes of degree at most $((C_W + 3)\bar{m} + 1)t$. This operation is injective on the corresponding spaces of functions on C , because Ω is nonzero in $\mathbb{Q}(C)$. Comparing affine Hilbert functions gives

$$\delta t + O(1) \leq d((C_W + 3)\bar{m} + 1)t + O(1),$$

and hence the asserted bound with $C_{\text{gr}} = C_W + 4$. $\square$

C A reduced basis in dimension three

The constant $64/27 = (4/3)(16/9)$ in Lemma 9.3 comes from two successive size reductions. We give the elementary argument for completeness.

Proof of Lemma 9.3. Write $\det \Lambda$ for the least nonzero $|\det|$ of a triple in Λ ; the triple $(\kappa \mathbf{e}_1, \kappa \mathbf{e}_2, \kappa \mathbf{e}_3)$ gives $\det \Lambda \leq \kappa^3$. A triple $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ with $|\det| = \det \Lambda$ is a basis: if not, some $x \in \Lambda$ has a Cramer coordinate that is not an integer, say the first, $\det(x, \mathbf{f}_2, \mathbf{f}_3)/\det(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$, and subtracting the integer part of that coordinate times $\mathbf{f}_1$ from x gives a triple with smaller nonzero determinant. Among all bases (all have $|\det| = \det \Lambda$) minimise $g_1 = |\mathbf{f}_1|^2$, and then the Gram determinant $g_2 = g_1|\mathbf{f}_2|^2 - \langle \mathbf{f}_1, \mathbf{f}_2 \rangle^2$. Put $g_3 = (\det \Lambda)^2$.

Choose an integer z_1 with $|2(\langle \mathbf{f}_1, \mathbf{f}_2 \rangle + z_1 g_1)| \leq g_1$. Then $(\mathbf{f}_2 + z_1 \mathbf{f}_1, \mathbf{f}_1, \mathbf{f}_3)$ is a basis, so $g_1 \leq |\mathbf{f}_2 + z_1 \mathbf{f}_1|^2$, and $g_1|\mathbf{f}_2 + z_1 \mathbf{f}_1|^2 = g_2 + (\langle \mathbf{f}_1, \mathbf{f}_2 \rangle + z_1 g_1)^2$. Hence $3g_1^2 \leq 4g_2$.

Let $\langle x, x' \rangle_1 = g_1 \langle x, x' \rangle - \langle \mathbf{f}_1, x \rangle \langle \mathbf{f}_1, x' \rangle$, so that $\langle \mathbf{f}_2, \mathbf{f}_2 \rangle_1 = g_2$, and Sylvester's identity gives $\langle \mathbf{f}_2, \mathbf{f}_2 \rangle_1 \langle \mathbf{f}_3, \mathbf{f}_3 \rangle_1 - \langle \mathbf{f}_2, \mathbf{f}_3 \rangle_1^2 = g_1 g_3$. Choose an integer z_2 with $|2(\langle \mathbf{f}_2, \mathbf{f}_3 \rangle_1 + z_2 g_2)| \leq g_2$ and put $y = \mathbf{f}_3 + z_2 \mathbf{f}_2$. Then $(\mathbf{f}_1, y, \mathbf{f}_2)$ is a basis with first vector $\mathbf{f}_1$, so $g_2 \leq g_1|y|^2 - \langle \mathbf{f}_1, y \rangle^2 = \langle y, y \rangle_1$, while $g_2 \langle y, y \rangle_1 = (\langle \mathbf{f}_2, \mathbf{f}_3 \rangle_1 + z_2 g_2)^2 + g_1 g_3$. Hence $3g_2^2 \leq 4g_1 g_3$.

Finally choose an integer z_3 with $|2(\langle \mathbf{f}_1, y \rangle + z_3 g_1)| \leq g_1$, and use the basis $\mathbf{f}_1, \mathbf{f}'_2 = \mathbf{f}_2 + z_1 \mathbf{f}_1, \mathbf{f}'_3 = y + z_3 \mathbf{f}_1$. The identities above give $3g_1|\mathbf{f}'_2|^2 \leq 4g_2$ and $4g_1 g_2 |\mathbf{f}'_3|^2 \leq g_2^2 + 4g_1 g_3 + g_1^2 g_2$. With the two inequalities this gives $9g_2 |\mathbf{f}'_3|^2 \leq 16g_3$. Since $|\mathbf{f}_1|^2 = g_1$, the product of the bounds $|\mathbf{f}'_2|^2 \leq \frac{4}{3}g_2/g_1$ and $|\mathbf{f}'_3|^2 \leq \frac{16}{9}g_3/g_2$ is the claim for the basis $\mathbf{f}_1, \mathbf{f}'_2, \mathbf{f}'_3$; rename $\mathbf{f}'_2, \mathbf{f}'_3$ as $\mathbf{f}_2, \mathbf{f}_3$. Replacing $\mathbf{f}_1$ by $-\mathbf{f}_1$ if necessary makes the determinant positive. For the coordinates, $x_1 \det \Lambda = \det(x, \mathbf{f}_2, \mathbf{f}_3)$, and Hadamard's inequality gives $\det(x, \mathbf{f}_2, \mathbf{f}_3)^2 \leq |x|^2 |\mathbf{f}_2|^2 |\mathbf{f}_3|^2$, so $27|\mathbf{f}_1|^2 x_1^2 (\det \Lambda)^2 \leq 64(\det \Lambda)^2 |x|^2$; likewise for x_2 and x_3 . $\square$

D Effective Hilbert irreducibility for monic families

We prove Theorem 10.5. The letters in this appendix are local to it.

Proof of Theorem 10.5. Write $\lg(x) = 1 + \log(x + 1)$. Implied constants depend only on r, d .

Step 1: integer roots in one parameter. Let $\mathcal{A} \in \mathbb{Z}[u][Z]$ be monic in Z , of Z -degree $e \geq 2$, irreducible over $\mathbb{Q}(u)$, of total degree at most D and height at most H_0 . We claim that

$$\#\{v \in [-B, B] \cap \mathbb{Z} : \mathcal{A}(v, Z) \text{ has an integer root}\} \ll_D \lg(H_0)^4 B^{1/e} (1 + \log B).$$

An integer root z of $\mathcal{A}(v, Z)$ satisfies $|z| \leq 1 + (D + 1)H_0 B^D$. Choose an integer $k > D$ with $k \asymp_D \lg(H_0)$, and apply the shear $(u, Z) \mapsto (u, Z + u^k)$. The polynomial $\mathcal{A}_k(u, Z) = \mathcal{A}(u, Z - u^k)$ is irreducible in $\mathbb{Q}[u, Z]$, and its total degree is exactly ke : the term $(Z - u^k)^e$ contributes $\pm u^{ke}$, while every other term has degree at most $D + k(e - 1) < ke$. The points $(v, z + v^k)$ are integer points of the plane curve $\mathcal{A}_k = 0$ in a box of size $B' \ll_D H_0 B^k$. By (27) with $\delta = ke$ there are $\ll (ke)^3 B^{1/(ke)} (\log B' + ke)$ of them. Here $B^{1/(ke)} \ll_D H_0^{1/(ke)} B^{1/e} \ll B^{1/e}$ by the choice of k , while $ke \ll \lg(H_0)$ and $\log B' \ll \lg(H_0)(1 + \log B)$. This proves the claim.

Step 2: one parameter. Let $r = 1$, and let $y_1, \dots, y_e$ be the roots of F in an algebraic closure of $\mathbb{Q}(u)$; they are integral over $\mathbb{Z}[u]$. For a nonempty proper subset $I \subset \{1, \dots, e\}$ and $1 \leq j \leq |I|$, let $\sigma_{I,j}$ be the j th elementary symmetric function of $(y_i)_{i \in I}$. Since F is irreducible, $\prod_{i \in I} (\zeta - y_i) \notin \mathbb{Q}(u)[\zeta]$, so some $\sigma_I := \sigma_{I,j(I)}$ lies outside $\mathbb{Q}(u)$. The resolvent $\prod_{|I'|=|I|} (Z - \sigma_{I',j(I)})$ is monic in Z , and its coefficients are fixed integer polynomials in the coefficients of F ; so it lies in $\mathbb{Z}[u][Z]$ with degree $O(1)$ and height $O(\mathcal{H}^{O(1)})$. Let $\mathcal{A}_I$ be its monic irreducible factor over $\mathbb{Q}(u)$ that vanishes at σ_I . By Gauss's lemma $\mathcal{A}_I \in \mathbb{Z}[u][Z]$; by Gelfond's factor-height inequality its height is $O(\mathcal{H}^{O(1)})$; and its Z -degree is at least two, because $\sigma_I \notin \mathbb{Q}(u)$.

Now suppose that $F(v, \zeta)$ is reducible over $\mathbb{Q}$. By Gauss's lemma it has a monic factor $g \in \mathbb{Z}[\zeta]$ of degree e' with $1 \leq e' < e$. Extend the specialisation $u \mapsto v$ to a ring homomorphism ϕ from $\mathbb{Z}[u][y_1, \dots, y_e]$ to $\overline{\mathbb{Q}}$, which is possible because this ring is integral over $\mathbb{Z}[u]$. Then $F(v, \zeta) = \prod_i (\zeta - \phi(y_i))$, so the roots of g , with multiplicity, are $\phi(y_i)$ for i in some I with $|I| = e'$, and the coefficients of g are $\pm \phi(\sigma_{I,j})$. In particular $\phi(\sigma_I)$ is an integer root of $\mathcal{A}_I(v, Z)$. Step 1, applied to the $O(1)$ polynomials $\mathcal{A}_I$, bounds the number of such v by $O(\lg(\mathcal{H})^4 B^{1/2} (1 + \log B))$.

Step 3: one parameter, the others free. Let $r \geq 2$. We claim that for all but $O(\lg(\mathcal{H})^4 B^{1/2} (1 + \log B))$ integers $v_1 \in [-B, B]$, the polynomial $F(v_1, u_2, \dots, u_r, \zeta)$ is irreducible over $\mathbb{Q}(u_2, \dots, u_r)$. Fix an integer $b > d$ and let $\mathcal{K}$ be the Kronecker map

$$u_2^{a_2} \dots u_r^{a_r} \zeta^{a_{r+1}} \longmapsto \omega^{a_2 + a_3 b + \dots + a_{r+1} b^{r-1}}.$$

It is a linear bijection from polynomials with all partial degrees less than b onto polynomials in ω of degree less than b^r , it commutes with specialising u_1 , and it is multiplicative on products whose partial degrees stay below b . Since ζ occupies the highest position, $\mathcal{K}(F) \in \mathbb{Z}[u_1][\omega]$ is monic. Factor $\mathcal{K}(F) = \prod_\nu \mathcal{A}_\nu$ into monic irreducible factors over $\mathbb{Q}(u_1)$. Again $\mathcal{A}_\nu$ lies in $\mathbb{Z}[u_1][\omega]$, with bounded degree and height $O(\mathcal{H}^{O(1)})$. By Step 2, outside an exceptional set of the stated size, every $\mathcal{A}_\nu(v_1, \omega)$ is irreducible over $\mathbb{Q}$.

Let v_1 be outside this set, and suppose $F(v_1, \cdot) = gh$ with $g, h \in \mathbb{Q}[u_2, \dots, u_r][\zeta]$ monic in ζ of positive ζ -degree. All partial degrees of g and h are at most those of F , hence less than b , so $\mathcal{K}(g)\mathcal{K}(h) = \prod_\nu \mathcal{A}_\nu(v_1, \omega)$, and by unique factorisation $\mathcal{K}(g) = \prod_{\nu \in S} \mathcal{A}_\nu(v_1, \omega)$ for some set S of indices. Put $g_S = \mathcal{K}^{-1}(\prod_{\nu \in S} \mathcal{A}_\nu)$ and define g_{S^c} likewise, so that $g_S(v_1) = g$ and $g_{S^c}(v_1) = h$. If $g_S g_{S^c} = F$ identically, this is a factorisation of F over $\mathbb{Q}(u_1, \dots, u_r)$ into factors of positive ζ -degree, which is impossible. Otherwise some coefficient of $g_S g_{S^c} - F$ is a nonzero polynomial in u_1 of bounded degree vanishing at v_1 . There are $O(1)$ sets S , so this adds $O(1)$ exceptional values.

Step 4: induction on r . The case $r = 1$ is Step 2. For $r \geq 2$, the exceptional values v_1 of Step 3 contribute at most their number times $(2B + 1)^{r-1}$. For every other $v_1 \in [-B, B]$, the family $F(v_1, u_2, \dots, u_r, \zeta)$ is monic, irreducible over $\mathbb{Q}(u_2, \dots, u_r)$, of the same degree bounds, and of height at most $(d + 1)\mathcal{H}B^d$. Since $\lg((d + 1)\mathcal{H}B^d) \ll \lg(\mathcal{H})(1 + \log B)$, the induction hypothesis bounds its reducible specialisations by $O(\lg(\mathcal{H})^4 B^{r-3/2}(1 + \log B)^{4r+1})$. Summing over the $2B + 1$ values of v_1 gives the theorem. $\square$

E Degree certificates along the line directions

This appendix records the certificates used in step 3 of the proof of Lemma 8.6. Fix one of the four directions w , write the nodes along the line as $s_j + tw_j$, and expand

$$\frac{P_4(x)R_4(y) - P_4(y)R_4(x)}{x - y} = \sum_{\alpha, \beta < 4} \mathcal{B}'_{\alpha\beta} M_\alpha(x) M_\beta(y), \quad M_\alpha(x) = \prod_{l < \alpha} (x - n_l),$$

with the nodes n_0, n_1, n_2 below. As for (10), $\det \mathcal{B}' = \text{Res}(P_4, R_4)$, and $\text{Res}(P_4, Z) = \gamma_0 \text{Res}(P_4, R_4)$ has the same degree in t . In each case the entry $\mathcal{B}'_{\alpha\beta}$ has t -degree at most $\varrho_\alpha + \kappa_\beta$, for all values of the γ_j , the γ'_j and the nodes s_j ; since $\sum_\alpha \varrho_\alpha + \sum_\beta \kappa_\beta = 6$, every term of the determinant has t -degree at most six.

Direction w	Nodes (n_0, n_1, n_2)	Row weights ϱ	Column weights κ
$(0, 1, 1, 0, 1)$	$(s_1, s_2, 0)$	$(1, 1, 0, 0)$	$(1, 1, 1, 1)$
$(0, 1, 1, 0, 0)$	$(s_1, 0, s_2)$	$(2, 1, 1, 0)$	$(1, 1, 0, 0)$
$(0, 0, 1, 1, 0)$	$(s_1, s_2, 0)$	$(1, 1, 1, 0)$	$(1, 1, 1, 0)$
$(0, 1, 0, 0, 1)$	$(0, s_1, s_2)$	$(2, 1, 1, 1)$	$(1, 0, 0, 0)$

Each entry is an explicit polynomial of degree at most four in t , so each bound is a finite expansion. We checked all sixteen entries in each case with a computer algebra system, and the bounds are formalised as `line_resultant_natDegree_le_six`. For generic coefficients the determinant has degree exactly six in each case.

F Certificates for the polar bound

F.1 The quartic initial form at the base point

At $(\rho, \varphi, \chi, \rho', \varphi', \chi') = (2, 5, -3, 3, 6, -5)$ the polar base point is $\mathbf{k}_* = (-1/7, -3/28, -1/2)$ and $\nabla f(\mathbf{k}_*) = (-4, 9, -5)$. Use tangent coordinates (ξ_1, ξ_2) with first-order displacement

$$\delta \mathbf{k} = ((9\xi_1 - 5\xi_2)/4, \xi_1, \xi_2).$$

Thus the two axes used in Lemma 10.2 have tangent vectors $(9/4, 1, 0)$ and $(-5/4, 0, 1)$. Direct substitution in the Newton quartic, its remainder and the norm polynomial gives

$$\begin{aligned} P_{4,*}(x) &= x^4 - \frac{20}{7}x^3 + \frac{757}{112}x^2 - \frac{669}{112}x + 2, \\ Z_1(x; \xi_1, \xi_2) &= \frac{77\xi_1 + 7\xi_2}{14} + \frac{-56\xi_1 - 124\xi_2}{14}x + \frac{-45\xi_1 + 157\xi_2}{14}x^2 + \frac{28\xi_1 - 28\xi_2}{14}x^3, \\ \Psi_1(\xi_1, \xi_2) &= \frac{3\xi_1 + 85\xi_2}{28}. \end{aligned}$$

The resultant is

$$\mathcal{R}_*(\xi_1, \xi_2) = \frac{32233977\xi_1^4 - 95047956\xi_1^3\xi_2 + 147873750\xi_1^2\xi_2^2 - 111199572\xi_1\xi_2^3 + 31702169\xi_2^4}{6272}.$$

In particular,

$$\mathcal{R}_*(1, 0)\Psi_1(0, 1)^4 - \mathcal{R}_*(0, 1)\Psi_1(1, 0)^4 = \frac{26291112152499}{60236288}.$$

The realising matrix in Lemma 10.2 shows that Θ is nonzero on the Plücker cone; no orbit with these parameters is required.

F.2 The irreducible sextic

The specialisation $E_\gamma = 24E_\star$ at $(\gamma_1, \gamma_2, \gamma_3) = (-7, 8, -4)$ reduces modulo 71 to a nonzero scalar multiple of

$$q(T) = T^6 + 20T^5 - 14T^4 + 15T^3 - 28T^2 + 5T - 4.$$

More precisely, $E_\star = 57q$ in $\mathbb{F}_{71}[T]$. The following certificates prove that q is irreducible. Repeated squaring modulo q gives

$$\begin{aligned} T^{71^2} &\equiv r_2(T) := 36T^5 + 48T^4 + 49T^3 + 53T^2 + 36T + 42, \\ T^{71^3} &\equiv r_3(T) := 37T^5 + 22T^4 + 60T^3 + 13T^2 + 66T + 35. \end{aligned}$$

For checking the iteration one may first use

$$T^{71} \equiv 14T^5 + 33T^4 + 31T^3 + 34T^2 + 48T \pmod{q}$$

and raise each remainder to the 71st power. Define

$$\begin{aligned} U_2 &= 20T^4 + 34T^3 + 31T^2 + 13T + 24, \\ V_2 &= 31T^5 + 61T^4 + 42T^3 + 61T^2 + 39T + 4, \\ U_3 &= 25T^4 + 11T^3 + 53T^2 + 59T + 48, \\ V_3 &= 7T^5 + 30T^4 + 65T^3 + 64T^2 + 25T + 40. \end{aligned}$$

Direct multiplication in $\mathbb{F}_{71}[T]$ verifies

$$U_2q + V_2(r_2 - T) = 1, \quad U_3q + V_3(r_3 - T) = 1.$$

Thus q has no irreducible factor of degree one or two, since such a factor would divide $T^{71^2} - T$, and none of degree three, since it would divide $T^{71^3} - T$. Every reducible sextic has an irreducible factor of degree at most three, so q is irreducible. Reduction modulo 71 preserves degree, and Gauss's lemma proves $E_\star$ irreducible over $\mathbb{Q}$.

Declaration of AI-generated research and authorship

This manuscript is the product of an AI-driven research process. The human author initiated the investigation and directed AI systems throughout, and wrote no part of the manuscript or of the Lean code. GPT-6-Astra (OpenAI) found the results and the key steps of their proofs. These include the orthogonal polynomials and adjacent labels behind Section 7, the first proof of a bound $a^{3/10-\delta}$ through resultants, the use of Lemma 7.3 to bound the number of states per pattern, the

sign identities of Lemma 7.5 and the slice count behind the 11/37 bound, the observation that the pattern integers form a Plücker vector, the weighting of minor vectors by the shortest vectors of their lattices behind the 45/152 bound, the joint-invariant graphs with their normalisation and valuation arguments, the exterior-lattice coordinates, and the polar identity and Hilbert sieve. The AI systems also developed the presentation, the formalisation and the writing. The systems were GPT-6-Astra (OpenAI), used through Codex, and Claude Opus 5.5 (Anthropic), used through Claude Code. Both contributed to the Lean formalisation and to the review and revision of the manuscript.

The Lean development gives a machine-checkable version of the results. Lean’s kernel checks the proofs, including the Mathlib results and the separate curve-counting library they use, relative to the axioms `propext`, `Classical.choice` and `Quot.sound` (Section 12); the sixteen selected statements are also replayed by two independent kernels. The conditional refinement described in Section 12 assumes two published theorems as explicit hypotheses. These automated checks are not independent human peer review.

This declaration distinguishes the provenance of the work from conventional AI-assisted writing. The human author accepts responsibility for the submission and for the accuracy of this disclosure.

Acknowledgements

We thank the Mathlib community, and the authors of `openai/math` for the formalised divisor bound.

References

- [1] O. Baraskar and I. Vukusic, *Bounds for sets of remainders*, [arXiv:2508.20853](#) (2025).
- [2] R. de la Bretèche and G. Tenenbaum, *Note on the mean value of the Erdős–Hooley Delta-function*, *Acta Arith.* **219** (2025), no. 4, 379–394, [doi:10.4064/aa250107-13-2](#).
- [3] W. Castryck, R. Cluckers, P. Dittmann and K. H. Nguyen, *The dimension growth conjecture, polynomial in the degree and without logarithmic factors*, *Algebra Number Theory* **14** (2020), no. 8, 2261–2294, [doi:10.2140/ant.2020.14.2261](#).
- [4] R. Cluckers, P. Dèbes, Y. I. Hendel, K. H. Nguyen and F. Vermeulen, *Improvements on dimension growth results and effective Hilbert’s irreducibility theorem*, *Forum Math. Sigma* **13** (2025), e153, [doi:10.1017/fms.2025.10096](#).
- [5] Z. Chase and M. Pandey, *On the length of Pierce expansions*, [arXiv:2211.08374](#) (2022).
- [6] P. Erdős and J. O. Shallit, *New bounds on the length of finite Pierce and Engel series*, *J. Théor. Nombres Bordeaux* **3** (1991), no. 1, 43–53, [doi:10.5802/jtnb.41](#).
- [7] J. Franke, Yu. I. Manin and Yu. Tschinkel, *Rational points of bounded height on Fano varieties*, *Invent. Math.* **95** (1989), 421–435, [doi:10.1007/BF01393904](#).
- [8] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: *Automated Deduction – CADE 28, Lecture Notes in Comput. Sci.* **12699**, Springer, 2021, 625–635, [doi:10.1007/978-3-030-79876-5_37](#).

- [9] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), 367–381, [doi:10.1145/3372885.3373824](https://doi.org/10.1145/3372885.3373824).
- [10] M. Nair and G. Tenenbaum, *Short sums of certain arithmetic functions*, Acta Math. **180** (1998), 119–144, [doi:10.1007/BF02392880](https://doi.org/10.1007/BF02392880).
- [11] OpenAI, openai/math, <https://github.com/openai/math>.
- [12] J. O. Shallit, *A discrete iteration related to Pierce expansions*, Open Problem Garden, https://www.openproblemgarden.org/op/a_discrete_iteration_related_to_pierce_expansions.
- [13] T. A. Pierce, *On an algorithm and its use in approximating roots of algebraic equations*, Amer. Math. Monthly **36** (1929), no. 10, 523–525, [doi:10.1080/00029890.1929.11987017](https://doi.org/10.1080/00029890.1929.11987017).
- [14] J. O. Shallit, *Metric theory of Pierce expansions*, Fibonacci Quart. **24** (1986), no. 1, 22–40, [doi:10.1080/00150517.1986.12429786](https://doi.org/10.1080/00150517.1986.12429786).
- [15] J. O. Shallit, *Talks*, <https://cs.uwaterloo.ca/~shallit/talks.html> (accessed October 2026).