hexagon:2610.00187v7

AN IRRATIONALITY MEASURE FOR CATALAN'S CONSTANT

SEAN LITSEY

ABSTRACT. Let $G = \sum_{j\ge 0}(-1)^j(2j+1)^{-2}$ be Catalan's constant. We prove that $|G - a/b| > b^{-270}$ for all integers a and all sufficiently large integers b. Hence the irrationality exponent of G is at most 270. The proof refines OpenAI's recent proof that G is irrational. That proof compares two bounds for a large determinant whose entries are linear forms in 1 and G. The lower bound comes from the denominators of the entries, and the upper bound from an integral representation. We improve both. On the arithmetic side we show that the matrix of rational parts has large corank modulo p for the primes in explicit intervals, which lowers the denominators. On the analytic side we bound a determinant that the energy method treats trivially. A phase inequality for Pick matrices, proved by Nevanlinna–Pick interpolation, compares it with a determinant of a simpler kind. That determinant is bounded by a finite Gram inequality and two logarithmic gases. Gauss quadrature at Chebyshev nodes of many short intervals extends the upper bound to an exponentially small neighbourhood of G, as a measure requires. The proof is formalized in the Lean proof assistant and kernel-checked.

CONTENTS

1. Introduction 2
2. The construction and the criterion 8
3. Denominators and nonvanishing 11
4. Corank modulo p 16
5. The neighbourhood of G 27
6. The energy method with prescribed columns 32
7. A bound for the row determinant 37
8. Proof of Theorem 1.1 52
9. Remarks on the method 53
10. Formalization 54
Disclosure of AI use 55
Appendix A. Counting on every subspace 55

Date: November 8, 2026.

2020 *Mathematics Subject Classification.* Primary 11J82; Secondary 11J72, 15A15, 31A15, 68V20.

Key words and phrases. Catalan's constant, irrationality measure, determinants, corank modulo p, Andréief's identity, logarithmic energy, Pick matrices, Nevanlinna–Pick interpolation, Gauss quadrature, Lean.

Appendix B. The bounds for the energy method 57
Appendix C. The reference measures of Section 7 59
References 61

1. Introduction

1.1. **The result.** Catalan's constant

$$G = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)^2} = L(2, \chi_{-4}) = 0.91596\,55941\,77219\ldots$$

is the value at $s = 2$ of the Dirichlet L-function of the nontrivial character modulo 4. Values of L-functions at integers where no formula in terms of π is available are notoriously hard to study. Apéry proved $\zeta(3)$ irrational in 1978 [1], and Calegari, Dimitrov and Tang proved the Calegari–Dimitrov–Tang constant $L_{-3} = L(2, \chi_{-3})$ irrational in 2024 [4]. For G the classical methods produce linear forms that do not decay fast enough, and its irrationality remained open. In October 2026 OpenAI [36] proved that G is irrational, with a new argument based on a large determinant, and verified the proof in Lean. The present paper turns that argument into an irrationality measure.

The *irrationality exponent* $\mu(\xi)$ of a real number ξ is the infimum of the real numbers μ such that $|\xi - a/b| > b^{-\mu}$ for all integers a and all sufficiently large integers b. Every irrational number has $\mu(\xi) \geq 2$, and almost every real number has $\mu(\xi) = 2$. An *irrationality measure* is an explicit upper bound for $\mu(\xi)$. The exponent does not change under $\xi \mapsto r\xi + r'$ with $r, r' \in \mathbb{Q}$ and $r \neq 0$.

Theorem 1.1. *For all integers a and all sufficiently large integers b,*

$$\left| G - \frac{a}{b} \right| > b^{-270}.$$

In particular $\mu(G) \leq 270$.

The proof uses the prime number theorem. The theorem and every result it depends on have been formalized in the Lean 4 proof assistant on top of OpenAI's formalization, and the proof is kernel-checked. The formal proof has no unproved steps and uses only the standard axioms (Section 10). The use of AI in this work is disclosed at the end of the paper.

1.2. **Background.** *Catalan's constant.* Irrationality proofs for values of L-functions at integers usually follow Apéry's proof for $\zeta(3)$. One constructs linear forms $q_n\xi - p_n$ with integer coefficients that tend to zero faster than the coefficients grow. If $|q_n\xi - p_n| = e^{-\alpha n + o(n)}$ and $|q_n| = e^{\beta n + o(n)}$ with $\alpha > 0$, then ξ is irrational and $\mu(\xi) \leq 1 + \beta/\alpha$. Beukers [2] obtained Apéry's forms for $\zeta(2)$ and $\zeta(3)$ as integrals of Legendre-type polynomials. Hata [21] showed how much is gained by dividing the forms by the primes that divide all their coefficients, an idea that goes back to Chudnovsky. Rhin and Viola

[39, 40] found finite groups acting on families of such integrals. The invariance yields a large common factor of the coefficients, and this *arithmetic method* gave the best measures for $\zeta(2)$ and $\zeta(3)$ for many years. Marcovecchio [33] applied it to $\log 2$, and Zeilberger and Zudilin [51] to π, using integrals of Salikhov [45]. For Catalan's constant the analogous linear forms exist but do not decay fast enough. Rivoal and Zudilin [42] proved that one of the numbers $\beta(2), \beta(4), \ldots, \beta(2m)$ is irrational for suitable m, where β is the Dirichlet beta function. This does not single out $G = \beta(2)$. Fischler [16] and Kumar and Tosi [27] obtained lower bounds for the number of irrational values among $\beta(2), \beta(4), \ldots, \beta(2m)$. These bounds again do not single out G. Zudilin [53] found an Apéry-like recurrence for G, whose solutions give linear forms that tend to zero too slowly. Related constructions and their limits are described in [52, 41, 24, 25, 56]. Nesterenko [35] developed a general construction of such approximations and compared their quality. Rational approximations to G with $|G - p_n/q_n| \le q_n^{-0.52}$ follow from these works. Eskandari, Murty and Nemoto [12] construct a mixed motive with period G and obtain linear forms in 1 and G from it. Eskandari [11] raises the exponent 0.52 to 0.62. An exponent below 1 does not imply irrationality.

Recent results. Since September 2026 a series of new results in this area has appeared, several of them obtained with AI systems. OpenAI proved $\mu(\pi) = 2$, with a Lean formalization [37]. Liu, Jiang and Zhang proved $\mu(\log r) = 2$ for every rational $r > 0$ with $r \neq 1$ [32]. Chudnovsky adapted OpenAI's method to periods of elliptic curves and proved $\mu(\Omega) = 2$ for the lemniscate period Ω [6]. Giraudin lowered the bound for $\mu(\pi^2)$ to 5.0193784 with two hypergeometric constructions of Zudilin [54] at new parameters [18], and the bound for $\mu(\zeta(3))$ to 5.5138800 [19]. Leibtag, Barkan, Kalisch, Weinbaum, Shalyt and Kaminer [30] showed that the classical constructions are one-dimensional trajectories in conservative matrix fields. Searching over other trajectories, they obtained $\mu(\pi^2) < 4.62675$ and $\mu(\zeta(3)) < 5.4197$, and further Apéry-like approximations to G, the best of which still has a negative approximation exponent. Prellberg [38] proved that $\mathrm{Li}_2(-1/4)$, $\mathrm{Li}_2(1/5)$ and $\mathrm{Li}_2(-1/3)$ are irrational. Fauzan proved that $\zeta(5)$ is irrational, using Hankel determinants [14]. Fasten [13], Firsching [15] and Romik [43] verified the irrationality statement in Lean. Romik's formalization takes the prime number theorem as an axiom. Fauzan's argument also gives an irrationality measure. Using its results, Giraudin proved $\mu(\zeta(5)) \le 171$ [20]. Lai combined this method with simultaneous approximations to show that 1, $\zeta(3)$ and $\zeta(4)$ are linearly independent over $\mathbb{Q}$ [28]. For L_{-3} the effective holonomy bounds of Calegari, Dimitrov and Tang give $\mu(L_{-3}) \le 24781$ [5, Thm. 5.1].

Determinant methods. OpenAI's argument [36] has a different shape. Instead of a sequence of linear forms it uses one determinant Δ_N of size $n = 48N$, whose entries are rational linear combinations of 1 and G. The entries are double integrals against the Cauchy kernel $1/(1 - ts)$. If G were rational, the denominators of the entries would force a lower bound

$|\Delta_N| \geq e^{-(\Lambda+o(1))N^2}$ along $N = p$ prime. On the other hand, Andréief's identity [17] writes Δ_N as an integral of a product of determinants, and an estimate for a discrete logarithmic energy gives $|\Delta_N| \leq e^{-(D+o(1))N^2}$. OpenAI proves $D > \Lambda$, with a margin of about 10^{-4} per unit of n^2, and proves that $\Delta_p \neq 0$ by reduction modulo p. Determinants of linear forms had been used for irrationality before, in Zudilin's criterion with Hankel determinants [55].

The first irrationality measure for G was obtained by Jha [22], who proved $\mu(G) \leq 10^8$. Carreira [8] treated the five integer parameters of the construction as free, stated conditions under which OpenAI's argument goes through, and proved $\mu(G) < 2294$. He also expressed the best constant of the energy estimate as the value of an equilibrium problem for two measures, and in [7] he records attempts to carry the method to $\beta(4)$. Krisman proved that one of $\beta(4), \beta(6), \ldots, \beta(36)$ is irrational, with a Lean formalization [26]. Bojnal [3] and Xu [50] obtained $\mu(G) < 2402$ and $\mu(G) \leq 727.03$ at other parameters. Xu's certificate is numerical and not formalized. Davis [9] reached $\mu(G) \leq 528.23288$ with new quadrature nodes, the zeros of $T_K(P(t))$ for a fixed polynomial P. The MultiAgency agent swarm regenerated P and reached $\mu(G) \leq 528.2315$ [34]. Both results are formalized in Lean. Our nodes are different, but we use two formal lemmas of [9] on the weights of nodal polynomials in Section 5. The other constructions are independent of the ideas of the present paper. In the third version of [8], Carreira bounds the generalized Vandermonde determinant by one term of the dual Cauchy identity, using the positivity of Schur polynomials. His bound carries the factor $\prod_l (1 + \tau s_l)^m$ with $\tau = 0.73$, of the same form as in Lemma 5.6, and he obtains $\mu(G) \leq 280.59$. His certificate is computer-assisted and not formalized. A preprint of Sun [46] claimed the irrationality of G, but Wachs [49] showed that its main estimate is not justified.

1.3. **The companion paper.** The companion paper [31] proves irrationality measures on three planes of constants. One Löwner structure produces all three. For all $r, r' \in \mathbb{Q}$, not both zero,

$$\mu(rL_{-3} + r'\pi^2) \leq 20.6622, \qquad \mu(r\zeta(3) + r'\pi^2) \leq 335.20,$$
$$\mu(r\zeta(3) + r'\pi^4) \leq 72.726.$$

The numbers on these planes were already known to be irrational, by unrelated methods. On the first plane the measure improves the published bound 24781. On the second it is, as far as we know, the first measure valid on the whole plane. On the third a better bound is known. At special points the bounds are smaller. For example, $\mu(L_{-3}) \leq 13.6839$ and $\mu(L_{-3} \pm \pi^2/9) \leq 9.4913$. The constructions use differences of values of the trigamma function ψ',

$$\Psi(s,t) = \psi'(s) - \psi'(t) = \sum_{n \geq 0} \Big(\frac{1}{(n+s)^2} - \frac{1}{(n+t)^2} \Big) \qquad (0 < s < t \leq 1).$$

Catalan's constant is one of them: $G = \frac{1}{16}\Psi(\frac{1}{4}, \frac{3}{4})$. The same structure reaches points off the planes. The companion paper proves measures for $G \pm 5\pi^2/48$, which are rational multiples of $\Psi(\frac{1}{4}, 1)$ and $\Psi(\frac{3}{4}, 1)$, and for $\Psi(\frac{k}{5}, 1)$ and $\Psi(\frac{k}{7}, 1)$. For these numbers they are the first proofs of irrationality that we know of. The determinants of the companion paper are Löwner matrices of digamma differences. They are much smaller than OpenAI's. That method also applies to $(\frac{1}{4}, \frac{3}{4})$, but there its denominators outgrow its decay, so it gives no measure for G. The present paper therefore refines OpenAI's construction. The two papers use the same criterion, the same names for the denominator exponent Λ, the decay rate D and the neighbourhood exponent ρ, and one formalization.

1.4. **What a measure requires.** Write $\Delta_N(X) = \det(\mathcal{A}_N^0 + X\mathcal{L}_N)$, so that OpenAI's determinant is $\Delta_N(G)$, and let Q_N be the integer polynomial obtained by clearing denominators. Suppose $|G - a/b|$ is small. Three properties of Q_N turn OpenAI's comparison into a contradiction.

(i) *Denominators.* The coefficients of Δ_N have a common denominator of size $e^{(\Lambda+o(1))N^2}$, and $\deg Q_N \le dN$.

(ii) *Smallness on a neighbourhood.* $|\Delta_N(x)| \le e^{-(D+o(1))N^2}$ for every x with $|x - G| \le e^{-\rho N}$, not only at $x = G$.

(iii) *Nonvanishing at the point.* $Q_p(a/b) \neq 0$ for a prime p of size about $\log b$.

OpenAI's proof gives nonvanishing only at $X = G$. We need it at every a/b with $p^2 \nmid b$, and we must find such a prime without losing a power of b. Section 2.4 shows that the squares of the larger primes can be recovered from the coefficients of Q_p.

1.5. **Outline of the proof.** Suppose that $|G - a/b| < b^{-\mu}$ with b large, and take a prime $p \approx \mu \log b/\rho$. Then $|a/b - G| \le e^{-\rho p}$, so by (ii) $|\Delta_p(a/b)| \le e^{-(D+o(1))p^2}$. By (i) the number $b^{dp}Q_p(a/b)$ is an integer, and by (iii) it is nonzero. Hence

$$1 \le \left|b^{dp}Q_p(a/b)\right| \le b^{dp}\, e^{(\Lambda - D + o(1))p^2}$$
$$= \exp\bigl(dp\log b - (\eta + o(1))p^2\bigr), \qquad \eta = D - \Lambda.$$

The right side is less than 1 as soon as $\mu > d\rho/\eta$. So $\mu(G) \le d\rho/\eta$ (Proposition 2.2 gives the precise form).

The three numbers come from different places. The degree d comes from the construction (Lemma 2.1). The neighbourhood exponent ρ comes from a quadrature formula for the coefficient of G (Section 5). The difference η is a small difference of two large numbers, $D \approx 6447$ and $\Lambda \approx 6393$ (Figure 1). For OpenAI's construction with our parameters, D and Λ differ by about 8. We lower Λ by 22.06 through the corank of a matrix modulo p (Section 4), and raise D by 24.64 through a bound for a determinant of row functions (Section 7). The result is $\eta \approx 54.69$, which is what makes $d\rho/\eta$ as small as 270.

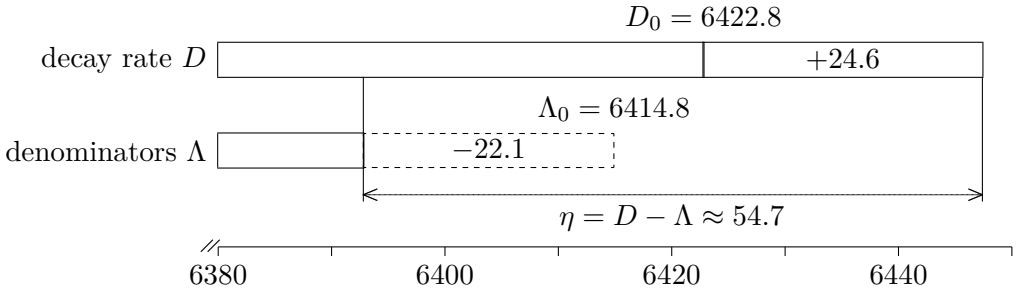


FIGURE 1. The exponent balance for the parameters $(58, 9, 3, 5, 3)$, per unit of N^2. The axis is broken at 6380. OpenAI's estimates give the decay rate D_0 and the denominator exponent Λ_0, which differ by about 8. The bound of Section 7 adds 24.6 to D_0, and the corank bounds of Section 4 remove 22.1 from Λ_0.

1.6. **New ideas.**

1. *Large corank modulo p.* Let $\nu_N(p)$ be the corank modulo p of p^2 times the matrix of rational parts. Each unit of corank lowers the denominator by a factor p (Proposition 3.3). We show that the corank is of order N for the primes in 113 intervals inside $(\frac{16}{7}N, \frac{53}{2}N]$. The kernel vectors are explicit on most of them. They come from antisymmetric combinations of rows, a reflection formula for Chebyshev polynomials modulo p, and the relation $(1-X)^p = 1-X^p$. Together these lower the denominator exponent by 22.06.

2. *One counting principle.* All these kernels arise in the same way. One takes a space of row combinations and imposes a few linear conditions that make every entry of the combined row vanish modulo p. Then the corank is at least the dimension minus the number of conditions (Theorem 4.1). The four closed formulas are instances. On the remaining intervals of primes we apply the principle to every constant subspace at once (Theorem A.1). Only upper bounds for ranks are then needed, and these come from forced intersections of row spaces. No value of a moment enters.

3. *A perturbation bound column by column.* For the neighbourhood of G we write the coefficient of G by a quadrature formula with $53N$ nodes. The nodes are Chebyshev nodes of 106 short intervals, and the Lagrange polynomials are bounded by the Green functions of the intervals. The perturbed determinant is then a sum of determinants in which some columns are replaced by values of the row polynomials at the nodes. We bound the weight of each replaced column and its generalized Vandermonde factor together. A weighted form of Hadamard's inequality with a parameter $\zeta = \frac{3}{4}$ makes the dependence on the column cancel, and one contour of radius $\frac{3}{4}$ treats all columns at once. The part of the weight that depends on the node is carried into the energy bound as an external field.

4. *Trial sequences that move with the configuration.* The energy bound with prescribed columns uses trial sequences that depend linearly on the

fraction of prescribed columns. Their fields are affine in the parameter, so the suprema are bounded by chords between two end points. A finite cover of the range of the fraction by cells then reduces every case to exact polynomial inequalities.

5. *A phase inequality for Pick matrices.* OpenAI's integrand contains a determinant $\det \mathsf{L}$ of size n that the energy method bounds by $e^{O(n \log n)}$. In the Pick matrix of the nodes the positive nodes carry the weight -5. We show that $\det \mathsf{L}$ loses only a factor 499^n when this weight is replaced by 0, as soon as one phase statistic of the nodes is large enough (Theorem 7.3). The proof interpolates the data of a singular Pick matrix by a finite Blaschke product and follows its argument along an arc of the unit circle.

6. *A complete log-gas bound.* With weight 0 the determinant factors through two Gram determinants. Fischer's inequality and a finite Andréief identity bound it by two logarithmic gases, one in the plane and one in the Joukowski coordinate. These are bounded by tangent inequalities against fixed reference measures. This raises the decay rate by 24.64. Numerical evaluation for small N suggests that this is most of the true size of $\det \mathsf{L}$ at the extremal configuration.

7. *Near-extremal configurations.* The phase statistic need not be large on every configuration. We show that it is large whenever the configuration comes close to the maximum of OpenAI's energy. On all other configurations OpenAI's bound already suffices.

We use the parameters $(58, 9, 3, 5, 3)$. The general corank theorems of Section 4 hold for all parameters, and Theorem 7.3 does not involve them.

1.7. **The constants.** Table 1 lists the constants for the parameters $(58, 9, 3, 5, 3)$, per unit of N or N^2. Without the corank gain and the bound

degree	d	53
cost of a prescribed column	$c = c_2 + \sigma$	3.93
weight of a column	ψ	50.11
neighbourhood exponent	$\rho = 58c + \psi$	278.37
decay rate (OpenAI's estimate)	D_0	6422.83
decay rate (with the bound for $\det \mathsf{L}$)	D	6447.47
denominator exponent (with the corank gain)	Λ	6392.78
difference	$\eta = D - \Lambda$	54.68
bound	$d\rho/\eta$	269.78 . . .

TABLE 1. The constants of the proof, truncated. The formal proof uses exact rational numbers.

for $\det \mathsf{L}$, η would be about 8, and the method would give an exponent above 1800.

1.8. **Organization.** Section 2 recalls OpenAI's construction for general parameters and proves the criterion that turns the three properties of §1.4 into a measure. Section 3 bounds the denominators and proves nonvanishing at every rational point. Section 4 proves the corank bounds, which give the arithmetic gain. Sections 5–7 treat the analytic side. Section 5 extends OpenAI's estimate to a neighbourhood of G. Section 6 recalls the energy method with prescribed columns, and Section 7 proves the bound for the row determinant. Section 8 assembles the proof, and Section 9 discusses the method. The certified numerical inequalities are collected in the appendices, and Section 10 describes the formalization.

1.9. **Notation.** We use the notation of OpenAI's paper [36] for the construction and refer to it for details. Citations of the form [36, Prop. 2.1] point to its numbered statements. We write log for the natural logarithm, v_p for the p-adic valuation and $\vartheta(x) = \sum_{p \le x} \log p$ for Chebyshev's function. T_k and U_k are the Chebyshev polynomials of the first and second kind, with $U_{-1} = 0$, $T_{-k} = T_k$ and $U_{-k-1} = -U_{k-1}$. We write $[X^k]F$ or $F[k]$ for the coefficient of X^k in a polynomial F, $y_+ = \max(y, 0)$, and $\operatorname{rev}_E F(X) = X^E F(1/X)$ for the reversal of a polynomial of degree at most E.

2. The construction and the criterion

2.1. **Designs.** OpenAI's determinant depends on five integer parameters. Following Carreira [8] we call a quintuple $(n_0, b_0, q_0, g_0, h_0)$ of integers with $n_0, b_0, q_0, h_0 \ge 1$, $g_0 \ge 0$ and $2g_0 < n_0$ a *design*. For $N \ge 1$ we write

$$n = n_0 N, \quad b = b_0 N, \quad q = q_0 N, \quad g = g_0 N, \quad h = h_0 N,$$
$$L = n + b + q, \quad C = L + g, \quad H = C + h, \quad A = b + q + 2g,$$

and $d_0 = n_0 - g_0$. OpenAI's choice is $(48, 7, 4, 4, 2)$. We work throughout with the design $(58, 9, 3, 5, 3)$, for which

$$n = 58N, \quad b = 9N, \quad q = 3N, \quad g = 5N, \quad h = 3N,$$
$$L = 70N, \quad C = 75N, \quad H = 78N, \quad A = 22N,$$

and $d_0 = 53$. The search that led to this design is not part of the proof.

2.2. **The determinant.** For $0 \le r < n$ the row polynomials are

$$P_r(t) = (1-t)^h t^{C-1} T_{|r-g|}(1/t), \qquad \deg P_r < H, \tag{2.1}$$

together with companion polynomials D_r built from U_{r-g-1} [36, (3)]. Thus P_r depends only on $|r-g|$. Put $f(t) = \sqrt{1-t^2}$ and $dm(t) = |t|\,dt/f(t)$ on $(-1, 1)$, and

$$A_r(t) = P_r(t) - \frac{3}{2}\,\mathbf{1}_{\{t>0\}}\,\frac{f(t)}{t} D_r(t), \qquad \varphi_k(s) = s^{b+k}(1-s)^q.$$

The determinant studied by OpenAI is that of the $n \times n$ matrix

$$\mathcal{A}_N(r,k) = \int_{-1}^{1} \int_{0}^{1} A_r(t)\,\frac{\varphi_k(s)}{1-ts}\,ds\,dm(t) \qquad (0 \le r, k < n). \tag{2.2}$$

The factor $(1-s)^q$ in φ_k amounts to a finite difference of order q applied to the columns. The entries are rational linear combinations of 1 and G. There are rational matrices $\mathcal{A}_N^0$ and $\mathcal{L}_N$ with

(2.3)
$$\mathcal{A}_N = \mathcal{A}_N^0 + G\,\mathcal{L}_N, \qquad \mathcal{L}_N(r,k) = \sum_{v=0}^{q} (-1)^v \binom{q}{v} \sum_i [t^i] P_r \cdot 4\, c_{(i-b-k-v)/2},$$

where $c_l = 4^{-l}\binom{2l}{l}$ for integers $l \ge 0$ and $c_z = 0$ otherwise [36, Prop. 2.1 and (20)]. The rational part is

(2.4)
$$\begin{aligned} \mathcal{A}_N^0(r,k) &= \sum_{v=0}^{q} (-1)^v \binom{q}{v} R(r, b+k+v), \\ R(r,j) &= \sum_{i<H} P_r[i]\, M(i,j) - \frac{3}{2} \sum_{i<H} D_r[i]\, Z(i,j) \end{aligned}$$

with rational numbers $M(i,j)$, the moments of m against the Cauchy kernel, and $Z(i,j)$, where $Z(i,i) = -\sum_{k\le i} k^{-2}$ and $Z(i,j) = (H_i - H_j)/(i-j)$ for $i \ne j$, $H_i = \sum_{k\le i} 1/k$ [36, Prop. 2.1]. We call the vectors $R(\cdot, j)$, $b \le j < L$, the *moment columns*. The columns of $\mathcal{A}_N^0$ are their differences of order q. The design enters (2.4) only through the integer coefficients $P_r[i], D_r[i]$ and the range of j. Put $\Delta_N(X) = \det(\mathcal{A}_N^0 + X\mathcal{L}_N) \in \mathbb{Q}[X]$, so that $\Delta_N(G) = \det \mathcal{A}_N$.

2.3. **The degree.**

Lemma 2.1. $\deg \Delta_N \le n - g$.

Proof. By (2.3) the row r of $\mathcal{L}_N$ depends only on P_r, and hence only on $|r-g|$. So the rows r and $2g-r$ of $\mathcal{L}_N$ agree for $0 \le r < g$, and $2g-r \le n-1$ because $2g < n$. Subtracting row r from row $2g-r$ for each $r < g$ does not change the determinant and makes g rows of the new coefficient of X vanish. A determinant $\det(B + XL')$ in which L' has $n-g$ nonzero rows has degree at most $n-g$ in X. □

2.4. **The criterion.**

The following criterion turns the three properties of §1.4 into a measure. It is the reason for every later estimate, so we state and prove it first.

Proposition 2.2. *Let $\xi \in \mathbb{R}$, $d \in \mathbb{N}$, and let $\eta, \rho, \lambda, \delta, \omega > 0$ satisfy*

$$\omega < \lambda, \quad \delta < \tfrac{1}{3}, \quad d \le (\lambda - \omega)\big(\eta + 2d(1-3\delta)\big),$$
$$1 < 2\big(\omega - \delta(2\lambda - \omega)\big), \quad \eta < 2d, \quad d + 4d\delta\lambda < \eta\lambda.$$

For each prime p let $Q_p \in \mathbb{Z}[X]$, and suppose that the following hold for all large p.

- *(a)* $\deg Q_p \le dp$.
- *(b)* $Q_p(a/b) \ne 0$ *whenever* $a \in \mathbb{Z}$, $b \ge 1$ *and* $p^2 \nmid b$.
- *(c) For all primes $p < p' \le 2p$ and all $i \le dp$, p'^{2i} divides the i-th coefficient of Q_p.*

(d) $|Q_p(x)| \le e^{-\eta p^2}$ *for* $|x-\xi| \le e^{-\rho p}$.

Then $|\xi - a/b| > b^{-\rho\lambda}$ *for all integers* a *and all sufficiently large* b.

Proof. Let a/b be a rational number with b large. Put $u_1 = \lambda \log b$ and $u_0 = (\lambda - \omega)\log b$, and suppose $|\xi - a/b| \le b^{-\rho\lambda}$. By the prime number theorem, $(1-\delta)t \le \vartheta(t) \le (1+\delta)t$ for all large t, and $u_0 \to \infty$ with b. If every prime $p \in (u_0, u_1]$ had $p^2 \mid b$, then $2(\vartheta(u_1) - \vartheta(u_0)) \le \log b$. But

$$\begin{aligned} 2\big(\vartheta(u_1) - \vartheta(u_0)\big) &\ge 2\log b\,\big((1-\delta)\lambda - (1+\delta)(\lambda-\omega)\big) \\ &= 2\log b\,\big(\omega - \delta(2\lambda - \omega)\big) > \log b \end{aligned}$$

by the fourth hypothesis. So there are primes in $(u_0, u_1]$ whose squares do not divide b. Let p be the *largest* one. Then every prime $p' \in (p, u_1]$ has $p'^2 \mid b$.

Let $T \le \min(u_1, 2p)$ with $T \ge p$, and let Π_T be the product of the p'^2 over the primes $p' \in (p, T]$. Then $\Pi_T \mid b$ and $\log \Pi_T = 2(\vartheta(T) - \vartheta(p))$. Consider the integer $b^{dp}Q_p(a/b) = \sum_i c_i a^i b^{dp-i}$. Each of its terms is divisible by p'^{2dp} for each of these p'. Indeed p'^{2i} divides c_i by (c), and $p'^2 \mid b$ gives $p'^{2(dp-i)} \mid b^{dp-i}$. So Π_T^{dp} divides the integer. Since $p^2 \nmid b$, $Q_p(a/b) \ne 0$ by (b). So $b^{dp}Q_p(a/b)/\Pi_T^{dp}$ is a nonzero integer, and

$$1 \le \Big(\frac{b}{\Pi_T}\Big)^{dp} |Q_p(a/b)|. \tag{2.5}$$

Since $\rho p \le \rho u_1 = \rho\lambda \log b$, we have $|a/b - \xi| \le e^{-\rho p}$. So $|Q_p(a/b)| \le e^{-\eta p^2}$, and (2.5) gives

$$\eta p \le d\big(\log b - 2(\vartheta(T) - \vartheta(p))\big). \tag{2.6}$$

Suppose first that $u_1 \le 2p$, and take $T = u_1$. With $\vartheta(u_1) - \vartheta(p) \ge u_1 - p - 2\delta u_1$, (2.6) reads

$$\eta p \le d(\log b - 2u_1 + 2p + 4\delta u_1).$$

Hence $(\eta - 2d)p \le d\log b - 2du_1 + 4d\delta u_1$. As $\eta < 2d$ and $p \le u_1$, the left side is at least $(\eta - 2d)u_1$. So $\eta u_1 \le d\log b + 4d\delta u_1$, that is, $\eta\lambda \le d + 4d\delta\lambda$. This contradicts the last hypothesis.

Suppose now that $2p < u_1$, and take $T = 2p$. Then $\vartheta(2p) - \vartheta(p) \ge (1-3\delta)p$, so (2.6) gives $p(\eta + 2d(1-3\delta)) \le d\log b$. Since $p > u_0 = (\lambda - \omega)\log b$, this gives $(\lambda - \omega)(\eta + 2d(1-3\delta)) < d$, which contradicts the third hypothesis. □

The hypotheses have the following roles. The fourth ensures that the interval $(u_0, u_1]$ contains a prime whose square does not divide b. The last makes the final inequality contradictory when all larger primes of the interval can be recovered. The proof takes the largest prime of the interval whose square does not divide b. The squares of the larger primes then divide b, and (c) recovers those below $2p$. Choosing any prime of the interval would lose a power of b. The chosen prime may be so small that $u_1 > 2p$, and then (c) recovers only the primes up to $2p$. The third hypothesis treats this case. A small prime p makes the decay $e^{-\eta p^2}$ weak, but it also makes the height b^{dp}

small, and the primes of $(p, 2p]$ alone already win. So λ need not exceed 2ω, and the bound $\rho\lambda$ may lie below ρ.

For Catalan's constant we take $d = 53$, $\omega = \frac{3}{5}$, $\delta = 2.9 \cdot 10^{-7}$ and $\lambda = 270/\rho$. The last hypothesis then reads $270 > \mu_*$ with $\mu_* = d\rho/(\eta - 4d\delta) = 269.78696\ldots$ The formal proof checks the six hypotheses with $\rho\lambda$ in place of 270 for every $\rho\lambda \in [269.787, 400]$.

3. Denominators and nonvanishing

This section proves the arithmetic properties of Δ_N: the denominators of its coefficients and its nonvanishing at rational points. Carreira [8] checked that OpenAI's argument goes through for parameters satisfying a few conditions. We need more. The entries must be controlled modulo p^2, since the corank bounds of Section 4 work there. And nonvanishing is needed at every rational point a/b with $p^2 \nmid b$, not only at G.

3.1. **Corank and valuation.** We use repeatedly the following elementary fact.

Lemma 3.1. *Let M be an $n \times n$ matrix with entries in $\mathbb{Z}_{(p)}$ whose reduction modulo p has rank at most $n - e$. Then $v_p(\det M) \geq e$.*

Proof. Factor the reduction through $\mathbb{F}_p^{n-e}$ and lift the two factors to integral matrices U, V, so that $M = UV + pE$ with E integral at p. Expand the determinant multilinearly in the columns. Every term that uses more than $n - e$ columns of UV vanishes, and every other term has at least e factors p. □

3.2. **The digit formula.** The reduction of $p^2\mathcal{A}_N^0$ modulo p is described by the base-p digits of the coefficients of the row polynomials. This is the one place where the arithmetic of the moments enters, and it does not depend on the parameters. For an odd prime p and $c \in \mathbb{F}_p^n$ put

$$P_c = \sum_r c_r P_r, \qquad D_c = \sum_r c_r D_r, \qquad W = (1 - X^2)^{(p-1)/2},$$

reading the P_r, D_r as polynomials in X.

Lemma 3.2 (digit formula)**.** *Let p be an odd prime with $H < p^2$, let $0 \leq j < H$, and write $j = Jp + \ell$ with $0 \leq \ell < p$. Then $p^2 R(r, j)$ is p-integral for every row r, and for every $c \in \mathbb{Z}^n$*

$$p^2 \sum_r c_r R(r, j) \equiv \sum_{u \geq 0} \Big(\alpha_{u,J}\, (XWP_c)[up + \ell] + \beta_{u,J}\, D_c[up + \ell] \Big) \pmod{p},$$

where $\alpha_{u,J} = M(u - 1, J)$ and $\beta_{u,J} = -\frac{3}{2} Z(u, J)$ are p-integral numbers that depend neither on c nor on ℓ.

Proof. Two facts about the numbers M and Z are needed, for $0 \leq i, j < p^2$. Write $j = Jp + \ell$, let $\varrho \in [0, p)$ be the residue of $j - i - 1$ modulo p, and put $u = (i + 1 + \varrho - \ell)/p$.

First, $p^2 M(i,j)$ is p-integral and $p^2 M(i,j) \equiv W[\varrho]\, M(u-1,J) \pmod p$. This is the digit reduction of the moments in [36, §3]. It is obtained from the p-adic analysis of their recursions and does not involve the parameters.

Second, $p^2 Z(i,j)$ is p-integral. It is congruent to $Z(\lfloor i/p \rfloor, J)$ if $i \equiv j \pmod p$ and to 0 otherwise. To see this, write $i' = \lfloor i/p \rfloor$. Since $i < p^2$, we have $H_i = p^{-1} H_{i'} + S_i$ and $H_i^{(2)} = p^{-2} H_{i'}^{(2)} + S_i'$ with S_i, S_i' p-integral, which gives the case $i = j$. If $i \not\equiv j$, then $p^2 Z(i,j) = p \cdot pZ(i,j) \equiv 0$. If $i \equiv j$ and $i \neq j$, then $i - j = (i' - J)p$. The integers $k \in (j, i]$ prime to p form $i' - J$ complete systems of residues modulo p, so $\sum_k 1/k \equiv (i' - J) \sum_{r=1}^{p-1} 1/r \equiv 0 \pmod p$. Hence $H_i - H_j \equiv (H_{i'} - H_J)/p$ and $p^2 Z(i,j) \equiv Z(i', J)$.

Now use (2.4). By the first fact, the part $\sum_i F[i]\, p^2 M(i,j)$ with $F = P_c$ is $\sum_i F[i]\, W[\varrho(i)]\, M(u(i)-1, J)$. For fixed u the indices with $u(i) = u$ are those for which $\varrho = up + \ell - 1 - i$ lies in $[0, p-1]$. Since W has degree $p-1$, the sum over these i is $(XFW)[up + \ell]$. This gives the α-terms. By the second fact the part with D_c is $\sum_{i \equiv \ell} D_c[i]\, Z(\lfloor i/p \rfloor, J) = \sum_u Z(u,J) D_c[up + \ell]$, which gives the β-terms. The p-integrality of $M(u-1,J)$ and $Z(u,J)$ for $u, J < p$ is part of the two facts. □

For a row combination $c \in \mathbb{F}_p^n$ and a moment column $j \in [b, L)$ let $R_c(j)$ be the reduction modulo p of $p^2 \sum_r c_r R(r,j)$. We call it the *value of c at the moment column j*. The columns of $\mathcal{A}_N^0$ are differences of order q of the moment columns, so c lies in the left kernel of the reduction as soon as the corresponding differences of the $R_c(j)$ vanish.

We use the digit formula in the following form. Let $0 \le k < n$, $s = b + k$, $J = \lfloor s/p \rfloor$ and $\varsigma = (s \bmod p) + q$. If $\varsigma < p$, the $q+1$ moment columns $s, \dots, s+q$ that enter column k have the same digit J and residues $(s \bmod p) + v < p$. Since $\sum_{v=0}^{q} (-1)^v \binom{q}{v} F[t+v] = [X^{t+q}]((X-1)^q F)$ for every polynomial F, Lemma 3.2 gives

$$p^2 \sum_r c_r \mathcal{A}_N^0(r,k) \equiv \sum_{u \ge 0} \Big(\alpha_{u,J}\, Q_c[up + \varsigma] + \beta_{u,J}\, \widehat{D}_c[up + \varsigma] \Big), \tag{3.1}$$

where

$$Q_c = (X-1)^q\, XW P_c, \qquad \widehat{D}_c = (X-1)^q D_c. \tag{3.2}$$

3.3. **The valuation at p.** For an odd prime p with $H < p^2$ the entries of $p^2 \mathcal{A}_N^0$ are p-integral by Lemma 3.2. Let $\mathcal{A}_{N,p}^0$ be its reduction modulo p. Let $\nu_N(p)$ be the corank of $\mathcal{A}_{N,p}^0$, that is, n minus its rank. The corank enters through the following statement.

Proposition 3.3. *Let p be an odd prime with $H < p^2$ and let $z \in \mathbb{Q}$ have denominator prime to p. Then $p^2(\mathcal{A}_N^0 + z\mathcal{L}_N)$ has p-integral entries, its reduction modulo p equals $\mathcal{A}_{N,p}^0$, and*

$$v_p(\Delta_N(z)) \ge \nu_N(p) - 2n.$$

Proof. The entries of $p^2 \mathcal{A}_N^0$ are p-integral, since each column of $\mathcal{A}_N^0$ is an integer combination of moment columns with $j < L \le H$. The entries of

$\mathcal{L}_N$ are p-integral for odd p, since $P_r[i] \in \mathbb{Z}$ and $c_l \in \mathbb{Z}[1/2]$. So $p^2\mathcal{L}_N \equiv 0 \pmod{p^2}$. Hence, for z with denominator prime to p, $p^2(\mathcal{A}_N^0 + z\mathcal{L}_N)$ has p-integral entries with the same reduction as $p^2\mathcal{A}_N^0$. By Lemma 3.1 its determinant has valuation at least $\nu_N(p)$, and $\det(p^2(\mathcal{A}_N^0 + z\mathcal{L}_N)) = p^{2n}\Delta_N(z)$. □

3.4. **Pole counts.** Let p be an odd prime with $\sqrt{H} < p \le H$. OpenAI's method [36, §3] bounds the valuation of the entries prime by prime. For any choice of the parameters it gives $v_p(\Delta_N(z)) \ge -\ell_N(p)$ for every z with denominator prime to p. Here $\ell_N(p)$ is the number of columns whose entries can have a pole at p. It is computed by grouping the coefficients of the row polynomials by their base-p digits. All differences below are truncated at 0. For $2p \le H$,

$$\ell_N(p) = 2n - \Big(\min(p,\, L-p,\, A) - \max(b,\, H-2p) - q\Big).$$

For $H < 2p \le 2H$, with $R = (L-b) - (\min(A, L-p) - b) - (\min(p, L) - \max(b, H-p))$, $P' = \min(A, L-p) - b$ and $E' = \min(b, H-p) - (L-p)$,

$$\ell_N(p) = \min(2n,\ n+R,\ 2R + P' + E').$$

These are the exponents of the valuation theorems of the formalization, which hold for all parameters. For the primes that interest us they agree with OpenAI's. For our parameters $\ell_N(p) = Nf(p/N)$ up to integer rounding, with a piecewise linear function f of $\theta = p/N$:

$$f(\theta) = \begin{cases} 116 & \theta \le 59/2, \\ 175 - 2\theta & 59/2 \le \theta \le 69/2, \\ 106 & 69/2 \le \theta \le 39, \\ 184 - 2\theta & 39 \le \theta \le 81/2, \\ 265 - 4\theta & 81/2 \le \theta \le 48, \\ 217 - 3\theta & 48 \le \theta \le 69, \\ 148 - 2\theta & 69 \le \theta \le 70, \\ 78 - \theta & 70 \le \theta \le 78, \end{cases}$$

and $f = 0$ for $\theta \ge 78$, since for $p > H$ the matrix $\mathcal{A}_N^0 + z\mathcal{L}_N$ is p-integral. Its integral over $[1, 78]$ is $24269/4$, and $f = 116 = 2n_0$ on $[0, 1]$. Since $f(\theta) = 2n_0$ for $\theta \le 59/2$, every prime of the range $(\frac{16}{7}N, \frac{53}{2}N]$ of the corank bounds of Section 4 has $\ell_N(p) = 2n$. There Proposition 3.3 improves the bound $-2n$ to $\nu_N(p) - 2n$.

The primes $p \le \sqrt{H}$ are not covered by these bounds. For them we use the cruder bound $6n \log_p H$ for the exponent, which contributes $O(\sqrt{N}\, N \log N) = o(N^2)$ in total.

3.5. **The prime 2.** At $p = 2$ the entries are not integral, and the bound comes from the expansion of the determinant into minors, as in [36, §3]. Its exponent is governed by a quadratic form in the numbers m, ℓ of columns of two exceptional types,

$$\tfrac{3}{2}m^2 + \tfrac{1}{2}\ell^2 - (H + C - b)\, m - C\,\ell + Cn.$$

The following identity shows that it is at least $-\Lambda_2$ whenever $m, \ell \geq 0$, $m + \ell \leq n$ and $C \geq n$, where

$$\Lambda_2 = \frac{(n + H - b)^2}{8} - \frac{n^2}{2}. \tag{3.3}$$

With $s = n - m - \ell$ and $m_* = (n + H - b)/4$,

$$\tfrac{3}{2}m^2+\tfrac{1}{2}\ell^2-(H+C-b)m-C\ell+Cn+\Lambda_2 = 2(m-m_*)^2+s\,(C-n+m)+\tfrac{1}{2}s^2 \geq 0.$$

This gives a 2-adic bound $v_2(\Delta_N(i)) \geq -\Lambda_2 - O(N \log N)$, uniform in the integer i. For our parameters $\Lambda_2 = (58 + 78 - 9)^2/8 - 58^2/2 = 2673/8$ in units of N^2.

3.6. **The common denominator.** The corank bounds of Section 4 give a function κ of $\theta = p/N$ (Theorem 4.10). We use it here as an input.

Proposition 3.4. *For the parameters* $(58, 9, 3, 5, 3)$ *there are positive integers* $\mathcal{D}_N$ *with the following properties.* $\mathcal{D}_N\Delta_N(i) \in \mathbb{Z}$ *for every integer* i*, and* $p'^{2n} \mid \mathcal{D}_N$ *for every prime* $N < p' \leq 2N$*. For every* $\varepsilon > 0$ *and all large* N

$$\log \mathcal{D}_N \leq (\Lambda + \varepsilon)N^2,$$

$$\Lambda = \frac{2673}{8}\log 2 + 116 + \frac{24269}{4} - \int_0^\infty \kappa(\theta)\, d\theta = 6392.7870\ldots$$

Here $\frac{2673}{8}\log 2$ comes from the prime 2, and $116 + \frac{24269}{4}$ is the area of f. The last term is the gain from the corank.

Proof. For primes $p \leq H$ with $H < p^2$, let $\nu_N^*(p)$ be the largest of the lower bounds for $\nu_N(p)$ of Section 4 that apply at p, and 0 if none applies. Put $e_N(p) = \min(\ell_N(p),\, 2n - \nu_N^*(p))$. For the primes with $p^2 \leq H$ use the crude bound above, and for $p = 2$ the 2-adic bound. Let $\mathcal{D}_N$ be the product of $p^{e_N(p)}$ over $p \leq H$. By Proposition 3.3 and the bounds of this section, $v_p(\Delta_N(i)) \geq -e_N(p)$ for every integer i and every p. For $p > H$ the matrix is p-integral. So $\mathcal{D}_N\Delta_N(i) \in \mathbb{Z}$. A prime $p' \in (N, 2N]$ has $\nu_N^*(p') = 0$, since no interval of Theorem 4.10 contains it, and $\ell_N(p') = 2n$. So $p'^{2n} \mid \mathcal{D}_N$. Here $N \geq 79$ guarantees $p' > \sqrt{H}$.

For the size of $\mathcal{D}_N$ we use the prime number theorem in the weighted form

$$\frac{1}{N}\sum_p F(p/N)\log p \longrightarrow \int_0^\infty F(\theta)\, d\theta \qquad (N \to \infty), \tag{3.4}$$

for bounded, Riemann integrable, compactly supported F. By (3.4) the sum of $\ell_N(p)\log p$ over the odd primes is $N^2\int_0^{78} f + o(N^2)$, with $\int_0^{78} f = 116 + 24269/4$. The primes in the intervals of Theorem 4.10 lie in $(16N/7, 53N/2]$

and have $\ell_N(p) = 116N$. So their exponent is at most $116N - N\kappa(p/N) + 2$. The sum of $(N\kappa(p/N) - 2)\log p$ over these primes is $N^2 \int \kappa + o(N^2)$, again by (3.4). The 2-adic part contributes $\Lambda_2 \log 2\, N^2 + o(N^2)$. □

The polynomial Δ_N has degree at most $n - g = d_0 N$ and takes values in $\mathcal{D}_N^{-1}\mathbb{Z}$ at the integers. An integer-valued polynomial of degree m has coefficients in $\frac{1}{m!}\mathbb{Z}$, by Newton's interpolation formula. So

$$Q_N = (d_0 N)!\, \mathcal{D}_N\, \Delta_N(X) \in \mathbb{Z}[X], \qquad \deg Q_N \le d_0 N, \tag{3.5}$$

and $\log(d_0 N)! = o(N^2)$. In this way the gain from the corank enters through the integer polynomial itself, not through a separate common factor of its coefficients.

The criterion needs the squares of the primes in $(N, 2N]$ to divide the coefficients of Q_N in a graded way.

Lemma 3.5. *For all large N and every prime $N < p' \le 2N$, the j-th coefficient of Q_N is divisible by p'^{2j}.*

Proof. For these primes the matrices $p'^2\mathcal{A}_N^0$ and $\mathcal{L}_N$ are p'-integral by Lemma 3.2, and p'^{2n} divides $\mathcal{D}_N$. The identity

$$Q_N(X/p'^2) = (d_0 N)!\, \frac{\mathcal{D}_N}{p'^{2n}} \det(p'^2\mathcal{A}_N^0 + X\mathcal{L}_N)$$

shows that $Q_N(X/p'^2)$ has p'-integral coefficients. Its j-th coefficient is $c_{N,j}/p'^{2j}$, where $c_{N,j}$ is the j-th coefficient of Q_N. □

3.7. **Nonvanishing.**

Proposition 3.6. *Suppose that the three $n_0 \times n_0$ matrices B_σ, $\sigma \in \{0, 1, -1\}$, defined below, are invertible. Then for all large primes p and every rational z such that pz has denominator prime to p, $\Delta_p(z) \ne 0$. In particular $\Delta_p(a/b) \ne 0$ whenever $a \in \mathbb{Z}$, $b \ge 1$ and $p^2 \nmid b$.*

The matrices are those of OpenAI's reduction modulo p at $N = p$, with general parameters in place of OpenAI's. Let B be the rational $(n_0 + 1) \times n_0$ matrix of the entries of $\mathcal{A}_1^0$, the rational part at $N = 1$, for the rows $0, \dots, n_0$. The row n_0 is an auxiliary row. Let B_σ have rows $B_r + \sigma B_{r+1}$, $0 \le r < n_0$. For our parameters B has 59 rows and 58 columns.

Proof. Let $p > 4H_0$, with $H_0 = H/N$, so that $H = H_0 p < p^2/4$. Let p also be prime to the numerators and denominators of the entries of B and of the three determinants. In the construction at $N = p$ write the row index $r = pr_0 + i$ with $0 \le r_0 < n_0$ and $0 \le i < p$. Since $q = q_0 p$, $C = C_0 p$, $h = h_0 p$ and $g = g_0 p$, the relation $(1 - t)^p \equiv 1 - t^p$ expresses the row polynomials modulo p through those at $N = 1$ and the successor row $r_0 + 1$. This is where the auxiliary row enters. It is valid for rows on both sides of g, because the polynomials are defined through T_{r-g} and U_{r-g-1} for signed indices. The column functions satisfy $s^{b_0 p + k}(1 - s)^{q_0 p} = s^\ell (s^p)^{b_0 + k_0}(1 - s^p)^{q_0}$ in $\mathbb{F}_p[s]$. So the columns of one residue class use the moment columns

$b_0 + k_0, \dots, b_0 + k_0 + q_0$ at $N = 1$ and never change the residue. By [36, Lemma 3.1 and eqs. (31), (35), (36)], which need only $H < p^2/4$, the reduction of $p^2\mathcal{A}_p^0$ modulo p is then, after ordering the rows by (i, r_0) and the columns by (ℓ, k_0), the matrix $a^{\mathsf{T}} \otimes B_0 + b^{\mathsf{T}} \otimes B_1$. Here a, b are the fixed $p \times p$ matrices of OpenAI's identity (51), with $\det a = 2^{-p(p-1)/2}$. The permutation matrix in this decomposition has the eigenvalue 0 once and ± 1 each $(p-1)/2$ times. Therefore

$$\det = (\det a)^{n_0} \det B_0 \det(B_0 + B_1)^{(p-1)/2} \det(B_0 - B_1)^{(p-1)/2} \neq 0$$

in $\mathbb{F}_p$. Hence $p^{2n}\Delta_p(0)$ is a p-adic unit. Let z be such that pz has denominator prime to p. The entries of $\mathcal{L}_p$ are p-integral, so the difference $p^2 z\mathcal{L}_p = p\,(pz)\,\mathcal{L}_p$ is p-integral with zero reduction. So the reduction of $p^2(\mathcal{A}_p^0 + z\mathcal{L}_p)$ is the same invertible matrix, and $v_p(\Delta_p(z)) = -2n$. Finally, if $p^2 \nmid b$, then $p(a/b)$ has denominator prime to p. □

For the parameters $(58, 9, 3, 5, 3)$ we verify that B_0, $B_0 + B_1$ and $B_0 - B_1$ are invertible by exhibiting their inverses modulo 1009. This prime divides no denominator of B, so the three determinants are nonzero rational numbers. Proposition 3.6 then applies to all large primes. The proof uses only $2g < n$ and $H < p^2/4$.

4. Corank modulo p

This section proves the gain in the denominators. The principle is that of the arithmetic method of Hata and Rhin–Viola, adapted to a determinant. If the reduction modulo p of $p^2\mathcal{A}_N^0$ has corank e, then $\Delta_N(z)$ gains a factor p^e at every z with denominator prime to p (Proposition 3.3). So we need explicit row combinations $c \in \mathbb{F}_p^n$ with $c^{\mathsf{T}}\mathcal{A}_{N,p}^0 = 0$.

Here is the heuristic. Modulo p the Chebyshev polynomials of index $kp+d$ factor through X^p, so rows placed symmetrically about the centres $g + kp$ are related. Suitable combinations of such rows have row polynomials with long runs of vanishing base-p digits. By the digit formula their values at most moment columns then vanish, and the few remaining values are removed by imposing linear conditions. Each surviving dimension is a unit of corank.

We first prove the counting inequality (Theorem 4.1). It says that a space of row combinations subject to a few linear conditions lies in the left kernel. We then derive four closed formulas from it, which give most of the gain. Finally we describe the further families and the resulting function κ. Three facts about characteristic p are used throughout: the Chebyshev identities

$$\begin{aligned} T_p = X^p, \qquad & U_{p-1} = (X^2-1)^{(p-1)/2}, \\ T_{kp} = T_k(X^p), \qquad & U_{kp-1} = (X^2-1)^{(p-1)/2}\, U_{k-1}(X^p) \end{aligned} \tag{4.1}$$

in $\mathbb{F}_p[X]$, the relation $(1-X)^p = 1 - X^p$, and the digit formula (3.1). The last two identities of (4.1) follow from $T_{kp} = T_k \circ T_p$ and $U_{kp-1} = (U_{k-1} \circ T_p)\, U_{p-1}$.

4.1. **Setting.** Fix parameters, an integer $N \geq 1$ and an odd prime p with $H < p^2$. For $0 \leq r < n$ put $m_r = r - g$, so that $-g \leq m_r \leq n-1-g$ and $n-1-g \geq g$. In the notation of (2.1),

$$P_r = (1-X)^h \operatorname{rev}_{C-1} T_{m_r}, \qquad D_r = (1-X)^h \operatorname{rev}_{C-1} U_{m_r-1},$$

where $T_{-m} = T_m$ and $U_{-m-1} = -U_{m-1}$, so that $D_g = 0$. For $c \in \mathbb{F}_p^n$ write $S_c = \sum_r c_r \operatorname{rev}_{C-1} T_{m_r}$ and $U_c = \sum_r c_r \operatorname{rev}_{C-1} U_{m_r-1}$, so that $P_c = (1-X)^h S_c$ and $D_c = (1-X)^h U_c$. The lowest degree of $\operatorname{rev}_{C-1} T_m$ is at least $C - 1 - |m| \geq A = b + q + 2g$, and that of $\operatorname{rev}_{C-1} U_{m-1}$ is at least $A+1$. Put $\chi = g + q + h$. Then the polynomials Q_c and $\widehat{D}_c$ of (3.2) have their coefficients in $[A+1,\, L+\chi+p-1]$ and $[A+1,\, L+\chi-1]$ respectively.

A vector c lies in the left kernel of $\mathcal{A}^0_{N,p}$ as soon as the right side of (3.1) vanishes for every column k. We only use columns whose $q+1$ moment columns share the same base-p digit.

4.2. **The counting inequality.** All the kernels of this section arise in the same way. A space of row combinations is chosen, a few linear conditions are imposed, and the conditions make every value vanish. The dimension of the space minus the number of conditions is then a lower bound for the corank. The following theorem states this once. It also allows $q \geq p$.

Theorem 4.1. *Let $N \geq 1$ and let p be an odd prime with $H < p^2$. Write $q = ap + r$ with $0 \leq r < p$. Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{F}_p$, $\iota\colon \mathcal{V} \to \mathbb{F}_p^n$ an injective linear map and Φ a linear map on $\mathcal{V}$. Suppose that for every $x \in \ker \Phi$ there is a polynomial $g_x \in \mathbb{F}_p[X]$ of degree less than r with*

$$\sum_{v=0}^{a} (-1)^{a-v} \binom{a}{v} R_{\iota(x)}(j + vp) = g_x(j) \qquad (b \leq j < L - ap).$$

Then $\nu_N(p) \geq \dim \mathcal{V} - \operatorname{rank} \Phi$.

For $q < p$ the hypothesis says that the conditions Φ force all values at the moment columns to vanish.

Proof. By (2.4), the entry of $c^{\mathsf{T}} \mathcal{A}^0_{N,p}$ at column k is $\sum_{v=0}^{q} (-1)^v \binom{q}{v} R_c(b+k+v)$. This is the polynomial $(1-T)^q$ applied to R_c, where T acts as the shift. Modulo p we have $(1-T)^p = 1 - T^p$, hence

$$(1-T)^q = (1-T)^r (1-T^p)^a.$$

This is Lucas's congruence for the binomial coefficients $\binom{q}{v}$. So the entry at k equals $\sum_{u=0}^{r} (-1)^u \binom{r}{u} \Sigma_c(b+k+u)$, where $\Sigma_c(j) = \sum_{v=0}^{a} (-1)^v \binom{a}{v} R_c(j+vp)$. The arguments satisfy $b \leq j$ and $j + ap \leq b + n - 1 + q = L - 1$. For $c = \iota(x)$ with $x \in \ker \Phi$ we have $\Sigma_c(j) = \pm g_x(j)$. The difference of order r of a polynomial of degree less than r vanishes identically. So $\iota(\ker \Phi)$ lies in the left kernel of $\mathcal{A}^0_{N,p}$. Since ι is injective, $\nu_N(p) \geq \dim \ker \Phi = \dim \mathcal{V} - \operatorname{rank} \Phi$. □

Suppose instead that the conditions only force the values on a set $\mathcal{C}$ of columns to vanish. The same argument, with the remaining $n - |\mathcal{C}|$ entries added to the conditions, gives $\nu_N(p) \geq \dim \mathcal{V} - \operatorname{rank} \Phi - (n - |\mathcal{C}|)$. The proofs of the four closed formulas below are of this form. In Theorems 4.3 and 4.5 the set $\mathcal{C}$ consists of all columns, and in Theorems 4.6 and 4.8 it is a block of columns.

4.3. **Four closed formulas.** Each closed formula exhibits explicit vectors in the left kernel. The rows used are symmetric about the centres $g + kp$, and the identities (4.1) make their combinations vanish modulo p. Put $e = (p+1)/2 - q - h$.

Theorem 4.2. *Let $N \geq 1$ and let p be an odd prime with $H < p^2$. Then $\nu_N(p)$ is at least each of the following quantities whenever the stated conditions hold. Negative values are replaced by 0.*

(a) (Theorem 4.3) If $2g < p$, $H \leq 4p$ and $k_T = \min(g,\, n-1-g-2p,\, C-3p) \geq 1$:

$$k_T - (n + q + g + h - 3p)_+.$$

(b) (Theorem 4.5) If $2h \leq p+1$, $L \leq 3p \leq H$, $n + q + g + h \leq 3p$ and $2p \leq n + q$:

$$\min(n,\, 2p + 2g + 1) - g - 2p - 1.$$

(c) (Theorem 4.6) If $e \geq 0$, $n \leq 4p + 2g + 1$, $L \leq 6p$ and $b + q \leq p + \varsigma_0$, with $\varsigma_0 = \max(q, H + q - 6p)$:

$$\min\Big(\frac{p-1}{2} - g,\ n + b + q + h - \frac{11p+1}{2}\Big).$$

(d) (Theorem 4.8) If $e \geq 0$ and $n \leq 4p + 2g + 1$: $m_1 - m_0 - e$, with

$$m_0 = \max(b + q,\, L + \chi - 4p), \qquad m_1 = \min(p,\, L - 3p,\, A + 1).$$

For our parameters these bounds apply for $12N < p \leq 26N$, where they give most of the gain (Section 4.10). We prove the four parts in turn.

4.4. **Antisymmetric combinations of rows.**

Theorem 4.3. *Let $2g < p$, $H \leq 4p$ and $k_T = \min(g,\, n-1-g-2p,\, C-3p) \geq 1$. Then $\nu_N(p) \geq k_T - (n + q + g + h - 3p)_+$.*

The rows used are the differences of rows placed symmetrically about the three centres g, $g + p$ and $g + 2p$.

Proof. For $1 \leq d \leq k_T$ and $k = 0, 1, 2$ let $\epsilon_k(d) = e_{g+kp+d} - e_{g+kp-d} \in \mathbb{F}_p^n$. These rows exist because $d \leq g$ and $g + 2p + d \leq n - 1$. The $3k_T$ vectors $\epsilon_k(d)$ are linearly independent, since their supports are disjoint. Indeed the supports lie in the intervals $[kp, kp + 2g]$ of row indices, which are disjoint

because $2g < p$. The row $g + kp \pm d$ has $m = kp \pm d$. By $T_{a+b} - T_{a-b} = 2(X^2-1)U_{a-1}U_{b-1}$ and $U_{a+b-1} - U_{a-b-1} = 2T_aU_{b-1}$ together with (4.1),

$$T_{kp+d} - T_{kp-d} = 2(X^2-1)^{(p+1)/2}\, U_{k-1}(X^p)\, U_{d-1},$$
$$U_{kp+d-1} - U_{kp-d-1} = 2\, T_k(X^p)\, U_{d-1}.$$

Put $R_d = 2(1-X)^h X^{C-d}\, \mathrm{rev}_{d-1}\, U_{d-1}$, the value of $D_{\epsilon_0(d)}$. Reverse at degree $C-1$ and use $\mathrm{rev}_{kp}(T_k(X^p)) = (\mathrm{rev}_k\, T_k)(X^p)$, with $T_0 = 1$, $T_1 = X$ and $T_2 = 2X^2 - 1$. One finds

$$D_{\epsilon_1(d)} = X^{-p}R_d, \qquad D_{\epsilon_2(d)} = 2X^{-2p}R_d - R_d,$$

$$XWP_{\epsilon_0(d)} = 0, \qquad XWP_{\epsilon_1(d)} = X^{-p}R_d - X^pR_d,$$
$$XWP_{\epsilon_2(d)} = 2X^{-2p}R_d - 2R_d.$$

The polynomial R_d is supported in degrees $[C-d, H-1]$, which lies in $[3p, 4p)$ because $d \le C - 3p$ and $H \le 4p$. So at each residue ℓ it has a single base-p digit $\varrho_\ell(d)$, in position $u = 3$. For $x = (x_0, x_1, x_2) \in \mathbb{F}_p^3$ the value of $\sum_k x_k\epsilon_k(d)$ at the moment column $Jp + \ell$ is therefore, by Lemma 3.2, $\varrho_\ell(d)\, f_J(x)$ with

$$f_J(x) = x_0\beta_{3,J} + x_1(\alpha_{2,J} - \alpha_{4,J} + \beta_{2,J}) + x_2(2\alpha_{1,J} - 2\alpha_{3,J} + 2\beta_{1,J} - \beta_{3,J}).$$

This depends neither on d nor on ℓ. The moment columns with $J \ge 3$ do not meet the support, since $3p + \ell \ge C - d \ge C - g = L$. Choose $x^* \ne 0$ with $f_1(x^*) = f_2(x^*) = 0$, which is two linear equations in three unknowns. For $c = \sum_d s_d \sum_k x_k^*\epsilon_k(d)$ the values at the columns with $J = 1, 2$ vanish. The columns with $J = 0$ have $\ell \ge b$, and the value there is $f_0(x^*)\sum_d s_d\varrho_\ell(d)$, the coefficient of $X^{3p+\ell}$ in $\sum_d s_dR_d$. These are at most $(H - 3p - b)_+ = (n+q+g+h-3p)_+$ linear conditions on $s = (s_d)$. The map $s \mapsto c$ is injective, so the solutions form a space of dimension at least $k_T - (n+q+g+h-3p)_+$ in the left kernel. □

4.5. **A reflection of Chebyshev polynomials.** The next bound rests on a reflection formula modulo p.

Lemma 4.4. *Let p be an odd prime and $\varpi = (p+1)/2$. In $\mathbb{F}_p[X]$, for every integer m,*

$$T_{p-m} = X^pT_m - (X^2-1)^{\varpi}\, U_{m-1}.$$

Proof. Both sides satisfy the three-term recurrence in m, so it suffices to check $m = 0$ and $m = 1$. For $m = 0$ it is $T_p = X^p$, by (4.1). For $m = 1$ use $T_{p+1} = XT_p + (X^2-1)U_{p-1}$ and $T_{p-1} = 2XT_p - T_{p+1}$. □

Theorem 4.5. *Let $2h \le p+1$, $L \le 3p \le H$, $n+q+g+h \le 3p$ and $2p \le n+q$, and put $n^* = \min(n, 2p+2g+1)$. Then $\nu_N(p) \ge n^* - g - 2p - 1$.*

Proof. We restrict c to the rows $r < n^*$, so that $m_r \le n^* - 1 - g \le 2p + g$. Apply Lemma 4.4 with $m = m_r$ and reflect at degree $p + C - 1$. For every such row this gives

$$(1 - X^2)^{\varpi} \operatorname{rev}_{C-1} U_{m-1} - X \operatorname{rev}_{C-1} T_m = -X^{p+1} \operatorname{rev}_{C-1} T_{p-m}.$$

The right side is divisible by $X^{p+C-|p-m|}$, hence by X^L since $|p - m| \le p + g = p + C - L$. Multiplying by $(1 - X)^h$ and summing, we get for every c supported on these rows

$$X^L \mid (1 - X^2)^{\varpi} D_c - X P_c. \tag{4.2}$$

Let $\mathcal{V}$ be the space of such c with $(1 - X^2)^{\varpi} \mid P_c$. Since $(1 - X)^h$ divides every P_r, this is $\varpi - h$ conditions at $X = 1$ and ϖ conditions at $X = -1$, that is, at most $p + 1 - h$ linear conditions. For $c \in \mathcal{V}$ write $P_c = (1 - X^2)^{\varpi} Y$ and $\widetilde{Q} = XY$. Since $(1 - X^2)^p = 1 - X^{2p}$,

$$XW P_c = (1 - X^{2p})\, \widetilde{Q}.$$

The polynomial P_c has no coefficients below X^{A^*}, $A^* = C - n^* + g$, and none from X^H on. So $\widetilde{Q}$ is supported in $[A^* + 1,\ H - p - 1]$. Dividing (4.2) by the unit $(1 - X^2)^{\varpi}$ modulo X^L gives $D_c[k] = \widetilde{Q}[k]$ for $k < L$.

Now let $H - 3p \le \ell < L - 2p$. The inequality $L - 2p \le A^* + 1$ is equivalent to $n^* \le 2p + 2g + 1$, and $L - 2p \le p$ since $L \le 3p$. By the supports, $\widetilde{Q}[\ell] = 0$, $\widetilde{Q}[2p + \ell] = 0$ and $\widetilde{Q}[3p + \ell] = 0$, and $D_c[3p + \ell] = 0$ because $3p + \ell \ge H$. So with $\varkappa = \widetilde{Q}[p + \ell]$ the base-p digits of $XW P_c$ at the residue ℓ are $(0, \varkappa, 0, -\varkappa, 0, \dots)$, and those of D_c are $(0, \varkappa, 0, 0, \dots)$. Hence, by Lemma 3.2, the value at every moment column $Jp + \ell$ with such an ℓ is $\varkappa$ times a constant depending on J only. On $\mathcal{V}$ these columns, which have $J \in \{0, 1, 2\}$, therefore span a space of dimension at most one for each residue ℓ.

Put $\ell_0 = H - 3p$ and $\ell_1 = L - 2p$. The hypotheses give $\ell_0 \le b \le \ell_1$. For $\ell_0 \le \ell < b$ the moment columns $j \in [b, L)$ with this residue have $J \in \{1, 2\}$, and for $b \le \ell < \ell_1$ they have $J \in \{0, 1, 2\}$. So the saving is $2\ell_1 - \ell_0 - b$. The number of linear conditions needed to make all values on the $n + q = L - b$ moment columns vanish on $\mathcal{V}$ is therefore $(p + 1 - h) + (n + q) - (2\ell_1 - \ell_0 - b)$. Substituting $\ell_1 = L - 2p$ and $\ell_0 = H - 3p$, this is $2p + 1 + g$. So the space of solutions in the span of n^* rows has dimension at least $n^* - g - 2p - 1$. □

4.6. **A block of columns: the digits 1 to 5.**

Theorem 4.6. *Let $e = (p + 1)/2 - q - h \ge 0$, $n \le 4p + 2g + 1$ and $L \le 6p$, and put $\varsigma_0 = \max(q, H + q - 6p)$ and $\varsigma_1 = L - 5p$. Assume $b + q \le p + \varsigma_0$. Then $\nu_N(p) \ge \varsigma_1 - \varsigma_0 - e$, that is,*

$$\nu_N(p) \ge \min\Big(\frac{p - 1}{2} - g,\ n + b + q + h - \frac{11p + 1}{2}\Big).$$

The first term is the minimum if $6p \le H$ and the second if $6p \ge H$.

In Theorems 4.3 and 4.5 the conditions are imposed on the rows. Here the argument concerns a block of columns. We first need an identity.

Lemma 4.7. *In $\mathbb{F}_p[X]$, for every integer k, $2X^p(X^2-1)^{(p-1)/2}T_k-(2X^{2p}-1)U_{k-1}=U_{2p-k-1}$. If $n\le 4p+2g+1$, then for every c, $X^L \mid 2XWS_c-(2-X^{2p})U_c$, and consequently*

$$2\,Q_c[i]=2\,\widehat{D}_c[i]-\widehat{D}_c[i-2p] \qquad (i<L). \tag{4.3}$$

Proof. For all integers m,k one has $U_{m-1}T_k-T_mU_{k-1}=U_{m-k-1}$, since both sides satisfy the same recurrence in k and agree for $k=0,1$. Take $m=2p$ and use (4.1): $T_{2p}=2X^{2p}-1$ and $U_{2p-1}=2T_pU_{p-1}=2X^p(X^2-1)^{(p-1)/2}$. Reflect at degree $2p+C-1$. The first term becomes $2XW\,\mathrm{rev}_{C-1}\,T_k$, the second $(2-X^{2p})\,\mathrm{rev}_{C-1}\,U_{k-1}$, and the right side $X^{2p+C-1}U_{2p-k-1}(1/X)$. For $k=m_r\ge -g$ with $2p-k-1\ge 0$ the right side is divisible by X^{C+k}, hence by X^L. If $2p-k-1<0$, it has degree $k-2p-1$ and is divisible by X^{4p+C-k}, hence by X^L when $k\le 4p+g$. This holds because $k\le n-1-g\le 4p+g$. Summing with weights c_r and multiplying by $(X-1)^q(1-X)^h$ gives (4.3). □

Proof of Theorem 4.6. Let $\mathcal{M}=[\varsigma_0,\varsigma_1)$ and $w=\varsigma_1-\varsigma_0$. We may assume $w>e$. The block consists of the columns $k(J,\varsigma)=Jp+\varsigma-(b+q)$ for $1\le J\le 5$ and $\varsigma\in\mathcal{M}$. These indices are nonnegative by the hypothesis $b+q\le p+\varsigma_0$, and smaller than n because $5p+\varsigma_1-1=L-1$. There are $5w$ of them. As $\varsigma<p$, formula (3.1) applies, and the value of c at $k(J,\varsigma)$ is $\sum_u \alpha_{u,J}Q_c[up+\varsigma]+\beta_{u,J}\widehat{D}_c[up+\varsigma]$.

Supports. For $\varsigma\in\mathcal{M}$ we have $p+\varsigma\le p+\varsigma_1-1=L-4p-1\le A$ by $n\le 4p+2g+1$. So the digits $u=0,1$ of Q_c and $\widehat{D}_c$ vanish. From $\varsigma_0\ge H+q-6p=L+\chi-6p$ we get that the digits $u\ge 6$ of $\widehat{D}_c$ and $u\ge 7$ of Q_c vanish. For $u\le 5$ we have $up+\varsigma\le L-1$, so (4.3) applies.

From Q to $\widehat{D}$. For $u=2,\dots,5$, (4.3) gives $\widehat{D}_c[up+\varsigma]=Q_c[up+\varsigma]+\frac{1}{2}\widehat{D}_c[(u-2)p+\varsigma]$, and for $u=2,3$ the last index is at most A, so that term vanishes. Hence if $Q_c[up+\varsigma]=0$ for $u=2,\dots,5$ and all $\varsigma\in\mathcal{M}$, then by induction $\widehat{D}_c[up+\varsigma]=0$ for $u=2,\dots,5$, and all digits of $\widehat{D}_c$ in $\mathcal{M}$ vanish.

A relation from the Frobenius map. We have $Q_c=(-1)^qX(1-X)^{q+h+(p-1)/2}(1+X)^{(p-1)/2}S_c$, and $q+h+(p-1)/2+e=p$. Since $(1-X)^p=1-X^p$,

$$(1-X)^e\,Q_c=(1-X^p)\,Y_c, \qquad Y_c=(-1)^qX(1+X)^{(p-1)/2}S_c,$$

and Y_c is supported in $[A+1,\,C+(p-1)/2]$. Compare the coefficients at $up+\varsigma$, subtract the coefficient at $(u-1)p+\varsigma$, and sum over $u=2,\dots,6$. The right side telescopes to $Y_c[6p+\varsigma]-Y_c[p+\varsigma]$, which vanishes for $\varsigma\ge\varsigma_0+e$. The second term vanishes because $p+\varsigma\le A$, and the first because $6p+\varsigma\ge H+q+e=C+(p+1)/2>\deg Y_c$. So, if $Q_c[up+\varsigma']=0$ for $u=2,\dots,5$ and all $\varsigma'\in\mathcal{M}$, then

$$\sum_{t=0}^{e}(-1)^t\binom{e}{t}Q_c[6p+\varsigma-t]=0 \qquad (\varsigma_0+e\le\varsigma<\varsigma_1).$$

This is a recurrence with unit leading coefficient. So if $Q_c[6p+\varsigma]=0$ for the first e values $\varsigma=\varsigma_0,\dots,\varsigma_0+e-1$, then it vanishes for all $\varsigma\in\mathcal{M}$.

Counting. Impose on c that the values at the $n-5w$ columns outside the block vanish, that $Q_c[up+\varsigma]=0$ for $2\le u\le 5$ and $\varsigma\in\mathcal{M}$ ($4w$ conditions), and that $Q_c[6p+\varsigma]=0$ for the first e values of ς. By the three steps above all the values at the block columns vanish. These are at most $n-5w+4w+e=n-(w-e)$ linear conditions, so the left kernel has dimension at least $w-e$. Finally, $w-e$ equals $\varsigma_1-q-e=n+b+q+h-(11p+1)/2$ if $\varsigma_0=q$, and $(p-1)/2-g$ if $\varsigma_0=H+q-6p$. □

4.7. **The digits 1 to 4.**

Theorem 4.8. *Let $e=(p+1)/2-q-h\ge 0$ and $n\le 4p+2g+1$, and put*

$$m_0=\max(b+q,\,L+\chi-4p),\qquad m_1=\min(p,\,L-3p,\,A+1).$$

Then $\nu_N(p)\ge m_1-m_0-e$.

Proof. Let $\mathcal{M}=[m_0,m_1)$ and $w=m_1-m_0>e$. For $\varsigma\in\mathcal{M}$ we have $\varsigma\le A$, so the digit $u=0$ of Q_c and $\widehat{D}_c$ vanishes. The inequality $\varsigma\ge L+\chi-4p$ gives that the digits $u\ge 4$ of $\widehat{D}_c$ and $u\ge 5$ of Q_c vanish. Finally $3p+\varsigma<L$. So (4.3) gives, with $D_i=\widehat{D}_c[ip+\varsigma]$ and $Q_i=Q_c[ip+\varsigma]$,

$$Q_1=D_1,\qquad Q_2=D_2,\qquad Q_3=D_3-\tfrac{1}{2}D_1.$$

So every value of c at a column with residue $\varsigma\in\mathcal{M}$ is a combination of the four functionals $D_1(\varsigma),D_2(\varsigma),D_3(\varsigma),Q_4(\varsigma)$ of c. For $\varsigma\in\mathcal{M}$ the columns $k=\varsigma-(b+q)$, $p+\varsigma-(b+q)$, $2p+\varsigma-(b+q)$ and $3p+\varsigma-(b+q)$ exist, since $b+q\le\varsigma<p$ and $3p+\varsigma<L$. So there are at least $4w$ block columns, and their entries lie in the span $\mathcal{F}$ of the $4w$ functionals.

We find $w-e$ linear relations among these functionals. As in the proof of Theorem 4.6, $(1-X)^eQ_c=(1-X^p)Y_c$ with Y_c a polynomial. Put $\sigma_c(\ell)=\sum_{u\ge0}Q_c[up+\ell]$ and sum the coefficients of this identity over the residue class of $\varsigma\ge e$ modulo p. The right side gives $\sum_u(Y_c[up+\varsigma]-Y_c[(u-1)p+\varsigma])=0$, since Y_c is a polynomial. The left side gives $\sum_{t=0}^{e}(-1)^t\binom{e}{t}\sigma_c(\varsigma-t)$. For $\varsigma\in\mathcal{M}$ with $\varsigma-e\ge m_0$ all $\varsigma-t$ lie in $\mathcal{M}$, and there $\sigma_c=Q_1+Q_2+Q_3+Q_4=D_1+D_2+D_3-\frac{1}{2}D_1+Q_4$. So

$$\sum_{t=0}^{e}(-1)^t\binom{e}{t}(\tfrac{1}{2}D_1+D_2+D_3+Q_4)(\varsigma-t)=0\qquad(m_0+e\le\varsigma<m_1).$$

These $w-e$ relations are independent, since $Q_4(\varsigma)$ occurs with coefficient 1 in the relation for ς and only with smaller arguments in the others. Hence $\dim\mathcal{F}\le 4w-(w-e)$, and the rank of the matrix is at most $(n-4w)+3w+e=n-(w-e)$. □

The proof uses neither a determinant nor a lower bound for N. Unlike the proof of Theorem 4.6, the telescoping here runs over all digits, so no vanishing of Y_c at the ends is needed.

4.8. **Decompositions modulo p.** The further families organize the row combinations as follows.

Definition 4.9. A *decomposition modulo p* of a family of row combinations consists of

(a) a linear map $x \mapsto c(x)$ from $(\mathbb{F}_p^K)^t$ to $\mathbb{F}_p^n$, injective on a subspace of admissible $x = (x_1, \dots, x_t)$;
(b) polynomials $\pi_1, \dots, \pi_t \in \mathbb{F}_p[X]$ of degree less than $e_{\max}$;
(c) linear maps $\mathcal{T}, \mathcal{G} \colon \mathbb{F}_p^K \to \mathbb{F}_p[Y]$,

such that for admissible x

$$(X-1)^h P_{c(x)} = \sum_d \pi_d(X)\, \mathcal{T}(x_d)(X^p),$$
$$(X-1)^h (X^2-1)^{\frac{p-1}{2}} D_{c(x)} = \sum_d \pi_d(X)\, \mathcal{G}(x_d)(X^p).$$

The identities (4.1) and $(1-X)^p = 1 - X^p$ provide such decompositions. By the digit formula, the value of $c(x)$ at the moment column $Jp + \ell$ is a linear form in the x_d:

$$\sum_{0 \le e' < e_{\max},\ p \mid H-1-\ell-e'} \Theta_{J,\,(H-1-\ell-e')/p}\Big(\sum_d \pi_d[e']\, x_d\Big).$$

Here $\pi_d[e']$ is the coefficient of $X^{e'}$, and each $\Theta_{J,s}$ is a linear form in J and the digit s, determined by $\mathcal{T}$ and $\mathcal{G}$.

4.9. **Further families.** The following lower bounds for $\nu_N(p)$ are proved for all parameters, by explicit choices of the space, the conditions and the columns in Theorem 4.1 or its variant. Each is a piecewise linear function of p and the parameters, positive on intervals described by linear inequalities. We describe them in words. The statements are those of the formalization.

(F1) *Frobenius products.* The row combinations have polynomials that are the product of a fixed factor, a free polynomial of bounded degree d_P in the Chebyshev variable, and a polynomial $\Xi(X^p)$ with exponents in an interval of K consecutive values. The conditions require Ξ to annihilate the functionals seen by the columns. If #seen of these functionals are independent, the bound is $(d_P + 1)(K - \#\text{seen})$. For our parameters this gives $53N - 2p - 1$ on $(26N, \frac{53}{2}N)$, where $2p + 2 \le 53N$.

(F2) *A truncated form of Theorem 4.6.* Without the hypothesis $n \le 4p + 2g + 1$, the bound $\varsigma_1 - \varsigma_0 - e_+ - (n - 4p - 2g - 1)_+$ holds whenever $L \le 6p$ and $b + q \le p + \varsigma_0$. For our parameters it gives $5p - 59N + 1$ on $(\frac{59}{5}N, 12N]$.

(F3) *A tail of the antisymmetric family.* With the four centres $g + kp$ $(0 \le k \le 3)$ and $k'_T = n - 1 - g - 3p$, the bound is $p - 1 - (2g + q + h)$, under $4p \le L$, $b \le p$, $k'_T \le g$, $k'_T \le (p-1)/2$, $3p + 1 + g \le n$ and $4p + b \le H$. For our parameters this is $p - 16N - 1$ on $(16N, \frac{118}{7}N]$.

(F4) *Two families of decompositions and their sum.* The first fixes a digit $s \le H/p$ and a range of blocks of rows, with polynomials π_d divisible by a power of $X-1$. The second has polynomials $(X^p-1)\pi$ with π supported in an interval of degrees. Under a separation condition their kernels are independent when both apply, at the same digit s or at the digits s and $s-1$. Then the sum of the two bounds is a lower bound. This is the one place where two bounds are added.

(F5) *Families defined by rational equations.* In five further families the coefficients of the decomposition satisfy a fixed system of linear equations with rational coefficients. A solution modulo p exists for every p prime to a fixed integer, and the bound is the dimension of the solution space minus the number of conditions. In two of these families the entries depend only on the residue of the column modulo p, so the difference of order q imposes $(p-q)_+$ conditions. When $h \ge p$ one writes $(X-1)^h = (X^p-1)(X-1)^{h-p}$, which imposes $h-p$ more conditions. These families give the bounds on $(\frac{58}{17}N, \frac{59}{17}N]$, $(\frac{29}{8}N, \frac{11}{3}N]$, $(\frac{16}{7}N, \frac{5}{2}N]$ and on six intervals in $(\frac{341}{25}N, \frac{809}{50}N]$.

(F6) *Block decompositions.* Write a row index as $kp+d$ and put $y = X^p$. For $h \ge p$ the identity

$$(X-1)^h W T_{kp+d} = \big((X-1)T_k\big)(y)\, Q_1 + \big((X-1)(X^2-1)U_{k-1}\big)(y)\, Q_2$$

and its analogue for U_{kp+d-1} hold in $\mathbb{F}_p[X]$, with $Q_1 = (X-1)^{h-p} W T_d$ and $Q_2 = (X-1)^{h-p} U_{d-1}$. At a residue with a single nonzero digit the values of all rows of a block are proportional, and the bound counts the blocks. For $h < p$ a difference of order $p-h$ of the coefficients plays the same role. For our parameters these give bounds on $(\frac{59}{21}N, 3N]$, where $h \ge p$, and on $(3N, \frac{16}{5}N]$.

For the intervals of primes that these families do not cover, we apply the counting inequality to every constant subspace at once. The space of row combinations is the tensor product of a space of polynomials with a subspace U of a fixed finite-dimensional profile space, and the conditions are Taylor coefficients composed with fixed linear maps. The corank is then at least $c \dim U$ minus a weighted sum of ranks on U, for an integer c (Theorem A.1). We never exhibit U. Forced intersections of row spaces show that this quantity is large for some subspace, and only upper bounds for ranks are needed. Appendix A states this form and describes how its hypotheses are verified.

4.10. **The corank profile for our parameters.** The bounds combine into the following statement.

Theorem 4.10. *For the parameters* $(58, 9, 3, 5, 3)$ *there are* 113 *intervals* $(a_i, b_i]$, *with* $\frac{16}{7} = a_1 < b_1 \le a_2 < \cdots < b_{113} = \frac{53}{2}$, *and a function* κ, *linear on each interval and* 0 *outside their union, with the following property. For all large* N *and every prime* p *with* $a_i N < p \le b_i N$ *for some* i,

$$\nu_N(p) \ge N\kappa(p/N) - 2.$$

On each interval the bound is given by one of the families of this section or by Theorem A.1. The area of κ is a rational number,

$$\int_0^\infty \kappa(\theta)\,d\theta = \frac{1002889739024620759352746227576740574893}{45460332302422237200000000000000000000} = 22.06076\ldots \;\geq\; 22.0607.$$

The contributions by ranges of $\theta = p/N$ are as follows. The third column counts the intervals on which the bound comes from Theorem A.1.

range of θ	intervals	by Theorem A.1	area
(16/7, 3]	11	0	0.06259
(3, 6]	54	24	0.25354
(6, 12]	31	12	0.71870
(12, 53/2]	17	1	21.02592
total	113	37	22.06076

Almost all of the area lies in $(12, \frac{53}{2}]$, where κ is as follows.

interval of θ	$\kappa(\theta)$	bound
$(12, 13]$	$\theta/2 - 5$	Thm. 4.2(c), $(p-1)/2 - g$
(13, 146/11]	$73 - 11\theta/2$	Thm. 4.2(c), second term
(341/25, 69/5]	$11\theta/2 - 75$	(F5)
(69/5, 14]	$\theta/2 - 6$	(F5)
(14, 711/50]	$64 - 9\theta/2$	(F5)
(30287/2000, 78/5]	$7\theta/2 - 53$	(F5)
(78/5, 16]	$25 - 3\theta/2$	(F5)
(16, 809/50]	$73 - 9\theta/2$	(F5)
$(809/50 + \tau,\ 1619/100 - \tau]$	$72.999999 - 9\theta/2$	Thm. A.1
(1619/100, 118/7]	$\theta - 16$	(F3)
(118/7, 69/4]	$9\theta/2 - 75$	Thm. 4.2(d)
(69/4, 35/2]	$\theta/2 - 6$	Thm. 4.2(d)
(35/2, 128/7]	$64 - 7\theta/2$	Thm. 4.2(d)
(64/3, 23]	$3\theta - 64$	Thm. 4.2(a)
(23, 24]	5	Thm. 4.2(a), then (b)
(24, 26]	$53 - 2\theta$	Thm. 4.2(b)
(26, 53/2]	$53 - 2\theta$	(F1)

In the rows of Theorem 4.2 the bound is $3p - (n + q + h)$ for part (a), which applies up to $\theta = 70/3$ and is followed by part (b), with the bound $n - g - 2p$ from $\theta = 24$. Part (d) gives $(9p - 2C)/2$, $(p-1)/2 + h - b$ and $(2(n+q+h) - 7p)/2$ on its three intervals. Here $\tau = 10^{-10}$. The intervals that come from Theorem A.1 are shortened by τ at both ends and their bounds are lowered by 10^{-6}, which absorbs the rounding of the integer orders in that theorem. The areas of the rows of this table add up to $21.02592\ldots$, and the 96 intervals below 12 contribute $1.03484\ldots$, so the total is $22.06076\ldots$.

The exact sum is computed in the formalization from the list of intervals. Figure 2 shows κ on $(12, \frac{53}{2}]$.

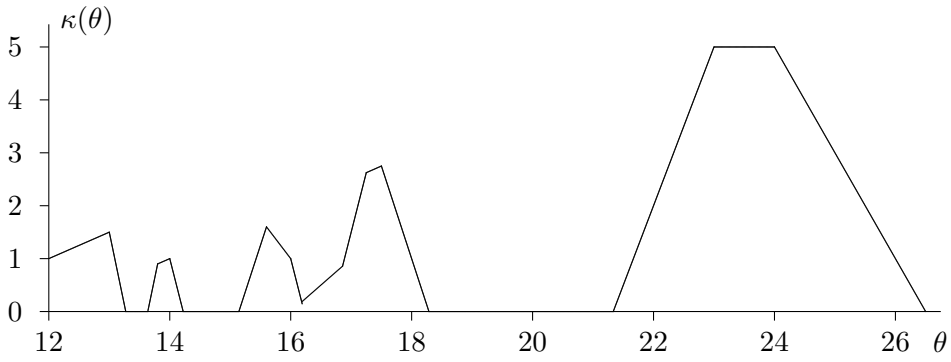


FIGURE 2. The corank profile $\kappa(\theta)$ on $(12, \frac{53}{2}]$, where $\nu_N(p) \geq N\kappa(p/N) - 2$ for $p = \theta N$. The bound on the pole count is $f(\theta) = 116$ on this whole range, so κ is the gain over OpenAI's denominators.

On each interval we use a single bound. The families come from different constructions, and we have not shown that their kernels are independent, except for the pair (F4). So the bounds are not added at a prime where several apply.

For Theorem 4.2 one checks the hypotheses on the intervals above. The hypotheses $e \geq 0$ and $n \leq 4p + 2g + 1$ both read $\theta \geq 12$, and $L \leq 6p$ reads $\theta \geq 35/3$. So Theorems 4.6 and 4.8 apply for $\theta \geq 12$. For $12 \leq \theta \leq 13$ one has $\varsigma_0 = H + q - 6p = 81N - 6p$, so the first branch of Theorem 4.6 applies. For $\theta \geq 13$ one has $\varsigma_0 = 3N$, and the second branch applies, which is positive until $\theta = 146/11$. In Theorem 4.8, $m_1 = p$ for $\theta \leq 35/2$ and $m_1 = 70N - 3p$ above, and $m_0 = 81N - 4p$ for $\theta \leq 69/4$ and $12N$ above. The value $m_1 - m_0 - e$ is $9\theta/2 - 75$, $\theta/2 - 6$ and $64 - 7\theta/2$ on the three intervals $(50/3, 69/4]$, $(69/4, 35/2]$ and $(35/2, 128/7]$, and it is nonpositive outside $(50/3, 128/7)$. Theorem 4.3 applies for $\theta > 10$ and $\theta \geq 39/2$, with $k_T = \min(5N, 52N - 1 - 2p, 75N - 3p)$. It gives $3p - (n + q + h) = 3\theta - 64$ up to $\theta = 23$ and the bound $g = 5N$ up to $\theta = 70/3$. From $\theta = 70/3$ to $\theta = 26$ the hypotheses of Theorem 4.5 hold. They read $3p \geq L$, $3p \leq H$, $2p \leq n + q$ and $2h \leq p + 1$. The bound is $\min(n, 2p + 2g + 1) - g - 2p - 1$, which is $5N$ for $\theta \leq 24$ and $53N - 2p - 1$ above. In each case the value of the theorem is the value in the table minus at most one unit, from the integer rounding. The same kind of check, with linear inequalities in N, p and the parameters, gives the bounds of Section 4.9 on their intervals, and Appendix A gives the others. This proves Theorem 4.10. The loss of 2 in its statement covers the integer rounding of all the bounds.

5. The neighbourhood of G

This section and the next two prove that Δ_N is small on a whole neighbourhood of G. The idea is to write the perturbation $\Delta_N(x) - \Delta_N(G)$ as a sum of determinants in which some columns of $\mathcal{A}_N$ are replaced by values of the row functions at prescribed nodes. Each such determinant is bounded by OpenAI's energy method, extended to configurations with prescribed points (Section 6). The loss caused by each replaced column, in the exponent, determines ρ. OpenAI's method has two cases. In the first it already gives the improved decay rate. In the second it falls short when few columns are replaced, and Section 7 closes that gap.

5.1. **The statement.** OpenAI's energy method bounds $|\Delta_N(G)|$ by $\exp(-(n_0^2(U - \frac{1}{2}\log 2) + o(1))N^2)$, where U is an upper bound for a discrete logarithmic energy (Section 6). For our parameters we certify $U = 2.255858836$, so OpenAI's estimate gives the decay rate $D_0 = n_0^2(U - \frac{1}{2}\log 2) = 6422.8355\ldots$ Let

$$\Gamma_0 = 24.64, \qquad U_0 = U + \Gamma_0/n_0^2 = 2.2631834\ldots,$$
$$D = n_0^2(U_0 - \tfrac{1}{2}\log 2) = 6447.4755\ldots$$

The number Γ_0 is the improvement of the decay rate proved in Section 7. Put

$$\rho = n_0 c + \psi, \qquad c = c_2 + \sigma, \tag{5.1}$$

with $c_2 = 3.596893$, $\sigma = 0.33871$ and $\psi = 50.111339578$, so that $\rho = 278.376313578$. Here c is the increase of the energy bound per prescribed column (Section 6), and ψ bounds the logarithmic size of the weight and the Vandermonde factor of one column, per unit of N (§5.4).

Theorem 5.1. *For our parameters, uniformly for real x with $|x-G| \le e^{-\rho N}$,*

$$|\Delta_N(x)| \le e^{-(D+o(1))N^2} \qquad (N \to \infty).$$

We use the following abbreviations for general parameters:

$$\alpha_A = \frac{A}{n}, \quad \beta_h = \frac{2h_0}{n_0}, \quad \gamma_H = \frac{H}{n}, \quad \gamma_d = \frac{d_0}{n_0},$$
$$a_0 = 1 - \frac{b_0 + q_0 + g_0}{n_0}, \quad a_1 = 3 - \frac{b_0 + q_0 + 2g_0}{n_0}.$$

For our parameters these are $\alpha_A = 11/29$, $\beta_h = 3/29$, $\gamma_H = 39/29$, $\gamma_d = 53/58$, $a_0 = 41/58$ and $a_1 = 76/29$.

5.2. **A quadrature formula for the coefficient of G.** Let $\varepsilon = x - G$. Then $\Delta_N(x) = \det(\mathcal{A}_N + \varepsilon\mathcal{L}_N)$. The question is how large the columns of $\mathcal{L}_N$ are compared with those of $\mathcal{A}_N$. By (2.3) the entries of $\mathcal{L}_N$ are linear functionals of the coefficients of the row polynomials. These polynomials have bounded degree, so the functionals can be written exactly as finite combinations of point evaluations. Let $t_1, \ldots, t_J$ be distinct points, with J at

least the number of degrees of freedom of the rows. Lagrange interpolation gives

$$\mathcal{L}_N(r,k) = \sum_{a=1}^{J} P_r(t_a)\,\beta_{a,k}, \tag{5.2}$$

where the weight $\beta_{a,k}$ is the value of the same functional on the a-th Lagrange basis polynomial ℓ_a. This is a quadrature formula of Gauss type. It makes the columns of $\mathcal{L}_N$ combinations of point values of the polynomials P_r. So the perturbed determinant can be treated by the energy method. The cost is the size of the weights.

We take $J = d_0 N = 53N$ nodes, which are Chebyshev nodes of short intervals. Let $[a_c, b_c]$, $1 \le c \le 106$, be fixed intervals with dyadic end points, 48 in $[-1, 0)$ and 58 in $(0, 1)$, adjacent within each side. Their union is $[-0.99931266\ldots, -0.37573294\ldots] \cup [0.19559072\ldots, 0.99999983\ldots]$. In the coordinate x with $t = 2x/(1+x^2)$ it is contained in

$$E_x = [-\tfrac{2409}{2500}, -\tfrac{19501}{100000}] \cup [\tfrac{98749}{1000000}, 1].$$

So the nodes avoid a neighbourhood of $x = 0$, where the fields of the energy method are largest. We pair the intervals $(2i-1, 2i)$ within each side, which gives $24 + 29 = 53$ pairs. The two members of a pair receive $K_c = \lceil N/2 \rceil$ and $\lfloor N/2 \rfloor$ nodes, so that $\sum_c K_c = 53N$. Let $\ell_c(z) = (2z - a_c - b_c)/(b_c - a_c)$ and $w_c = b_c - a_c$. The nodes in $[a_c, b_c]$ are the zeros of $T_{K_c} \circ \ell_c$. They lie in the open interval, so all nodes are distinct.

Let $G(w) = \log|w + \sqrt{w^2 - 1}|$, with the branch of modulus at least 1, be the Green function of $\mathbb{C} \setminus [-1, 1]$. Put $G_c = G \circ \ell_c$ and

$$L(z) = \tfrac{1}{2} \sum_c (G_c(z) + \log(w_c/4)). \tag{5.3}$$

Lemma 5.2. *There is a constant C such that for all N, every node t_a and every $z \in \mathbb{C}$ with $|z| \le 1$, the Lagrange polynomial ℓ_a of t_a satisfies $|\ell_a(z)| \le CN\, e^{N(L(z) - L(t_a))}$.*

Proof. We use three facts about T_K. First, $|T_K(w)| \le e^{KG(w)}$, since $2T_K(w) = \phi^K + \phi^{-K}$ with $\phi = w + \sqrt{w^2 - 1}$. Second, $|T_K(v)| = \cosh(KG(v)) \ge \frac{1}{2} e^{KG(v)}$ for real $|v| \ge 1$. Third, the Lagrange polynomial λ_j of the j-th of the K Chebyshev nodes satisfies $|\lambda_j(w)| \le K e^{(K-1)G(w)}$ for all $w \in \mathbb{C}$. For the third, write $w = \frac{1}{2}(u + u^{-1})$ with $|u| = e^{G(w)} \ge 1$ and the node as $\frac{1}{2}(v + v^{-1})$ with $|v| = 1$. Then

$$\frac{T_K(w)}{w - \frac{1}{2}(v + v^{-1})} = \Big(\sum_{i<K} u^i v^{K-1-i} \Big) \Big(\sum_{i<K} (uv)^{-i} \Big),$$

which has modulus at most $K^2 |u|^{K-1}$. The product of the differences of the j-th node from the others has modulus at least $K/2^{K-1}$. This gives the third fact.

Let t_a be the j-th node of the interval c. Grouping the factors of ℓ_a by intervals gives

$$\ell_a(z) = \lambda_j(\ell_c(z)) \prod_{c' \neq c} \frac{T_{K_{c'}}(\ell_{c'}(z))}{T_{K_{c'}}(\ell_{c'}(t_a))}.$$

For $c' \neq c$ the node lies outside the open interval $(a_{c'}, b_{c'})$, so $|\ell_{c'}(t_a)| \geq 1$. The three facts and $G_c(t_a) = 0$ give $|\ell_a(z)| \leq K_c 2^{105} e^{\Lambda(z)-\Lambda(t_a)}$ with $\Lambda = \sum_c K_c G_c$. Since $|K_c - N/2| \leq \frac{1}{2}$ and $\sum_c G_c$ is bounded on the closed unit disc, $\Lambda(z) - \Lambda(t_a) \leq \frac{N}{2} \sum_c (G_c(z) - G_c(t_a)) + O(1)$. The constants $\log(w_c/4)$ cancel in $L(z) - L(t_a)$, and $K_c \leq N$. □

The term $-L(t_a)$ depends on the node. We bound it by a piecewise linear function. On each interval $[a_d, b_d]$ we have $G_d = 0$, and for $c \neq d$ the function $|\ell_c|$ is affine with values at least 1. Since arccosh is concave and increasing on $[1, \infty)$, the function L is concave on each interval. Let $k_0 < \cdots < k_{865}$ be fixed dyadic knots that include all end points a_c, b_c, and let v_i be rational numbers with $v_i \geq -L(k_i) - C_0$, where $C_0 = 33.052$. Let $\hat{\delta}$ be the piecewise linear function through the points (k_i, v_i). By concavity,

$$-L(t) \leq C_0 + \hat{\delta}(t) \tag{5.4}$$

for every t in the union of the intervals, and so at every node. One value $L(k_i)$ needs no logarithm per term: $2L(k_i)$ is the logarithm of a product of 106 factors, each an algebraic number with one square root.

The weight of column k is a coefficient functional of ℓ_a multiplied by the column function. After removal of the common factor of the rows it equals

$$\beta'_k = 4\,[t^0]\,F_k(t)\,(1-t^{-2})^{-1/2}, \qquad F_k(t) = t^{\,A-b-q-k}(1-t)^h(t-1)^q \ell_a(t),$$

where $(1-t^{-2})^{-1/2} = \sum_l c_l t^{-2l}$ with $c_l = \binom{2l}{l} 4^{-l}$. The next lemma bounds it along a contour.

Lemma 5.3 (keyhole bound)**.** *Let P be a real polynomial, $E \geq 0$, and $F(t) = t^{-E}P(t)$. Let $0 < r \leq 1$, and suppose that $|F| \leq M$ on the circle $|t| = r$ and $|F| \leq M'$ on the segments $r \leq |t| \leq 1$ of the real axis. Then*

$$\Big| 4 \sum_{l \geq 0} c_l\,[t^{E+2l}]P \Big| \leq 4(1+E)\max(M, M').$$

Proof. Write $F = \Pi + R$ with Π a polynomial and $R = \sum_{i=1}^{E} \pi_i t^{-i}$ the principal part at 0. The functional sees only Π. Since $c_l = \frac{1}{\pi}\int_{-1}^{1} t^{2l}(1-t^2)^{-1/2}dt$, its value is $\frac{4}{\pi}\int_{-1}^{1} \Pi(t)(1-t^2)^{-1/2}dt$, which is at most $4\max_{[-1,1]}|\Pi|$. By Cauchy's estimate $|\pi_i| \leq Mr^i$, so $|R(t)| \leq EM$ for $r \leq |t| \leq 1$, and $|\Pi| \leq M' + EM$ there. For $|t| < r$ the maximum principle gives $|\Pi(t)| \leq \max_{|t'|=r}|\Pi| \leq (1+E)M$. □

The formal proof replaces the integral by the finite Gauss–Chebyshev identity $\frac{1}{M}\sum_{j<M} \cos^{2l}\frac{(2j+1)\pi}{2M} = c_l$ for $l < M$, and Cauchy's estimate by the

finite Cauchy formula at roots of unity. So no integration is needed. For $0 \le y \le 1$ and $z \in \mathbb{C}$ let

$$\Phi_y(z) = (10 - 58y)\log|z| + 6\log|1 - z| + L(z).$$

Here $10 = (A - b - q)/N$ and $6 = (q + h)/N$. With $y = k/n$, Lemmas 5.2 and 5.3 give

$$|\beta_k'| \le e^{(w(y) - L(t_a) + o(1))N}, \qquad w(y) = \max_{z \in \Gamma} \Phi_y(z), \tag{5.5}$$

where Γ is the circle $|z| = \frac{3}{4}$ together with the segments $\frac{3}{4} \le |z| \le 1$ of the real axis. One radius serves all columns and all nodes. The bound for this maximum is in §5.4.

5.3. **Column expansion and Andréief's identity.** Write $\mathcal{P}(r, a) = P_r(t_a)$. Expanding $\det(\mathcal{A}_N + \varepsilon\mathcal{P}\beta)$ multilinearly in the columns gives a sum over maps f from the columns to $\{\varnothing, 1, \dots, J\}$ of $\prod_{f(k) \ne \varnothing} \varepsilon\beta_{f(k),k} \det M_f$. Here M_f is the matrix whose k-th column is the k-th column of $\mathcal{A}_N$ if $f(k) = \varnothing$, and the vector $(P_r(t_{f(k)}))_r$ otherwise.

Lemma 5.4. *Suppose that* $|\det M_f| \le M_0 \prod_{f(k) \ne \varnothing} C_k$ *for every* f*, and that* $|\beta_{a,k}| \le B_k$*. Then*

$$|\det(\mathcal{A}_N + \varepsilon\mathcal{P}\beta) - \det \mathcal{A}_N| \le M_0\Big(\prod_k (1 + J|\varepsilon| B_k C_k) - 1\Big).$$

The lemma allows a different constant for each column. Each M_f has $n' = n - m$ columns that are integrals and m columns that are point values. Andréief's identity extends to this situation.

Lemma 5.5. *Let* m_1, m_2 *be finite measures, let* F_r $(r < n' + m)$*,* F_i' $(i < n')$ *and* k *be integrable functions, and let* w_j $(j < m)$ *be vectors indexed by* $r < n' + m$*. Let* B *be the matrix with* $B(r, i) = \iint F_r(t) k(t, s) F_i'(s)\, dm_1(t)\, dm_2(s)$ *for* $i < n'$ *and* $B(r, n' + j) = w_j(r)$ *for* $j < m$*. For* $t = (t_0, \dots, t_{n'-1})$ *let* $\mathcal{W}(t)$ *be the matrix whose column* $c < n'$ *is* $(F_r(t_c))_r$ *and whose column* $n' + j$ *is* w_j*. Then*

$$(n'!)^2 \det B = \iint \det \mathcal{W}(t) \cdot \det[k(t_i, s_j)]_{i,j<n'} \det[F_i'(s_j)]_{j,i<n'}\, dm_1^{n'}(t)\, dm_2^{n'}(s).$$

Proof. For $m = 0$ this is Andréief's identity applied twice [17]. In general, expand $\det B$ multilinearly in its first n' columns. This introduces one pair of variables (t_i, s_i) for each of them. Then antisymmetrize, first in the t variables and then in the s variables. The last m columns are constant throughout. □

We apply the lemma to M_f with $F_r = A_r$ and with $w_j = (P_r(t_{a_j}))_r$ at the prescribed nodes t_{a_j}. With $k(t, s) = 1/(1 - ts)$ the middle determinant is a Cauchy determinant. The third is a generalized Vandermonde determinant

$\prod_j s_j^b(1-s_j)^q \det[s_j^{E_i}]$. Here the exponents $E_i < n$ are those of the columns that remain integrals, and the missing exponents are the replaced columns. We bound the generalized Vandermonde determinant by a product over the missing exponents.

Lemma 5.6. *Let $0 < \zeta < 1$, let $s_1, \dots, s_{n'} \ge 0$, let $E_1 < \cdots < E_{n'} < n$ be integers, and let $\{k_1, \dots, k_m\}$ be the complement of $\{E_i\}$ in $\{0, \dots, n-1\}$. Then*

$$|\det[s_l^{E_i}]| \le \prod_l (1+\zeta s_l)^m \prod_{j=1}^{m} \mathcal{B}_\zeta(k_j) \prod_{i<l} |s_i - s_l|, \qquad \mathcal{B}_\zeta(k) = \sum_i \zeta^{-(n'-k+i)},$$

where the sum runs over the $i < m$ with $0 \le n' - k + i \le n'$. Moreover $\mathcal{B}_\zeta(k) \le \zeta^{-(n-1-k)}/(1-\zeta)$.

Proof. Let $P(X) = \prod_l (X - s_l) = \sum_r (-1)^r e_r(s) X^{n'-r}$. Multiplying the matrix by a unitriangular matrix built from P gives $|\det[s_l^{E_i}]| = \prod_{i<l} |s_i - s_l| \cdot |\det Q|$, where Q is the $m \times m$ matrix with entries $Q_{j,i} = [X^{k_j - i}]P$. So $|Q_{j,i}| = e_{n'-k_j+i}(s)$, and this is 0 unless $0 \le n' - k_j + i \le n'$. Since the s_l are nonnegative, $\prod_l (1+\zeta s_l) = \sum_r \zeta^r e_r(s)$ is a sum of nonnegative terms. So $e_r(s) \le \zeta^{-r} \prod_l (1+\zeta s_l)$, and the row j of Q has ℓ^1 norm at most $\prod_l (1+\zeta s_l)\, \mathcal{B}_\zeta(k_j)$. Hadamard's inequality, applied to the rows of Q in the ℓ^1 norm, gives the first claim. The exponents $n' - k + i$ in $\mathcal{B}_\zeta(k)$ are distinct and at most $n - 1 - k$, which gives the second. □

5.4. **The weight of one column.** The bound of Lemma 5.6 has two parts. The factor $\prod_l (1+\zeta s_l)^m$ depends on the free points. With $m = \alpha n$ it equals $\exp\bigl(n \sum_l \alpha \log(1+\zeta s_l)\bigr)$. So it adds the field $\alpha \log(1+\zeta s)$ to the column field of the energy, and we treat it in Section 6. The other factor belongs to the replaced columns. With $y = k/n$ we have $\zeta^{-(n-k)} = e^{58N(1-y)\log(1/\zeta)}$. By (5.5), the replaced column k at the node t_a therefore contributes at most

$$e^{(w(y)+58(1-y)\log(1/\zeta) - L(t_a) + o(1))N}.$$

We take $\zeta = \frac{3}{4}$, the radius of the contour Γ. Two inequalities bound w, with $D_\Gamma = 3.2506$:

$$6 \log|1-z| + L(z) \le D_\Gamma \qquad (|z| = \tfrac{3}{4}), \tag{5.6}$$

$$6 \log|1-x| + L(x) - 10 \log \tfrac{3}{4} \le D_\Gamma \qquad (x \in \mathbb{R},\ \tfrac{3}{4} \le |x| < 1). \tag{5.7}$$

Appendix B proves them. On the circle, (5.6) gives $\Phi_y(z) \le (10-58y)\log\frac{3}{4} + D_\Gamma$. On a real segment put $l = \log\frac{3}{4}$ and $u = \log|x|$, so that $l \le u \le 0$. Then $(10-58y)(u-l) \le 10(u-l) \le -10l$ for $0 \le y \le 1$, and (5.7) gives the same bound. The end points $|x| = 1$ follow by continuity. Hence

$$w(y) + 58(1-y)\log\tfrac{4}{3} \le (10-58y)\log\tfrac{3}{4} + D_\Gamma - 58(1-y)\log\tfrac{3}{4} = D_\Gamma - 48\log\tfrac{3}{4}$$

for every y. The dependence on the column cancels exactly. This is the reason for choosing the parameter of Lemma 5.6 equal to the radius of

the contour. With (5.4), the column k at the node t_a contributes at most $e^{(\psi+\hat{\delta}(t_a)+o(1))N}$ as soon as

$$\psi > C_0 + D_\Gamma - 48\log\tfrac{3}{4} = 50.1113394\ldots$$

We take $\psi = 50.111339578$, which leaves a strict margin for the terms $o(1)$. The factor $e^{N\hat{\delta}(t_a)}$ depends on the node. We do not bound it here. It is carried into the energy bound, where it raises the field of each prescribed point y by $\hat{\delta}(\mathsf{t}(y))/58$, with $\mathsf{t}(y) = 2y/(1+y^2)$.

The remaining part of ρ is $n_0 c$, the increase of the energy bound caused by the prescribed columns. It comes from Section 6, and with it $\rho = 58 \cdot 3.935603 + 50.111339578 = 278.376313578$.

6. The energy method with prescribed columns

This section recalls OpenAI's energy method, extends it to determinants with prescribed columns, and proves Theorem 5.1 assuming the bound of Section 7. Throughout, $\alpha_1 = 51/500$.

6.1. **OpenAI's energy.** By Andréief's identity, $\det \mathcal{A}_N$ is an integral over n points $x_i \in (-1,1)$ and n points $s_j \in (0,1)$. OpenAI bounds the integrand by $e^{n^2 E}$, where E is a discrete logarithmic energy of the points, written in the coordinate x with $t = 2x/(1+x^2)$ [36, §5]. It has the form

$$n^2 E = \sum_i V_1(x_i) + \sum_j V_2(s_j) + \sum_{i<i'} \big(\log|x_i - x_{i'}| + \log(1 - x_i x_{i'})\big) + \sum_{j<j'} \log|s_j - s_{j'}| - \sum_{i,j} \log(1 - 2x_i s_j + x_i^2),$$

schematically, with external fields V_1, V_2 of order n. The first two sums are the external fields. The next two are the logarithmic interactions among points of the same kind, and the last is the interaction between the two kinds. There are two cases, according to which of OpenAI's two bounds for the determinant of the row functions is used [36, Props. 5.2, 5.3]. The case with doubled interaction is the one in which OpenAI's condition $\sum_i (1-x_i^2)/(1+x_i^2) \le n - 1 - 2g$ [36, Eq. (67)] holds, and the constrained case is its complement. In the *constrained case* the row field acquires a term with a multiplier λ_1 for the constraint that defines the case. In the *case with doubled interaction* the interaction among the x_i has weight 2 instead of 1. This case also bounds a determinant $\det \mathsf{L}$ of size n by $e^{O(n\log n)}$, which is the factor that Section 7 improves.

The energy is bounded from above by duality. Given trial sequences $u = (u_k)$ and $v = (v_k)$, one writes the interaction terms as quadratic forms in the moments $\frac{1}{n}\sum_i T_k(x_i)$ and $\frac{1}{n}\sum_j s_j^k$. Then one completes the squares against u and v. What remains is the sum of the norms $\|u\|^2 = \sum_k u_k^2/k$ and $\|v\|^2$ and, for each point, a function of one variable. With $\mathsf{T}(u,x) =$

$\sum_k u_k T_k(x)/k$ and $\mathsf{S}(v,x) = \sum_k v_k x^k/k$, these functions are, in the case with doubled interaction,

$$X(x) = \alpha_A \log|x| + \beta_h \log(1-x) - \gamma_H \log(1+x^2) - 4\mathsf{T}(u,x) - \mathsf{S}(v,x),$$
$$Y(s) = \tfrac{b_0}{n_0} \log s + \tfrac{q_0}{n_0} \log(1-s) + 2\mathsf{T}(v,s).$$

In the constrained case, with a trial sequence u' in place of u, they are

$$X_1(x) = \alpha_A \log|x| + \beta_h \log(1-x) - (\gamma_H - \tfrac{1}{2}) \log(1+x^2) + \lambda_1 (\tfrac{2g_0}{n_0} - \tfrac{2x^2}{1+x^2}) - 2\mathsf{T}(u',x) - \mathsf{S}(v,x),$$

with the same Y. If $X \le M_X$ and $Y \le M_Y$ on the whole domains, the energy of every configuration is at most a constant determined by M_X, M_Y and the norms. In the language of potential theory the extremal configuration is described by an equilibrium problem for two interacting measures in an external field [44]. The best trial sequences are the moments of its solution, as Carreira observed [8]. Any trial sequences give a valid upper bound.

6.2. **Prescribed points.** When m of the n columns are point values, only $n-m$ points of each kind are integrated over, and the energy is still normalized by n. The prescribed points enter through the row determinant, now evaluated at the free points together with the prescribed ones. There are two kinds of prescribed points: the quadrature nodes, and points $y \in (0,1)$. In the coordinate x the nodes lie in the set E_W of the x with $|x| \le 1$ and $2x/(1+x^2)$ in one of the intervals of Section 5. We have $E_W \subset E_x$. Let $\alpha = m/n$ be the fraction of prescribed columns and α_2 the fraction of prescribed points of the second kind. Each missing free point changes the energy bound by a bounded amount. The total change is linear in m, and its slope $c = c_2 + \sigma$ enters ρ.

Two terms from Section 5 are added. The factor $\prod_l (1+\zeta s_l)^m$ of Lemma 5.6 adds the field $\alpha g(s)$, with $g(s) = \log(1+\frac{3}{4}s)$, to the column field. Since $g > 0$ on $(0,1]$, the value with αg is at most the value with γg for every $\gamma \ge \alpha$. The factor $e^{N\hat\delta}$ of the weights adds $\hat\delta(\mathsf{t}(x))/58$ to the field of every prescribed point x, where $\mathsf{t}(x) = 2x/(1+x^2)$. Indeed $n^2 \cdot \frac{1}{n} \sum_j \hat\delta(\mathsf{t}(y_j))/58 = N \sum_j \hat\delta(\mathsf{t}(y_j))$.

Every step of OpenAI's argument remains valid uniformly in the number of free points. The regularization of the logarithmic interactions, the diagonal corrections and the completion of squares only involve sums over the points present. For the prescribed points of the first kind the relevant field is

$$F_1(x) = -\gamma_d \log(1+x^2) - 2\mathsf{T}(u,x) + \tfrac{1}{58}\hat\delta(\mathsf{t}(x)), \qquad x \in E_W.$$

For those of the second kind it is

$$F_2(y) = -\gamma_d \log(1+y^2) + \frac{n_0 - 2g_0 - 1}{n_0} \log y - u_1 y + \tfrac{1}{58}\hat\delta(\mathsf{t}(y))$$

for $y \in E_W$ with $0 < y < 1$, where u_1 is the first coefficient of u. The coefficient $(n_0 - 2g_0 - 1)/n_0$ is a lower bound, valid for all $N \ge 1$, of $(n-1-2g)/n$. The exact coefficient of $-\log(1+x^2)$ in both fields is

$(n-g-1)/n = \gamma_d - 1/n$. Since $0 \le \log(1+x^2) \le \log 2$ for $|x| \le 1$, the exact fields exceed F_1 and F_2 by at most $\log 2/n$, which tends to zero. By Lemma 5.6 the generalized Vandermonde determinant contributes the factor $\prod_{i<l} |s_i - s_l|$ and the column field αg to the energy, and the factors $\mathcal{B}_\zeta$ are assigned to the columns. The energy bound then reduces to a quadratic inequality in α and α_2. The maximum over α_2 has a closed form.

Lemma 6.1 (energy inequality with prescribed points)**.** *Let $\ell = \log 2$ and let $\mathcal{N}, \mathcal{F}, \mathcal{R}, \mathcal{F}_2$ be real numbers. Put*

$$V(\alpha) = (-(a_0 - \tfrac{1}{2}) + (a_1 - 1)\alpha - \alpha^2)\ell + \mathcal{N} + (1-\alpha)\mathcal{F} + \alpha\mathcal{R} + \frac{(\mathcal{F}_2 + \ell - \mathcal{R})^2}{4\ell} + U_0 - \tfrac{1}{2}\ell - c_2\alpha.$$

Then for all real α and all $\alpha_2 \ge 0$

$$(-(a_0 - \tfrac{1}{2}) + (a_1 - 1)\alpha - \alpha^2 + \alpha_2 - \alpha_2^2)\ell + \mathcal{N} + (1-\alpha)\mathcal{F} + (\alpha - \alpha_2)\mathcal{R} + \alpha_2\mathcal{F}_2 \le -U_0 + \tfrac{1}{2}\ell + c_2\alpha + V(\alpha).$$

Proof. The difference of the two sides is $\alpha_2(\mathcal{F}_2 + \ell - \mathcal{R}) - \ell\alpha_2^2$, and $tb - \ell t^2 \le b^2/(4\ell)$ for all real t and b. □

So the energy bound holds at the slope $c = c_2 + \sigma$ on every range of α on which $V(\alpha) \le \sigma\alpha$. The lemma is applied with $\mathcal{N}$ a bound for $2\|u\|^2 + \frac{1}{2}\|v\|^2$, $\mathcal{F}$ a bound for $M_X + M_Y$, with the column field γg included in Y, $\mathcal{R}$ a bound for the supremum of F_1 and $\mathcal{F}_2$ a bound for the supremum of F_2. Keeping $\mathcal{R}$ and $\mathcal{F}_2$ apart is much better than one cap for both, because the moved trial sequences of §6.3 lower the first field and raise the second. The analogue for the constrained case has no prescribed points of the second kind. Its inequality is

$$(-(a_0 - \tfrac{1}{2}) + (a_1 - 1)\alpha - \alpha^2)\ell + n_1 + (1-\alpha)\mathcal{F}_1 + \alpha\mathcal{M}_1 \le -U_0 + \tfrac{1}{2}\ell + c\alpha,$$

with n_1 a bound for the norms of the trial sequences, $\mathcal{F}_1 = M_{X_1} + M_Y$, and $\mathcal{M}_1$ a bound for $\lambda_1(\frac{2g_0}{n_0} - \frac{2x^2}{1+x^2}) + \hat{\delta}(\mathsf{t}(x))/58$ on E_W. Since $1 - \frac{2x^2}{1+x^2} = \sqrt{1 - \mathsf{t}(x)^2}$, this bound is a maximum of affine functions of λ_1 over the segments of the chord table, with $\hat{\delta}$ replaced by its larger end value. So it is convex in λ_1.

The squares that remain. Section 7 uses what the completion of squares leaves, so we record it for the case with doubled interaction. Let μ_F, μ_M and μ_S be the normalized counting measures of the free points, the prescribed points of the first kind and the prescribed points of the second kind. Their total masses are $1 - \alpha$, $\alpha - \alpha_2$ and α_2. Put $W = \mu_F + \mu_M$, $A_k(\mu) = \int T_k \, d\mu$ and $\langle a, b\rangle = \sum_k a_k b_k / k$, so that $\|a\|^2 = \langle a, a\rangle$. Three products carry the interactions in the variable x. The Vandermonde product of the free points comes from the Cauchy determinant. The row determinant gives the Vandermonde product of the free points and the prescribed points of the first kind, and the Vandermonde product of the points y of the second kind. These last points enter the row determinant through $1/y$, so it also

gives the cross factors $|w-1/y| = (1-wy)/y$, for w a point of W. Since $-\log|x-x'| = \log 2 + 2\sum_k T_k(x)T_k(x')/k$ and $\log(1-wy) \le -wy$, the part of E quadratic in the Chebyshev moments is, apart from the diagonal terms and the multiples of $\log 2$, at most

$$-\|A(\mu_F)\|^2 - \|A(W)\|^2 - \|A(\mu_S)\|^2 - A_1(W)A_1(\mu_S).$$

For any measure μ we have $-\|A(\mu)\|^2 = \|u\|^2 - 2\langle u, A(\mu)\rangle - \|A(\mu)-u\|^2$. Using this for μ_F and W, and writing $A_1(W) = u_1 + (A_1(W)-u_1)$, we find that the display equals

$$2\|u\|^2 - 2\langle u, A(2\mu_F+\mu_M)\rangle - u_1A_1(\mu_S) - \mathcal{Q},$$

where

$$\mathcal{Q} = \|A(\mu_F)-u\|^2 + \|A(W)-u\|^2 + \|A(\mu_S)\|^2 + (A_1(W)-u_1)A_1(\mu_S).$$

The first three terms are the term $2\|u\|^2$ of $\mathcal{N}$ and the terms of X, F_1 and F_2 that contain u. So the energy bound obtained from Lemma 6.1 improves by $\mathcal{Q}$. To bound $\mathcal{Q}$ from below, let e_1 be the first unit vector and put $a = A(\mu_F) - u$ and $b = A(W) - u + \frac{1}{2}A_1(\mu_S)e_1$. Since $\|e_1\| = 1$,

$$\|b\|^2 = \|A(W)-u\|^2 + (A_1(W)-u_1)A_1(\mu_S) + \tfrac{1}{4}A_1(\mu_S)^2.$$

Also $\frac{1}{4}A_1(\mu_S)^2 \le \|A(\mu_S)\|^2$, so $\mathcal{Q} \ge \|a\|^2 + \|b\|^2 \ge 2\|\frac{1}{2}(a+b)\|^2$. The average of a and b is $A(\mu_F + \frac{1}{2}\mu_M) - u + \frac{1}{4}A_1(\mu_S)e_1$. This gives

$$\mathcal{Q} \ge \Delta, \qquad \Delta = 2\sum_k \frac{1}{k}\bigl(A_k(\mu_F+\tfrac{1}{2}\mu_M) + \tfrac{1}{4}A_1(\mu_S)\delta_{k1} - u_k\bigr)^2. \tag{6.1}$$

The weight $\frac{1}{2}$ of μ_M records that these points occur in one of the two Vandermonde products of the x variables, and the weight $\frac{1}{4}$ is half of the $\frac{1}{2}$ in b. The diagonal terms are treated by OpenAI's regularization, which replaces the moments by damped moments with an error that tends to zero, as in the proof of Lemma 7.9. We suppress this.

6.3. **The trial sequences.** The trial sequences have the form

$$u_k = c_k + \operatorname{Re}\sum_j r_j z_j^k \qquad (k \ge 1;\ c_k = 0 \text{ for } k > 160), \tag{6.2}$$

with 160 Chebyshev coefficients and finitely many geometric terms. The geometric terms have a simple meaning. Since $\sum_{k\ge1} z^k T_k(x)/k = -\frac{1}{2}\log(1-2xz+z^2)$, a term (z,r) contributes $-\frac{1}{2}\operatorname{Re} r\log(1-2xz+z^2)$ to the potential $\mathsf{T}(u,x)$. If $x_0 + i\epsilon_0 = \frac{1}{2}(z+z^{-1})$, then $1-2xz+z^2 = 2z(x_0+i\epsilon_0-x)$. So the term is, up to a constant, the logarithmic potential of a point charge of complex strength r at $x_0+i\epsilon_0$ and its conjugate. The extremal measure has a gap in its support around $x=0$. Its potential has a logarithmic singularity at $x=0$ and square-root behaviour at the edges of the gap. A Chebyshev expansion converges slowly near such points, while a charge at a small distance ϵ_0 reproduces the local behaviour with one term. This is also why the quadrature nodes of Section 5 avoid a neighbourhood of $x=0$.

Moving trial sequences. The best trial sequence depends on α. We use families

$$u(\tau) = u + \tau\,\mathrm{d}u, \qquad v(\tau) = v + \tau\,\mathrm{d}v \qquad (0 \le \tau \le 1),$$

with fixed finitely supported increments and with $\tau = \tau(\alpha)$ an explicit piecewise linear function. The fields X, Y, F_1 and F_2 are affine in τ. So the supremum of a field at τ is at most the chord between the suprema at $\tau = 0$ and $\tau = 1$, or, for the kept base fields, the supremum at $\tau = 0$ plus τ times the supremum of the increment. The norms are quadratic in τ. All bounds then become polynomials in α on each piece of $\tau(\alpha)$. Appendix B gives the charges, the increments and the certified bounds.

6.4. **The two cases at the improved rate.** Theorem 5.1 asks for the rate $U_0 = U + \Gamma_0/n_0^2$ instead of U, at the slope c. In both cases this reduces to finitely many polynomial inequalities in α, each of the form $V(\alpha) \le \sigma\alpha$ on a closed interval with polynomial caps. Each is proved by a cover of the interval by subintervals on which the leading term dominates, in exact rational arithmetic. The bound at an interior α of a cell uses the caps at the right end point of the cell, since $g > 0$. A finite cover by cells and the uniform continuity of the caps then give the bound for all α, with an error that tends to zero.

In the constrained case we use the increment with 24 Chebyshev coefficients and $\tau = \min(1, \frac{50}{11}\alpha)$. The inequality holds on $[0, 1]$, which is covered by the 6 pieces with end points $0, \frac{11}{50}, \frac{3}{10}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$. In the case with doubled interaction we use the increment with 40 Chebyshev coefficients and $\tau = 1$. The inequality of Lemma 6.1 holds on $[\alpha_1, 1]$, covered by the 6 pieces with end points $\alpha_1, \frac{1}{5}, \frac{3}{10}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$. For $\alpha < \alpha_1$ it fails, because the extra term $U_0 - U$ is absorbed by the term $c\alpha$ only once α is not small. So a new argument is needed only in the case with doubled interaction with fewer than $\alpha_1 n$ prescribed columns. Section 7 supplies it, with the same slope c. Precisely, it proves the following.

Proposition 6.2. *In the case with doubled interaction, for every configuration with $m \le \alpha_1 n$ prescribed columns at the nodes $t_{a_1}, \dots, t_{a_m}$, the integrand of Lemma 5.5, with the generalized Vandermonde determinant replaced by the bound of Lemma 5.6, is at most*

$$\exp\Big(n^2(-U_0 + \tfrac{1}{2}\log 2 + c\alpha + o(1)) - N\sum_j \hat{\delta}(t_{a_j})\Big)\, C_*^m \prod_j \mathcal{B}_\zeta(k_j), \qquad \alpha = m/n,$$

with an absolute constant C_*.

Proof of Theorem 5.1, given Proposition 6.2. By Proposition 6.2 and the discussion above, in both cases a determinant M_f with m prescribed columns $k_1, \dots, k_m$ at the nodes t_{a_j} is at most

$$C_*^m \prod_j \mathcal{B}_\zeta(k_j)\, e^{n_0 cNm - N\sum_j \hat{\delta}(t_{a_j})} \exp\big(n_0^2 N^2(-U_0 + \tfrac{1}{2}\log 2 + o(1))\big).$$

The factor C_*^m comes from the expansion of the row functions at the prescribed points. By (5.5) and §5.4, the weight of the column k_j at the node t_{a_j}, times $\mathcal{B}_\zeta(k_j)$, is at most $e^{(\psi+\hat\delta(t_{a_j})+o(1))N}$. The terms $\hat\delta$ cancel. So for every column k and every node a the weight and the factor of M_f that belong to the pair (k,a) have the product

$$\log(B_k C_k) \le \bigl(n_0 c + \psi + o(1)\bigr)N = (\rho + o(1))N,$$

and Lemma 5.4 applies with this bound for each pair. For $|\varepsilon| \le e^{-\rho N}$ Lemma 5.4 gives a factor $(1+O(N^5))^n = e^{o(N^2)}$. With $D = n_0^2(U_0 - \frac12 \log 2)$ this is the claim. □

7. A bound for the row determinant

This section proves Proposition 6.2, which completes the proof of Theorem 5.1. In the case with doubled interaction OpenAI's integrand contains a determinant $\det \mathsf{L}$ of size n that the energy method bounds by $e^{O(n\log n)}$. We replace this determinant by a simpler one, with a loss of only $e^{O(N)}$, and we bound the simpler one by a Gram determinant. The argument has three steps.

The first step is a comparison. Let $F(w)$ be the Pick matrix of the nodes with weight w on the positive nodes. Then $\det \mathsf{L}$ is $\det F(-5)$ up to a known factor, and a positivity argument bounds it by $499^n \det F(0)$ as soon as $F(-\frac{1}{100}) \ge 0$ (§7.1). The second step shows that $F(-\frac{1}{100}) \ge 0$ whenever a phase inequality holds for the nodes (§7.2, §7.3), and that this inequality holds on every configuration for which the energy bound is not already sufficient (§7.4). The third step bounds $\det F(0)$, which is a Vandermonde determinant of a Laurent system. It splits into a Vandermonde determinant of the positive nodes and a determinant of folded functions at the negative nodes. Fischer's inequality and Andréief's identity turn the latter into two log-gas integrals (§7.5). Reference measures then bound these integrals, and a numerical inequality closes the proof (§7.6).

7.1. The row determinant and the comparison.

In the case with doubled interaction OpenAI's integrand contains the determinant of the row functions at points $z_1,\dots,z_n \in (-1,1)\setminus\{0\}$, which we call the nodes [36, Prop. 5.2]. They are the free points x_i and the prescribed points y_j together. We write $\det\varphi(z)$ for this determinant. With $\mathsf{e} = n-1-2g$ it factors as

$$\det\varphi(z) = C_{n,g}\, V(z) \prod_a |z_a|^{-(n-1-g)} \det \mathsf{L}(z), \tag{7.1}$$

where $C_{n,g}$ depends only on n, g and the number of nodes of each kind, and $V(z) = \prod_{a<b}(z_b - z_a)$. To describe L let

$$B(z) = \prod_a \frac{z - z_a}{1 - z_a z}, \qquad \mathcal{K}_{ab} = \frac{1}{1 - z_a z_b}, \qquad D_a = B'(z_a) z_a^{\mathsf{e}},$$

so that B is the Blaschke product of the nodes and $\mathcal{K}$ is the Pick matrix of the Szegő kernels. Let E_- and E_+ be the diagonal blocks of $\operatorname{diag}(D_a)$ at the negative and the positive nodes. For $w \in \mathbb{R}$ let

$$F(w) = \mathcal{K} + \operatorname{diag}(E_-, wE_+).$$

The computation of [36, Prop. 5.2] gives, for every real vector (w_a), the identity

$$\det\bigl(z_a^{g-r} + w_a z_a^{r-g}\bigr)_{a,\,0\le r<n} = c(z)\det\bigl(\mathcal{K} + \operatorname{diag}(w_a D_a)\bigr), \tag{7.2}$$

where $c(z)$ does not depend on w. If all nodes are free, the rows of OpenAI's determinant are $\frac{1}{2}(z^{g-r} + z^{r-g})$ at a negative node and $-\frac{1}{4}z^{g-r} + \frac{5}{4}z^{r-g}$ at a positive node. So $\det \mathsf{L} = \det F(-5)/\det \mathcal{K}$. The value $-5 = 1 - 6$ comes from the factor $\frac{3}{2}$ in the row functions. OpenAI's argument uses only $|\det \mathsf{L}| \le e^{O(n\log n)}$, which is negligible at the scale e^{N^2}. Numerical evaluation suggests that $\log|\det \mathsf{L}|$ is of order $-N^2$ near the extremal configuration (Remark 9.1).

A prescribed point has the row $\frac{1}{2}(z^{g-r} + z^{r-g})$ at either sign. At a positive prescribed point this row is -2 times the row of a free positive node plus 3 times the row $(z^{r-g})_r$. Let Ξ be the matrix of all rows at the nodes. Expanding its determinant in these rows gives

$$|\det \Xi| \le \sum_S 2^{m_+ - |S|} 3^{|S|} |\det \Xi_S|, \tag{7.3}$$

where m_+ is the number of positive prescribed points and S runs over the subsets of these points. In Ξ_S a node of S has the row $(z^{r-g})_r$, a negative node has the row $\frac{1}{2}(z^{g-r} + z^{r-g})$, and every other positive node has the row of a free positive node. The total weight in (7.3) is at most 5^m. A node of S is what Section 6 calls a prescribed point of the second kind. Let Ξ_S^0 be the same matrix with the rows of the positive nodes outside S replaced by $(z^{g-r})_r$.

Both Ξ_S and Ξ_S^0 are obtained from (7.2). The left side of (7.2) is affine in each w_a. Taking the coefficient of $\prod_{a\in S} w_a$ gives

$$\frac{|\det \Xi_S|}{|\det \Xi_S^0|} = 4^{-|P\setminus S|}\frac{|\det F(-5)_{S^c}|}{\det F(0)_{S^c}}, \tag{7.4}$$

where P is the set of positive nodes and the index S^c denotes the principal submatrix on the complement of S. The case $w = 1$ is a Pick matrix.

Lemma 7.1. *Let $H = \operatorname{diag}(B'(z_a))$ and let X be a real diagonal matrix. If $\mathcal{K} - X\mathcal{K}X \ge 0$, then $\mathcal{K} + HX \ge 0$. In particular $F(1) \ge 0$ for every configuration.*

Proof. First we show that $\mathcal{K}H^{-1}\mathcal{K} = H$. For a real vector c put $\phi_c(z) = B(z)\sum_a c_a/(z - z_a)$. It has node values $\phi_c(z_a) = c_a B'(z_a)$. Partial fractions give $\phi_c = \sum_a d_a k_a$ with $k_a(z) = 1/(1 - z_a z)$, for a real vector d, and comparing node values gives $\mathcal{K}d = Hc$. On the unit circle $|\phi_c(z)| = |\sum_a c_a/(z - z_a)| =$

$|\sum_a c_a k_a(\bar z)|$. So $\|\phi_c\|^2_{H^2} = c^{\mathsf T}\mathcal K c$, and also $\|\phi_c\|^2_{H^2} = d^{\mathsf T}\mathcal K d = c^{\mathsf T} H\mathcal K^{-1}Hc$. Hence $H\mathcal K^{-1}H = \mathcal K$, which is the claim.

Let $T = H^{-1}\mathcal K X$ and $A = I + T$. Then $\mathcal K T = HX$, $T^{\mathsf T}\mathcal K = XH$ and $T^{\mathsf T}\mathcal K T = X\mathcal K X$. So

$$0 \le A^{\mathsf T}\mathcal K A = \mathcal K + 2HX + X\mathcal K X.$$

Adding $\mathcal K - X\mathcal K X \ge 0$ gives $2(\mathcal K + HX) \ge 0$. For the second statement take $X = \operatorname{diag}(z_a^{\mathsf e})$. The entries of $\mathcal K - X\mathcal K X$ are $\sum_{j<\mathsf e}(z_a z_b)^j$, which is a Gram matrix. □

Lemma 7.2. *If $F(-\frac{1}{100}) \ge 0$, then $|\det F(-5)_T| \le 499^{|T|}\det F(0)_T$ for every set T of nodes. Hence $|\det \Xi_S| \le 499^n |\det \Xi^0_S|$ for every S.*

Proof. Put $Y = F(0)$ and $Q = \operatorname{diag}(0, E_+)$. By hypothesis and Lemma 7.1, $Y + Q \ge 0$ and $Y - \frac{1}{100}Q \ge 0$. A convex combination of the two gives $Y \ge 0$. Let $Z = F(-5) = Y - 5Q$. Then

$$499Y - Z = 5(Y+Q) + 493Y \ge 0, \qquad 499Y + Z = 500\big(Y - \tfrac{1}{100}Q\big) \ge 0.$$

So $-499Y \le Z \le 499Y$, and the same holds for principal submatrices. If $Y_T > 0$, all eigenvalues of $Y_T^{-1/2} Z_T Y_T^{-1/2}$ have modulus at most 499, which is the claim. The singular case follows by replacing Y_T by $Y_T + tI$ and letting $t \to 0$. The second statement follows from (7.4). □

So everything reduces to two questions: when is $F(-\frac{1}{100}) \ge 0$, and how large is $\det \Xi^0_S$?

7.2. **A phase criterion for positivity.** The matrix $F(-c)$ is a Pick matrix with a small negative weight on the positive nodes. The following theorem gives a criterion for its positivity that depends only on the positions of the positive nodes. It holds for all configurations and has no relation to our parameters.

Theorem 7.3. *Let $z_1, \dots, z_n \in (-1,1)\setminus\{0\}$ be distinct, with k_+ positive and k_- negative entries. Let $t \ge 0$ be an integer, let $c = \frac{1}{100}$, and let $D_a = B'(z_a) z_a^t$. For a positive node p put $\rho_p = (1-p)/(1+p)$. If*

$$t\arctan 4 - \sum_{p>0} \arctan(4\rho_p) > \frac{\pi}{2}(k_- + 1) + \frac{17}{4}\,\frac{c^2}{1-c^2}\,k_+, \tag{7.5}$$

then $\mathcal K + \operatorname{diag}(E_-, -cE_+) \ge 0$.

The number $\arctan 4$ is the half-angle of the arc of the unit circle around $z = -1$ with end point $z_0 = (4i-1)/(4i+1)$. The proof compares two phases on this arc. We apply the theorem with $t = \mathsf e$.

7.3. **Proof of Theorem 7.3.** Let $f_a(\gamma) = \gamma z_a^t$ at a negative node and $f_a(\gamma) = -cz_a^t$ at a positive node. Let $W(\gamma)_{ab} = (1 - f_a(\gamma)f_b(\gamma))\mathcal{K}_{ab}$ be the Pick matrix of these data. By Lemma 7.1 it suffices to prove $W(1) \geq 0$. With $X = \operatorname{diag}(z_a^t)$ we have

$$W(-c) = (1 - c^2)\mathcal{K} + c^2(\mathcal{K} - X\mathcal{K}X),$$

and $\mathcal{K} - X\mathcal{K}X \geq 0$ as in Lemma 7.1. So $W(-c) > 0$. If $W(1)$ is not positive semidefinite, then there is $\gamma \in (-c, 1)$ with $W(\gamma) \geq 0$ singular. Fix such a γ and write $W = W(\gamma)$, $f_a = f_a(\gamma)$. Note that $|f_a| < 1$.

Let $\mathcal{T}$ be a minimal set of nodes such that the principal submatrix $W_{\mathcal{T}}$ is singular. Since the diagonal entries of W are positive, $|\mathcal{T}| \geq 2$. By minimality $W_{\mathcal{T}}$ has a kernel vector v with all entries nonzero, and every proper principal submatrix is positive definite. From now on only the nodes of $\mathcal{T}$ are used, and k_+ and k_- count these nodes. Inequality (7.5) still holds, because removing a node lowers the right side or raises the left side. If all nodes of $\mathcal{T}$ have the same sign, the data are a constant of modulus less than 1 times z^t, and $W_{\mathcal{T}} > 0$ as above. So $k_+ \geq 1$ and $k_- \geq 1$.

We construct a Blaschke product that interpolates the data. Put

$$\varphi(\zeta) = \sum_{a\in\mathcal{T}} \frac{v_a}{1 - z_a\zeta}, \qquad \psi(\zeta) = \sum_{a\in\mathcal{T}} \frac{v_a f_a}{1 - z_a\zeta}.$$

The relation $Wv = 0$ says that $\varphi(z_b) = f_b\psi(z_b)$ for $b \in \mathcal{T}$. For $|\zeta| < 1$ let $\beta_a = (z_a - \zeta)/(1 - z_a\bar{\zeta})$. From $(1 - z_a\bar{\zeta})(1 - z_b\zeta) - (z_a - \zeta)(z_b - \bar{\zeta}) = (1 - z_az_b)(1 - |\zeta|^2)$ we get

$$\frac{1 - |\zeta|^2}{(1 - z_a\bar{\zeta})(1 - z_b\zeta)} = \frac{1 - \beta_a\bar{\beta}_b}{1 - z_az_b}.$$

Multiplying by $v_av_b(1 - f_af_b)$ and summing gives

$$(1 - |\zeta|^2)(|\psi(\zeta)|^2 - |\varphi(\zeta)|^2) = u^*W_{\mathcal{T}}u, \qquad u_a = v_a\bar{\beta}_a, \tag{7.6}$$

because $v^{\mathsf{T}}W_{\mathcal{T}}v = 0$. So $|\varphi| \leq |\psi|$ in the disc. For $|\zeta| = 1$ one has

$$\frac{1}{(1 - z_a\bar{\zeta})(1 - z_b\zeta)} = \frac{1}{1 - z_az_b}\Big(\frac{1}{1 - z_a\bar{\zeta}} + \frac{1}{1 - z_b\zeta} - 1\Big),$$

and summation with $W_{\mathcal{T}}v = 0$ gives $|\varphi| = |\psi|$ on the circle. Since the functions $1/(1 - z_a\zeta)$ are independent, $\varphi \neq 0$ and so $\psi \neq 0$. Hence $h = \varphi/\psi$ is a rational function with real coefficients, with $|h| \leq 1$ in the disc and $|h| = 1$ on the circle. So it is a finite Blaschke product. Both φ and ψ are polynomials of degree at most $|\mathcal{T}| - 1$ divided by $\prod_a(1 - z_a\zeta)$, so the degree of h is at most $|\mathcal{T}| - 1$. Thus

$$h = \varepsilon z^s\frac{P}{P^*}, \qquad \varepsilon = \pm 1, \quad P^*(z) = z^{\deg P}P(1/z),$$

where P is a real polynomial with all roots in the open disc and $P(0) \neq 0$. Put $d = s + \deg P$. Then $d + 1 \leq k_+ + k_-$.

We have $h(z_b) = f_b$ for $b \in \mathcal{T}$. This is clear if $\psi(z_b) \neq 0$. If $\psi(z_b) = 0$, then $\varphi(z_b) = 0$, and (7.6) at $\zeta = z_b$ gives $u^* W_{\mathcal{T}} u = 0$ with $u_b = 0$ and $u \neq 0$. This contradicts the positive definiteness of $W_{\mathcal{T}\setminus\{b\}}$. We show that h cannot exist.

The case $t \leq s$. The real polynomial $\varepsilon z^{s-t} P + cP^*$ is nonzero, because $|P| = |P^*|$ on the unit circle and $c < 1$. It vanishes at the k_+ positive nodes and has degree at most $d - t$. So $k_+ + t \leq d \leq k_+ + k_- - 1$, that is, $t < k_-$. This contradicts (7.5), which gives $t \arctan 4 > \frac{\pi}{2} k_-$.

The case $s < t$. Put $T = t - s \geq 1$ and $r = h/z^t = \varepsilon P/(z^T P^*)$. Then $r = \gamma$ at the k_- negative nodes and $r = -c$ at the k_+ positive ones. On the unit circle $zh'/h = t + \Omega$, with $\Omega = zr'/r$ real.

Lemma 7.4. *$\Omega \geq 0$ on the unit circle.*

Proof. Let $v = \deg P$. The numerator of zr'/r is the self-reciprocal real polynomial $zP'P^* - zP^{*\prime}P - TPP^*$ of degree at most $2v$. So it equals $z^v R(z + 1/z)$ with $\deg R \leq v \leq k_+ + k_- - 1$, and $\Omega = R(2\cos\theta)/|P|^2$ at $z = e^{i\theta}$. Rolle's theorem gives a critical point of r between consecutive nodes of the same sign, hence $k_+ + k_- - 2$ distinct roots of R outside $[-2, 2]$. So R has at most one further root. Suppose that $R < 0$ somewhere in $[-2, 2]$. If $R(2) \geq 0$ and $R(-2) \geq 0$, then R has two further roots in $[-2, 2]$. If $R(2) < 0$, then $|r|$ decreases at $z = 1$ on the real axis, so $|r| > 1$ just to the left of 1. Since $|r| < 1$ at the largest node, r takes the value $r(1) = \pm 1$ on the right of that node. Rolle's theorem gives one further critical point, and a further root of R. The same holds at -2. If exactly one of $R(\pm 2)$ is negative, this root comes with a sign change of R in $[-2, 2]$. If both are negative, there are two such roots. In every case R has two further roots, which contradicts the degree bound. □

Let $p_1, \ldots, p_m$ be the positive nodes, so $m = k_+$ in the rest of this proof, and let $k = k_-$. We compare the phase of h with the phase of an explicit interpolant of degree m. Put $\lambda = c^2$ and $w_a = p_a^t$. Put $\mathsf{P}_{ab} = (1 - \lambda w_a w_b)/(1 - p_a p_b)$ and $M = \mathsf{P}^{-1}$, and write $\langle a, b\rangle = a^{\mathsf{T}} M b$. Let $\xi_z(a) = 1/(1 - z p_a)$, $g_1 = 1$ and $g_2 = -cw$. With $F_i(z) = g_i \xi_z$, $C_j = g_j \xi_{-1}$ and $J = \operatorname{diag}(1, -1)$ define

$$\Theta_{ij}(z) = \delta_{ij} - (1 + z) J_j \langle F_i(z), C_j\rangle.$$

So $\Theta(-1) = I$. Direct expansion gives the kernel identity

$$\sum_k \Theta_{ik}(z) J_k \Theta_{jk}(u) = \delta_{ij} J_i - (1 - zu)\langle F_i(z), F_j(u)\rangle. \tag{7.7}$$

The vector $F_1(p_a) - f_a F_2(p_a)$ is row a of P, where now $f_a = -cw_a$. This gives the node relation $\Theta_{1j}(p_a) = f_a \Theta_{2j}(p_a)$. Let $D(z) = \prod_a (1 - p_a z)$, $b(z) = \prod_a (z - p_a)$ and $\kappa = (-1)^m$. The functions $e_{ij} = D\Theta_{ij}$ are real polynomials of degree at most m. Their determinant vanishes at the p_a by the node relation and at the $1/p_a$ because $(e_{ij}(1/p_a))$ has rank one. So it

equals κbD. Taking $u = 1/z$ in (7.7) gives

$$e_{11}(z) = \kappa z^m e_{22}(1/z), \qquad e_{12}(z) = \kappa z^m e_{21}(1/z).$$

On the unit circle (7.7) with $i = j = 2$ and $u = \bar{z}$ gives $|\Theta_{22}|^2 = 1+|\Theta_{21}|^2$. Let $\sigma = h(-1) = \pm 1$ and $e = \sigma e_{21} + e_{22}$. Then $W = e/D = \Theta_{22}(1 + \sigma\Theta_{21}/\Theta_{22})$ does not vanish on the circle, and $W(-1) = 1$.

The size of the argument of Θ_{22}. Let T_l be Θ_{22} with λ replaced by l, so $T_0 = 1$. Let $\mathcal{K}^{(p)}$ be the matrix with entries $1/(1 - p_a p_b)$, let $Z = \operatorname{diag}(w_a)$, and let $M_l = (\mathcal{K}^{(p)} - lZ\mathcal{K}^{(p)}Z)^{-1}$. For $0 \le l < l' < 1$ and real vectors a, b the resolvent identity gives

$$l'\, a^{\mathsf{T}} Z M_{l'} Z b - l\, a^{\mathsf{T}} Z M_l Z b = (l' - l)(M_{l'} Z a)^{\mathsf{T}} \mathcal{K}^{(p)} (M_l Z b).$$

Since $\mathcal{K}^{(p)} - Z\mathcal{K}^{(p)}Z \ge 0$, the right side has modulus at most $(l' - l)\|a\|\|b\|/((1 - l)(1 - l'))$, where $\|a\|^2 = a^{\mathsf{T}}(\mathcal{K}^{(p)})^{-1}a$. For $c = 0$ the identity (7.7) at $z = \zeta$ and $u = \bar{\zeta}$ gives $(1 - |\zeta|^2)\|\xi_\zeta\|^2 = 1 - |b(\zeta)/D(\zeta)|^2$. Since $1 - \prod t_a \le \sum(1 - t_a)$ for $t_a \in [0, 1]$, this yields

$$\|\xi_\zeta\|^2 \le \sum_a \frac{1 - p_a^2}{|1 - p_a\zeta|^2} \qquad (|\zeta| \le 1).$$

In particular $\|\xi_{-1}\|^2 \le m$. For z on the arc from z_0 to -1 one has $\operatorname{Re} z \le \frac{15}{17}$, and a quadratic inequality in p gives $|1 + z|^2(1 - p^2)/|1 - pz|^2 \le (\frac{17}{4})^2$. So $|1 + z|\,\|\xi_z\|\,\|\xi_{-1}\| \le \frac{17}{4}m$, and $|T_{l'}(z) - T_l(z)| \le \frac{l'-l}{(1-l)(1-l')} \cdot \frac{17}{4}m$ on the arc. Since $|T_l| \ge 1$, a telescoping product over a fine grid of $[0, \lambda]$ shows that the continuous argument of Θ_{22} on the arc, normalized to vanish at -1, has modulus at most $\mathcal{X} = \frac{17}{4}m\lambda/(1 - \lambda)$. The factor $1 + \sigma\Theta_{21}/\Theta_{22}$ has positive real part. So the continuous argument $\ell(\theta)$ of $W(e^{i\theta})$, with $\ell(\pi) = 0$, satisfies $|\ell(\theta_0)| \le \frac{\pi}{2} + \mathcal{X}$, where $z_0 = e^{i\theta_0}$ and $\pi - \theta_0 = 2\arctan 4$.

The phases. By Lemma 7.4, $h(e^{i\theta}) = \sigma e^{-i\phi(\theta)}$ with $\phi(\theta) = \int_\theta^\pi (t + \Omega) \ge t(\pi - \theta)$. Let $\psi(\theta) = m(\pi - \theta) + 2\sum_a \arg(1 - p_a e^{i\theta})$. Then $(-z)^m\bar{e}/e = \exp(-i(\psi + 2\ell))$. At θ_0, $1 - pz_0 = (1 + p)(1 + 4i\rho_p)/(1 + 4i)$. So $\psi(\theta_0) = 2\sum_p \arctan(4\rho_p)$.

The root count. Let $\hat{e}(z) = z^m e(1/z)$ and

$$\mathcal{N} = \varepsilon z^s P e - \sigma\kappa P^* \hat{e},$$

a real polynomial of degree at most $d+m$. Put $a = \phi - \psi - 2\ell$, so $a(\pi) = 0$. At every θ with $a(\theta) \in 2\pi\mathbb{Z}$ the polynomial $\mathcal{N}$ vanishes at $e^{\pm i\theta}$. It also vanishes at every p_a by the node relation and the reversal formulas, at every $1/p_a$ because $z^{d+m}\mathcal{N}(1/z) = \pm\mathcal{N}(z)$, and at -1. Suppose that $a(\theta_0) > \pi k$ and let $L = \lfloor k/2 \rfloor$. If $\mathcal{N} \ne 0$, the intermediate value theorem gives L conjugate pairs of roots on the open arc. So $2L + 2m + 1 \le d + m \le 2m + k - 1$, which contradicts $2L + 2 > k$. If $\mathcal{N} = 0$, then $e^{ia} = 1$ on the arc, which contradicts the intermediate value theorem at the value π. Hence $a(\theta_0) \le \pi k$, and

$$2t \arctan 4 \le \phi(\theta_0) \le \pi k + 2\sum_p \arctan(4\rho_p) + \pi + 2\mathcal{X}.$$

This contradicts (7.5), and Theorem 7.3 is proved. □

7.4. **The alternative on near-extremal configurations.** We now show that (7.5) holds with $t = \mathsf{e}$ and $c = \frac{1}{100}$ on every configuration for which the energy bound of Section 6 does not already suffice. The idea is that such a configuration has Chebyshev moments close to the trial sequence u, and the left side of (7.5) is then computed by a fixed test function.

Let $\phi_H(w) = \frac{\pi}{2} - \arctan(4(1-w)/(1+w))$ for $0 \le w < 1$. Then $\arctan(4\rho_p) = \frac{\pi}{2} - \phi_H(p)$. With $\mathsf{e} = 48N - 1$, $n = 58N$ and $k_- = n - k_+$, the inequality (7.5) follows for large N from

$$\frac{1}{n}\sum_{p>0}\phi_H(p) \ge \arctan\tfrac{1}{4} + \tfrac{5}{29}\arctan 4 + \tfrac{17}{4}\,\frac{c^2}{1-c^2} + \xi, \qquad \xi = \tfrac{1}{1000}, \tag{7.8}$$

whose right side is $0.474992\ldots$. Let $f = \frac{F_0}{2} + \sum_{k\ge1} F_k T_k$ be a polynomial of degree 32 with dyadic coefficients and denominator 2^{31}, where $F_0 = 992223195/2^{30}$. It satisfies $f \le \phi_H$ on $(0,1)$, $f \le 0$ on $[-1,0]$, and

$$\tfrac{1}{4}F_0 + \tfrac{1}{4}|F_1|y \le \phi_H(y) \qquad (0 < y < 1). \tag{7.9}$$

Also $\sum_k kF_k^2 \le 7877/10000$. On this range of α the case with doubled interaction uses the moving trial sequence $u(\tau) = u + \tau\,\mathrm{d}u$ of §6.3, with the increment of 40 Chebyshev coefficients and $\tau = \alpha/\alpha_1$. We have $\sum_k F_k u_k \ge 879/10000$ and $\sum_k F_k \mathrm{d}u_k \ge -0.010526$.

Fix a configuration and a set S. Let μ_F, μ_M and μ_S be the measures of §6.2, with the points of S as the prescribed points of the second kind, and let β be the mass of μ_S. By (6.1), the bound of Lemma 6.1 with $\alpha_2 = \beta$ improves by the number Δ of (6.1). Divide $[0,\alpha_1]$ into 64 equal cells. On the cell j the left side of the inequality of Lemma 6.1, minus the right side at the slope c, is at most a rational number Δ_j. These numbers lie between 10^{-6} and 0.008365. So either $\Delta \le \Delta_j$, or the energy bound gives the conclusion of Proposition 6.2 for Ξ_S directly.

Suppose that $\Delta \le \Delta_j$. Since $\phi_H \ge 0$ on $(0,1)$ and $f \le \phi_H$, while $f \le 0$ on the negative half, we have $\frac{1}{2}f \le \phi_H$ on positive prescribed points and $\frac{1}{2}f \le 0$ on negative ones. With (7.9),

$$\begin{aligned}\frac{1}{n}\sum_{p>0}\phi_H(p) &\ge \int f\,d(\mu_F + \tfrac{1}{2}\mu_M) + \tfrac{1}{4}F_0\beta + \tfrac{1}{4}|F_1|A_1(\mu_S)\\ &\ge \tfrac{1}{2}F_0 - \tfrac{1}{4}F_0\alpha + \sum_k F_k u_k(\tau) - \sqrt{\tfrac{1}{2}\Delta_j\sum_k kF_k^2},\end{aligned}$$

by the Cauchy–Schwarz inequality. The right side is affine in α on each cell. At the end points of every cell it exceeds 0.4752, which is larger than the right side of (7.8). The smallest margin is 0.0050. So (7.8) holds, and Theorem 7.3 gives $F(-\frac{1}{100}) \ge 0$. All the numerical facts of this subsection are part of the formal proof, which also carries the regularization of the moments that we have suppressed here. We have proved the following.

Proposition 7.5. *For N large, every configuration with $m \le \alpha_1 n$ prescribed columns in the case with doubled interaction, and every S, satisfies at least one of two statements. Either the energy bound for Ξ_S gives the conclusion of Proposition 6.2, or $F(-\frac{1}{100}) \ge 0$ and therefore $|\det \Xi_S| \le 499^n |\det \Xi_S^0|$.*

7.5. **Gram determinants.** It remains to bound $\det \Xi_S^0$. Let $\mathsf{D} = n-1-g = 53N-1$. Let $\xi_1, \dots, \xi_{k_+}$ be the positive nodes, with $\xi = z$ outside S and $\xi = 1/z$ on S, so that all these rows are $(\xi^{g-r})_r$. Let $x_1, \dots, x_{k_-}$ be the negative nodes, and put $u_b = J(x_b)$, where $J(z) = \frac{1}{2}(z + 1/z)$. Let $s = \mathsf{D} + 1 - k_-$ and $\omega(z) = \prod_p (z - \xi_p)$. If $k_- > \mathsf{D}+1$, the determinant vanishes, because the negative rows lie in a space of dimension $\mathsf{D}+1$. So we may assume $s \ge 0$. The Laurent polynomials with exponents in $[-\mathsf{D}, g]$ that vanish at all ξ_p are spanned by the functions $z^{i-\mathsf{D}}\omega(z)$ with $0 \le i < k_-$. So elimination gives

$$|\det \Xi_S^0| = \prod_p |\xi_p|^{-\mathsf{D}} \, |V(\xi)| \, \left|\det\left(\mathcal{F} z^{i-\mathsf{D}} \omega\,(x_b)\right)_{i,b}\right|, \tag{7.10}$$

where $\mathcal{F}f(x) = \frac{1}{2}(f(x) + f(1/x))$ is the fold. The fold sends z^d to $T_{|d|}(J(z))$. So the fold of a Laurent polynomial of degree at most D is a polynomial of degree at most D in $u = J(z)$.

We measure these functions in a finite Laurent-orthogonal measure. Let $\nu = \nu_N$ be the positive measure on 16 concentric circles of radii $r_j = e^{-\sigma_j}$, with total masses $e^{-N\tau_j}$, which is uniform on the roots of unity of order $2\mathsf{D}+1$ on each circle (scaled by r_j). We assume that $\sigma_0 = \tau_0 = 0$ and $10\sigma_j \le \tau_j$ for all j. The Laurent monomials with exponents in $[-\mathsf{D}, \mathsf{D}]$ are orthogonal in $L^2(\nu)$, with $\|z^c\|^2 = A_c = \sum_j e^{-N\tau_j} r_j^{2c}$. Let $G_0 = A_0$ and $G_a = (A_a + A_{-a})/4$ for $a \ge 1$, which is the squared norm of $T_a(J(z))$.

Proposition 7.6. *Let ℓ_a be the leading coefficient of T_a, and suppose that the numbers $\theta_d \ge 1$ for $-\mathsf{D} \le d \le g$ satisfy $G_a \sum_d (A_d \theta_d)^{-1} \le 1$ for each $a \le \mathsf{D}$, the sum running over the d with $|d| = a$. Then*

$$(\det \Xi_S^0)^2 \prod_{a=0}^{D} G_a \cdot k_-!\, s! \le \Big(\prod_p \xi_p^{-\mathsf{D}} V(\xi)\Big)^2 \Big(\prod_{a=0}^{\mathsf{D}} \ell_a \, V(u)\Big)^2 \cdot \prod_d \theta_d \cdot Z_+ Z_-,$$

where

$$Z_+ = \int \prod_i |z_i|^{-2\mathsf{D}} |\omega(z_i)|^2 |V(z)|^2 \, d\nu^{k_-},$$

$$Z_- = \int \prod_l |\pi(Jz_l)|^2 |V(Jz)|^2 \, d\nu^{s},$$

with $\pi(u) = \prod_b (u - u_b)$.

Proof. Let $\mathcal{P}$ be the space of polynomials of degree at most D in u, with the inner product for which the T_a are orthogonal of squared norm G_a. Consider the two families $f_i = \mathcal{F} z^{i-\mathsf{D}}\omega$ for $0 \le i < k_-$ and πu^l for $0 \le l < s$. They have $\mathsf{D}+1$ members together. The second family spans the kernel of the evaluation at the u_b. So the determinant of the coefficients of the joint family in the basis (T_a) equals $\det(f_i(x_b))$ divided by $\prod_a \ell_a \cdot V(u)$, up to sign. Multiplying its square by $\prod_a G_a$ gives the Gram determinant of the joint family. Fischer's inequality bounds it by the product of the Gram determinants of the two families.

Let $f = \sum_d f_d z^d$ be a combination of the functions $z^{i-\mathsf{D}}\omega$. The coefficient of T_a in $\mathcal{F}f$ is $\sum_{|d|=a} f_d$, and the Cauchy–Schwarz inequality with the hypothesis gives $\|\mathcal{F}f\|^2 \le \sum_d \theta_d A_d f_d^2$. Raising the weights of a diagonal form by factors $\theta_d \ge 1$ raises a Gram determinant by at most $\prod_d \theta_d$, by the Cauchy–Binet formula. So the first Gram determinant is at most $\prod_d \theta_d$ times the Gram determinant of the functions $z^{i-\mathsf{D}}\omega$ in $L^2(\nu)$. Andréief's identity [17] in its finite form writes the two Gram determinants as $Z_+/k_-!$ and $Z_-/s!$. The inequality follows with (7.10). □

The constants θ_d and G_a. Put $\theta_d = 1$ for $d \le 0$ and $\theta_d = A_{-d}/A_d$ for $1 \le d \le g$. For $a > g$ the sum in the hypothesis has the single term $d = -a$, and $G_a/A_{-a} \le \frac12$ because $A_a \le A_{-a}$. For $1 \le a \le g$ the left side of the hypothesis is $\frac12(A_a + A_{-a})/A_{-a} \le 1$. Since $g = 5N$ and $10\sigma_j \le \tau_j$, each term of A_{-a} is at most 1, so $A_{-a} \le 16$ for $a \le g$. Also $A_a \ge 1$ by the term $j = 0$. Hence $\prod_d \theta_d \le 16^{5N}$. Let $\tau^*(q) = \max_j(2q\sigma_j - \tau_j)$. Then $G_a \ge \frac14 e^{N\tau^*(a/N)}$ for $a \ge 1$. Since τ^* is nondecreasing,

$$\sum_{a=0}^{\mathsf{D}} \log G_a \ge N^2 \int_0^{53} \tau^*(q)\,dq - 2\mathsf{D}\log 2 - N\tau^*(53). \tag{7.11}$$

The factor $\prod_a \ell_a^2$ equals $2^{\mathsf{D}(\mathsf{D}-1)}$.

A tangent bound for Vandermonde determinants. The integrals $Z_\pm$ are compared with reference measures by the following lemma, which is the negative definiteness of the logarithmic energy on neutral measures in a form exact for finitely many points. Let $d(z,w)^2 = |z-w|^2/((1+|z|^2)(1+|w|^2))$ and $\mathsf{h}(z) = \frac12 \log(1+|z|^2)$. For a finite positive measure q on $\mathbb{C}$ write $U^q(z) = \int \log|z-w|\,dq(w)$ and $I(q) = \int U^q\,dq$. Let $H_L = \sum_{j\le L} 1/j$.

Lemma 7.7. *Let q be a finite positive measure on $\mathbb{C}$ with $\int \mathsf{h}\,dq < \infty$ and $q\{w : d(z,w)^2 \le t^4\} \le Ct$ for all z and all $0 < t \le 1$. For every $\eta > 0$ there is L such that for every $m \ge 0$, all $\nu_0 \ge 0$ with $\nu_0 q(\mathbb{C}) = m$ and all $z_1, \dots, z_m \in \mathbb{C}$,*

$$|V(z)|^2 \le \exp\Bigl(2\nu_0 \sum_i U^q(z_i) - \nu_0^2 I(q) + 2\nu_0 m\eta + \tfrac{m}{2} H_L\Bigr).$$

Proof. We may assume that the z_i are distinct. Let $\mathsf{n}(z)$ be the stereographic image of z on the unit sphere. Then $\mathsf{C}(z,w) = 1 - d(z,w)^2 = \frac12(1 + \mathsf{n}(z)\cdot\mathsf{n}(w))$

is a sum of four products $b_i f_i(z) f_i(w)$ with $b_i \geq 0$, and so is every power of C, with more terms. Let $K_L = -\frac{1}{2}\sum_{j\leq L} \mathsf{C}^j/j$. Then $K_L = -\sum_\phi b_\phi F_\phi \otimes F_\phi$ is a finite sum with $b_\phi \geq 0$. For real $S_\phi = \sum_i F_\phi(z_i)$ and $M_\phi = \int F_\phi \, dq$ the inequality $(S_\phi - \nu_0 M_\phi)^2 \geq 0$ gives, after summation,

$$\sum_{i,j} K_L(z_i, z_j) \leq 2\nu_0 \sum_i \int K_L(z_i, w)\, dq(w) - \nu_0^2 \iint K_L \, dq\, dq. \tag{7.12}$$

For $z \neq w$ we have $\log d(z,w) = \frac{1}{2}\log(1-\mathsf{C}) = -\frac{1}{2}\sum_{j\geq 1} \mathsf{C}^j/j \leq K_L(z,w)$, and $K_L(z,z) = -\frac{1}{2}H_L$. So $\sum_{i\neq j} \log d(z_i, z_j)$ is at most the left side of (7.12) plus $\frac{m}{2}H_L$. Also $\iint K_L \geq \iint \log d$. For the remaining term we show that $\int K_L(z,\cdot)\, dq \leq \int \log d(z,\cdot)\, dq + \eta$ for large L, uniformly in z. Fix $h \in (0,1)$. On $\{d^2 \geq h\}$ the difference $K_L - \log d = \frac{1}{2}\sum_{j>L} \mathsf{C}^j/j$ is at most $(1-h)^L/(2Lh)$. On $\{d^2 < h\}$ it is at most $-\log d$, and the ball hypothesis on the shells $2^{-4(i+1)} < d^2 \leq 2^{-4i}$ bounds the integral of $-\log d$ over this set by a quantity that tends to 0 with h, uniformly in z. This proves the claim.

Finally $\log|z-w| = \log d(z,w) + \mathsf{h}(z) + \mathsf{h}(w)$. So $\sum_{i\neq j} \log|z_i - z_j| = \sum_{i\neq j} \log d(z_i, z_j) + 2(m-1)\sum_i \mathsf{h}(z_i)$. Expanding U^q and $I(q)$ in the same way, and using $\nu_0 q(\mathbb{C}) = m$, the terms in $\int \mathsf{h}\, dq$ cancel. What remains is $2m\sum_i \mathsf{h}(z_i)$ on the right side against $2(m-1)\sum_i \mathsf{h}(z_i)$ on the left side. Since $\mathsf{h} \geq 0$, this proves the lemma. □

Apply the lemma to the $a+c$ points consisting of a fixed base points b_i and c points Z_l. If Z_l runs over a finite set with weights $\mathrm{wt} \geq 0$, summing over tuples gives, with $\nu_0 q(\mathbb{C}) = a + c$ and the same η and L,

$$|V(b)|^2 \sum_Z \prod_l \mathrm{wt}(Z_l) \prod_{l,i} |Z_l - b_i|^2 \, |V(Z)|^2 \tag{7.13}$$
$$\leq \exp\Big(2\nu_0 \sum_i U^q(b_i) - \nu_0^2 I(q) + 2\nu_0(a+c)\eta + \tfrac{a+c}{2} H_L\Big) \Big(\sum_z \mathrm{wt}(z) e^{2\nu_0 U^q(z)}\Big)^c.$$

The exponential on the right does not depend on the Z_l, and the last sum separates them. This is the reason for completing the gas by the base points: the determinant of the base points is then bounded together with the gas.

The two gases. Let q_+ be a measure of mass 1 in the plane of z, and q_- a measure of mass $\frac{53}{58}$ in the plane of u, both satisfying the hypotheses of Lemma 7.7. Let κ_+ and c_- be constants. We assume the angular bounds

$$U^{q_+}(r_j e^{i\theta}) + \tfrac{53}{58}\sigma_j - \tfrac{\tau_j}{116} \leq \kappa_+, \qquad U^{q_-}\big(J(r_j e^{i\theta})\big) - \tfrac{\tau_j}{116} \leq c_-, \tag{7.14}$$

for all j and θ. Apply (7.13) to Z_+ with $q = q_+$, $\nu_0 = n$, base points ξ_p and the k_- gas points z_i. Since $r_j^{-2\mathsf{D}} \leq e^{2\cdot 53 N\sigma_j}$, the sum in (7.13) is at most $16e^{2n\kappa_+}$. Apply (7.13) to Z_- with $q = q_-$, $\nu_0 = n$, base points u_b and the s gas points Jz_l. The sum is at most $16e^{2nc_-}$. Insert these in Proposition 7.6, use (7.11), and take logarithms. Each positive node contributes $\mathsf{g}(\xi) = -\mathsf{D}\log|\xi| + n(U^{q_+}(\xi) - \kappa_+)$ and each negative node $\mathsf{g}(x) = n(U^{q_-}(Jx) - c_-)$. We have proved the following statement.

Proposition 7.8. *For every $\eta > 0$ there is C such that for all N, all nodes and all S with $\det \Xi_S^0 \neq 0$,*

$$2\log|\det \Xi_S^0| \le 2\sum_a \mathsf{g}(a) + \mathsf{K}_N + 4n^2\eta + Cn,$$

where $\mathsf{g}(a)$ is the contribution of node a and

$$\mathsf{K}_N = 2n^2\kappa_+ + 2\cdot 53Nn\,c_- + \mathsf{D}(\mathsf{D}+1)\log 2 - N^2\int_0^{53}\tau^* - n^2\big(I(q_+)+I(q_-)\big).$$

The constant is correct because the k_- points of Z_+ give $2nk_-\kappa_+$ and the s points of Z_- give $2nsc_-$. In K_N these are written as $2n^2\kappa_+$ and $2\cdot 53Nnc_-$, and the differences are absorbed by the terms $-n\kappa_+$ and $-nc_-$ of g. The term $\mathsf{D}(\mathsf{D}+1)\log 2$ is the sum of $\mathsf{D}(\mathsf{D}-1)\log 2$ from $\prod_a \ell_a^2$ and $2\mathsf{D}\log 2$ from (7.11).

7.6. **The final inequality.** We now bound the other factors of the integrand. Let $n' = n-m$, with free points x_i and s_j and prescribed points y_j, and let $\mathcal{S}$ be the product of the factors of the integrand other than the row determinant and the factors $\mathcal{B}_\zeta$ of Lemma 5.6. Put $\mathsf{t}(x) = 2x/(1+x^2)$ and $t_i = \mathsf{t}(x_i)$. In the normalization of §7.1 a free point carries the weight $|t|^{75N-1}(1-t)^{3N}$. The prescribed columns have lost the common factor $|t|^{22N}(1-t)^{3N}$ of the rows, so a prescribed point carries $|\mathsf{t}(y)|^{\mathsf{D}}$. The Cauchy determinant of Lemma 5.5 is $\pm V(t)V(s)/\prod_{i,j}(1-t_is_j)$, and the generalized Vandermonde determinant gives $\prod_j s_j^{9N}(1-s_j)^{3N}|V(s)|$. Hence

$$\mathcal{S} = \prod_j |\mathsf{t}(y_j)|^{\mathsf{D}} \prod_i |t_i|^{75N-1}(1-t_i)^{3N} \cdot \prod_j s_j^{9N}(1-s_j)^{3N}\,\frac{|V(t)|\,V(s)^2}{\prod_{i,j}(1-t_is_j)}.$$

Now $1-t = (1-x)^2/(1+x^2)$, $1-ts = (1-2xs+x^2)/(1+x^2)$ and

$$t_i - t_{i'} = \frac{2(x_i-x_{i'})(1-x_ix_{i'})}{(1+x_i^2)(1+x_{i'}^2)}.$$

Each x_i occurs in $n'-1$ differences and in n' kernels, so the exponent of $1+x_i^2$ is $-(75N-1)-3N-(n'-1)+n' = -(78N-2)$. Taking logarithms gives

$$\log\mathcal{S} = \left(n'(75N-1) + \binom{n'}{2}\right)\log 2 + \sum_j \mathsf{w}(y_j) + \sum_i \mathsf{W}_N(x_i) + \sum_j \big(9N\log s_j + 3N\log(1-s_j)\big) + n^2\mathsf{Q}(x,s), \tag{7.15}$$

where $\mathsf{w}(y) = \mathsf{D}(\log 2 + \log|y| - \log(1+y^2))$, $\mathsf{W}_N(x) = (75N-1)\log|x| + 6N\log(1-x) - (78N-2)\log(1+x^2)$ and

$$n^2\mathsf{Q}(x,s) = \sum_{i<i'} \log(|x_i - x_{i'}|(1-x_ix_{i'})) + 2\sum_{j<j'}\log|s_j - s_{j'}| - \sum_{i,j}\log(1-2x_is_j+x_i^2).$$

This is the energy of §6.1 with the interaction among the x_i at weight 1. The second copy of that interaction is contained in the determinant.

For sequences $a = (a_k)$ and $v = (v_k)$ put $\|a\|^2 = \sum_k a_k^2/k$. Let b be a real number.

Lemma 7.9. *Let a, v be sequences with $\sum|a_k|/k$, $\sum|v_k|/k$, $\|a\|$ and $\|v\|$ finite. For every $\varepsilon > 0$ there is ν_1 such that for all $n \geq \nu_1$, all $n' \leq n$ and all configurations of distinct points,*

$$\mathsf{Q}(x,s) \leq \tfrac{3}{2}b^2\log 2 + \|a\|^2 + \tfrac{1}{2}\|v\|^2 - \tfrac{3}{2}\left(\tfrac{n'}{n} - b\right)^2\log 2 + \frac{1}{n}\sum_i \mathsf{R}_*(x_i) + \frac{1}{n}\sum_j \mathsf{C}_*(s_j) - \frac{1}{2n^2}\sum_i \log(1-x_i) - \frac{1}{n^2}\sum_j \log(1-s_j) + \varepsilon,$$

where $\mathsf{R}_*(x) = -b\log 2 - 2\mathsf{T}(a,x) - \mathsf{S}(v,x)$ and $\mathsf{C}_*(s) = -2b\log 2 + 2\mathsf{T}(v,s)$.

Proof. Use the expansions $-\log|x-x'| = \log 2 + 2\sum_k T_k(x)T_k(x')/k$, $-\log(1-xx') = \sum_k x^k x'^k/k$ and $-\log(1-2xs+x^2) = 2\sum_k x^k T_k(s)/k$. Let $A_k = \frac{1}{n}\sum_i T_k(x_i)$, $B_k = \frac{1}{n}\sum_i x_i^k$ and $C_k = \frac{1}{n}\sum_j T_k(s_j)$, and let $f = n'/n$. If the diagonal terms are ignored, these expansions give

$$\mathsf{Q} = -\tfrac{3}{2}f^2\log 2 - \|A\|^2 - \tfrac{1}{2}\|B - 2C\|^2.$$

The diagonal terms diverge. OpenAI's regularization [36, §5] replaces the kernels by Abel-damped series with a parameter e. It gives the identity above for the damped moments, up to an error at most $-\frac{1}{2n^2}\sum_i\log(1-x_i) - \frac{1}{n^2}\sum_j\log(1-s_j) + O_e(1/n) + 14e$. For the trial sequences, $-\|A\|^2 \leq \|a\|^2 - 2\sum_k a_kA_k/k$ and $-\frac{1}{2}\|B-2C\|^2 \leq \frac{1}{2}\|v\|^2 - \sum_k v_k(B_k - 2C_k)/k$. Removing the damping from the linear terms costs $O(e)$ times the ℓ^1 norms of the trial sequences, so we first replace a and v by long prefixes, with an error at most $\varepsilon/4$ uniformly in the points. Finally $-\frac{3}{2}f^2\log 2 = -\frac{3}{2}(f-b)^2\log 2 + \frac{3}{2}b^2\log 2 - 3bf\log 2$, and the last term is split between R_* and C_*. Choosing the prefix, then e, then ν_1 proves the lemma. The order of these choices makes the bound uniform in the configuration. □

We apply the lemma with $b = 1$. Let X be the inverse of $x \mapsto x + 1/x$ with values in $[-1,1]\setminus\{0\}$. Let R and C be the row and column references

of Appendix C, and let $L(\mu) = \int \log|X|\,d\mu$. We take

$$a_k = \int T_k(X)\,dR, \qquad v_k = \int X^k\,dR - 2\int T_k\,dC,$$

and put $\mathsf{I}_0 = \frac{3}{2}\log 2 + \|a\|^2 + \frac{1}{2}\|v\|^2$. The expansions in the proof of Lemma 7.9 give

$$\mathsf{R}_*(x) = \int (\log|x-X| + \log(1-xX))\,dR - \int \log(1-2xs+x^2)\,dC(s),$$

$$\mathsf{C}_*(s) = 2U^C(s) - \int \log(1-2Xs+X^2)\,dR.$$

For nonzero x and w and for $0 < s < 1$ we have

$$x + \frac{1}{x} - w - \frac{1}{w} = \frac{(x-w)(xw-1)}{xw}, \qquad 1 - 2xs + x^2 = |x|\,|x + \tfrac{1}{x} - 2s|. \tag{7.16}$$

Since R and C have mass 1, this gives

$$\mathsf{R}_*(x) = L(R) + U^{R-2\#C}(x + 1/x), \qquad \mathsf{C}_*(s) = -L(R) + 2U^C(s) - U^R(2s).$$

So $9N\log s + 3N\log(1-s) + n\mathsf{C}_*(s) = n\mathsf{C}(s)$ for the column field C of the appendix.

The moved sequences. Lemma 7.9 holds for any sequences, so we may move a and v with the fraction α. Let $\mathrm{d}a^0, \mathrm{d}a^1, \mathrm{d}v^0, \mathrm{d}v^1$ be fixed sequences supported on $k \le 160$, with dyadic entries of denominator 2^{40}. Put $\alpha_e = \frac{1}{25}$, $\tau = \min(1, \alpha/\alpha_e)$ and

$$a(\tau) = a + \mathrm{d}a^0 + \tau\,\mathrm{d}a^1, \qquad v(\tau) = v + \mathrm{d}v^0 + \tau\,\mathrm{d}v^1.$$

The increment $(\mathrm{d}a^0, \mathrm{d}v^0)$ moves the sequences already at $\alpha = 0$. The fields change by finite sums: R_* gains $h^0 + \tau h^1$ with $h^i = -2\mathsf{T}(\mathrm{d}a^i, \cdot) - \mathsf{S}(\mathrm{d}v^i, \cdot)$, and C_* gains $k^0 + \tau k^1$ with $k^i = 2\mathsf{T}(\mathrm{d}v^i, \cdot)$. The polarization identity gives

$$\mathsf{I}_0(\tau) = \mathsf{I}_0 + (L_0 + Q_{00}) + \tau(L_1 + 2Q_{01}) + \tau^2 Q_{11},$$

with $L_i = 2\langle a, \mathrm{d}a^i\rangle + \langle v, \mathrm{d}v^i\rangle$ and $Q_{ij} = \langle \mathrm{d}a^i, \mathrm{d}a^j\rangle + \frac{1}{2}\langle \mathrm{d}v^i, \mathrm{d}v^j\rangle$. Only L_0 and L_1 involve the references. We also include the column field γg of §6.2, with $\gamma = \tau\alpha_e$ for $\alpha \le \alpha_e$ and $\gamma = \alpha$ for $\alpha_e \le \alpha \le \alpha_1$, so that $\gamma \ge \alpha$. The moved row field at τ is the convex combination $(1-\tau)(F_\pm + h^0) + \tau(F_\pm + h^0 + h^1)$, and the moved column field is affine in τ and γ. So five numbers bound all of them, the bounds at the end points:

$$F_\pm + h^0 \le c_0^{\mathsf{R}}, \qquad F_\pm + h^0 + h^1 \le c_1^{\mathsf{R}},$$

$$\mathsf{C} + k^0 \le c_0^{\mathsf{C}}, \qquad \mathsf{C} + k^0 + k^1 + \alpha_e g \le c_1^{\mathsf{C}}, \qquad \mathsf{C} + k^0 + k^1 + \alpha_1 g \le c_2^{\mathsf{C}}.$$

For $\alpha \le \alpha_e$ the row and column caps are the chords $(1-\tau)c_0^{\mathsf{R}} + \tau c_1^{\mathsf{R}}$ and $(1-\tau)c_0^{\mathsf{C}} + \tau c_1^{\mathsf{C}}$. For $\alpha_e \le \alpha \le \alpha_1$ they are c_1^{R} and the chord of c_1^{C} and c_2^{C} at α. No supremum of an increment is ever taken separately from the supremum of the base field.

We turn to the gas potentials. Let $\nu_\pm$, $A_\pm$ and $\tilde{q}_-$ be as in Appendix C, so that $q_+ = X_\# G_+ + \nu_+$ and $q_- = J_\#\tilde{q}_-$ with $\tilde{q}_- = X_\# G_- + \nu_-$. Let $0 < x < 1$ and $Y = x + 1/x$. By (7.16), $\log|x - X(Y')| = \log|Y - Y'| +$

$\log x + \log|X(Y')| - \log(1 - xX(Y'))$. On the support of G_+ we have $X > 0$, so $-\log(1-xX) = \sum_k x^k X^k/k$ has positive terms. Those with $k \le 900$ integrate to $\mathsf{P}(x)$ and the others to at most E. Hence

$$U^{q_+}(x) \le L(G_+) + U^{G_+}(Y) + m_+ \log x + \mathsf{P}(x) + \mathsf{E} + A_+(x). \tag{7.17}$$

Let now $-1 < x < 0$. We have $\log|Jx - Y'/2| = \log|Y - Y'| - \log 2$ on the support of G_-. By (7.16), $\log|Jx - Jw| = \log|x - w| + \log|1 - xw| - \log 2 - \log|x| - \log|w|$ on the support of ν_-. Put $c_M = \int \log|u|\, d\tilde{q}_-(u)$. Integrating, and using the mass $\frac{53}{58}$ of $\tilde{q}_-$, we find

$$\begin{aligned} U^{q_-}(Jx) = L(G_-) + U^{G_-}(Y) + m_- \log|x| + A_-(x) \\ - \tfrac{53}{58}(\log 2 + \log|x|) - c_M. \end{aligned} \tag{7.18}$$

Define κ_- by $c_- = \kappa_- - \frac{53}{58}\log 2 - c_M$. Since $n \cdot \frac{53}{58} = 53N$, this says

$$\mathsf{g}(x) = n\big(L(G_-) + U^{G_-}(Y) + m_- \log|x| + A_-(x) - \kappa_-\big) - 53N \log|x|. \tag{7.19}$$

Let $F_\pm$ be the free row fields of the appendix. Since $n \cdot \frac{11}{29} = 22N$, $n \cdot \frac{39}{29} = 78N$ and $n \cdot \frac{3}{29} = 6N$, the definition reads

$$\begin{aligned} nF_\sigma = n\mathsf{R}_* + n\big(L(G_\sigma) + U^{G_\sigma}(Y) + m_\sigma \log|x| - \kappa_\sigma \\ + [\sigma = +](\mathsf{P} + \mathsf{E}) + A_\sigma\big) + 22N \log|x| - 78N \log(1 + x^2) + 6N \log(1 - x). \end{aligned}$$

Recall that $\mathsf{g}(x) = -\mathsf{D}\log x + n(U^{q_+}(x) - \kappa_+)$ for $x > 0$. Add W_N to this and use (7.17) for $0 < x < 1$ and (7.19) for $-1 < x < 0$. The terms in $\log(1-x)$ cancel, the exponents of $1+x^2$ leave 2, and the coefficient of $\log|x|$ is $75N - 1 - 22N - \mathsf{D} = 0$ in the first case and $75N - 1 - 22N - 53N = -1$ in the second. So

$$\mathsf{W}_N(x) + n\mathsf{R}_*(x) + \mathsf{g}(x) \le nF_+(x) + 2\log(1 + x^2), \tag{7.20}$$

$$\mathsf{W}_N(x) + n\mathsf{R}_*(x) + \mathsf{g}(x) = nF_-(x) + 2\log(1 + x^2) - \log|x|. \tag{7.21}$$

Let R and C_0 be upper bounds for the moved row and column fields at the given α. The added fields are bounded, so they do not affect the following argument, which we write for $F_\pm$ and C. The term $-\frac{1}{2}\log(1-x)$ of Lemma 7.9 and the term $-\log|x|$ of (7.21) are unbounded. They are controlled by the field itself. Let H_0 be the supremum of $F_+ - \frac{3}{29}\log(1-x) - \frac{11}{29}\log|x|$ on $(0,1)$ and of $F_- - \frac{3}{29}\log(1-x) - \frac{11}{29}\log|x|$ on $(-1,0)$. It is finite. Indeed, after the subtraction F_σ consists of constants, the term $-\frac{39}{29}\log(1+x^2) \le 0$, the polynomial P, the term A_σ and $U^{G_\sigma + R - 2_\# C}(x + 1/x) + m_\sigma \log|x|$. The term A_σ is bounded because the auxiliary measures have bounded densities on circles. The measures G_σ, R and $2_\# C$ are piecewise uniform with compact support, so their potentials are continuous on the real line. This settles the limits $x \to \pm 1$, where $x + 1/x \to \pm 2$, and every compact subset of $(-1,1) \setminus \{0\}$. As $x \to 0$ we have $|x + 1/x| \to \infty$. The measure $R - 2_\# C$ has mass 0 and G_σ has mass m_σ, so $U^{G_\sigma + R - 2_\# C}(x + 1/x) + m_\sigma \log|x|$ tends to

0. Now split $nF = (n - \frac{29}{6})F + \frac{29}{6}F$. Since $\frac{29}{6} \cdot \frac{3}{29} = \frac{1}{2}$ and $\frac{29}{6} \cdot \frac{11}{29} = \frac{11}{6} \geq 1$, and since $\log|x| \leq 0$,

$$nF - \tfrac{1}{2}\log(1-x) - \log|x| \leq (n - \tfrac{29}{6})\mathsf{R} + \tfrac{29}{6}H_0.$$

The columns are treated in the same way. Let H_1 be the supremum of $\mathsf{C}(s) - \frac{3}{58}\log(1-s)$ on $(0,1)$. It is finite, because U^C and $U^R(2s)$ are continuous and $\log s \leq 0$. Since $\frac{58}{3} \cdot \frac{3}{58} = 1$,

$$n\mathsf{C}(s) - \log(1-s) \leq (n - \tfrac{58}{3})\mathsf{C}_0 + \tfrac{58}{3}H_1.$$

So every free point contributes at most $n\mathsf{R} + O(1)$ and every free column at most $n\mathsf{C}_0 + O(1)$.

Let now y be a prescribed point. All prescribed points lie in E_x, so $y \geq 98749/10^6$ or $y \leq -19501/10^5$. Let H_+, H_- and H_S be the fields of the prescribed rows in the appendix. For $y > 0$ outside S, (7.17) and $53N = \mathsf{D}+1$ give

$$\mathsf{w}(y) + \mathsf{g}(y) \leq \mathsf{D}\log 2 + n\mathsf{H}_+(y) + \log(1+y^2).$$

For $y < 0$, (7.19) gives

$$\mathsf{w}(y) + \mathsf{g}(y) = \mathsf{D}\log 2 + n\mathsf{H}_-(y) - \log|y| + \log(1+y^2).$$

For $y \in S$ the base point is $\xi = 1/y$, so $\mathsf{g} = \mathsf{D}\log y + n(U^{q_+}(1/y) - \kappa_+)$. We have $\log|1/y - w| = -\log y + \log|1 - yw|$, the measure q_+ has mass 1, and $\int \log(1-yX)\,dG_+ \leq -\mathsf{P}(y)$ because all the moments are nonnegative. Hence $U^{q_+}(1/y) \leq -\log y - \mathsf{P}(y) + \int \log|1-yw|\,d\nu_+(w)$. With $2\mathsf{D} = 106N - 2$ this gives

$$\mathsf{w}(y) + \mathsf{g} \leq \mathsf{D}\log 2 + n\mathsf{H}_S(y) - 2\log y + \log(1+y^2).$$

The factor $e^{-N\hat{\delta}(\mathsf{t}(y))}$ of Proposition 6.2 adds $\hat{\delta}(\mathsf{t}(y))/58$ to these fields. Let M be an upper bound for $\mathsf{H}_\pm + \hat{\delta}\circ\mathsf{t}/58$ on E_W and for $\mathsf{H}_S + \hat{\delta}\circ\mathsf{t}/58$ on $E_W \cap (0,1)$. Since $\log(1+y^2) \leq \log 2$ and $|y|$ is bounded below on E_x, every prescribed row, with its factor $e^{-N\hat{\delta}}$, contributes at most $\mathsf{D}\log 2 + n\mathsf{M} + O(1)$.

Now add (7.15), Lemma 7.9 multiplied by n^2, and half of Proposition 7.8. With $n' = (1-\alpha)n$ and $\eta, \varepsilon \to 0$, this gives

$$\log(\mathcal{S}\,|\det \Xi_S^0|) \leq n^2(\mathsf{B}(\alpha) + o(1)),$$

uniformly in the configuration, with

$$\mathsf{B}(\alpha) = \mathsf{I} - \tfrac{3}{2}\alpha^2\log 2 + (1-\alpha)(\mathsf{R} + \mathsf{C}_0) + \alpha\mathsf{M} + \log 2\Big(\tfrac{75}{58}(1-\alpha) + \tfrac{1}{2}(1-\alpha)^2 + \tfrac{53}{58}\alpha\Big). \tag{7.22}$$

Here $\mathsf{I} = \mathsf{I}_0(\tau) + \mathsf{I}_{\text{gas}}$, and $\frac{1}{2}\mathsf{K}_N/n^2$ tends to

$$\mathsf{I}_{\text{gas}} = \kappa_+ + \tfrac{53}{58}(\kappa_- - c_M) - \tfrac{1}{2}\big(\tfrac{53}{58}\big)^2\log 2 - \tfrac{1}{2}(I(q_+) + I(q_-)) - \frac{1}{2\cdot 58^2}\int_0^{53}\tau^*,$$

with κ_- and c_M as above. The three terms in $\log 2$ of (7.22) come from the exponent $n'(75N-1)+\binom{n'}{2}$ and from the term $\mathsf{D}\log 2$ of each prescribed row. Appendix C gives the reference measures and the bounds

$$\begin{gathered}\mathsf{I}_0+\mathsf{I}_{\mathrm{gas}}\le 2.366997,\qquad \mathsf{M}\le -0.976,\\ c_0^{\mathsf{R}}=-2.579740254,\quad c_1^{\mathsf{R}}=-2.574009165,\\ c_0^{\mathsf{C}}=-2.946694099,\quad c_1^{\mathsf{C}}=-2.888907917,\quad c_2^{\mathsf{C}}=-2.85421372,\end{gathered}\tag{7.23}$$

with $\kappa_-=5463909/10^8$ and $\kappa_+=14762657/10^8$, and the moved values

$$L_0\le -7.5421576\cdot 10^{-5},\quad L_1\le -0.04943292757,$$
$$Q_{00}\le 1.6623\cdot 10^{-7},\quad Q_{01}\le 9.5232\cdot 10^{-7},\quad Q_{11}\le 1.002863\cdot 10^{-3}.$$

Let $V_S(\alpha)=\mathsf{B}(\alpha)+U_0-\frac12\log 2-c_2\alpha$, with I, R and C_0 the moved values above and $\log 2$ replaced by an upper bound. On each of the two pieces $[0,\alpha_e]$ and $[\alpha_e,\alpha_1]$ it is a polynomial in α with rational coefficients. We have $V_S(0)=-1.81\ldots\cdot 10^{-5}$, and the cover of §6.4 proves $V_S(\alpha)\le\sigma\alpha$ on both pieces. The inequality is tight near $\alpha=0.0033$, where $V_S(\alpha)/\alpha$ reaches 0.33867.

Proof of Proposition 6.2. Fix a configuration with $m\le\alpha_1 n$. By (7.3) it suffices to bound each of the at most 2^m terms $\mathcal{S}\,|\det\Xi_S|$, at the cost of the factor 5^m. By Proposition 7.5 each term either satisfies the bound directly, or is at most $499^n\mathcal{S}\,|\det\Xi_S^0|$. In the second case the final inequality gives $\exp(n^2(-U_0+\frac12\log 2+c\alpha+o(1))-N\sum_j\hat\delta(t_{a_j}))$, since $499^n=e^{o(n^2)}$ and $c=c_2+\sigma$. The factor $\prod_l(1+\zeta s_l)^m$ of Lemma 5.6 is contained in the column field, and the factors $\mathcal{B}_\zeta$ give $\prod_j\mathcal{B}_\zeta(k_j)$. □

8. Proof of Theorem 1.1

OpenAI's energy method gives the decay rate $D_0=n_0^2(U-\frac12\log 2)=6422.8355\ldots$ for our parameters. Section 7 raises it to $D=6447.4755\ldots$ With $\Lambda=6392.7870\ldots$ the polynomial $Q_N=(d_0N)!\,\mathcal{D}_N\Delta_N$ has integer coefficients, degree at most $53N$, and

$$|Q_N(x)|\le e^{-\eta N^2}\qquad(|x-G|\le e^{-\rho N}),\qquad \eta=54.68745.$$

Two further properties are needed to choose the prime in the criterion: nonvanishing of $Q_p(a/b)$ when $p^2\nmid b$ (Proposition 3.6), and divisibility of the i-th coefficient of Q_N by p'^{2i} for every prime $N<p'\le 2N$ (Lemma 3.5). We now check the hypotheses of Proposition 2.2.

Take $d=53$, and let the polynomials Q_p of the proposition be the Q_N with $N=p$. Let

$$\rho=278.376313578,\quad \eta=54.68745,\quad \omega=\tfrac35,\quad \delta=2.9\cdot 10^{-7},$$
$$\lambda=\frac{270}{\rho}=0.9699101\ldots$$

(a) $\deg Q_N\le 53N$ by (3.5) and Lemma 2.1.

(b) By Proposition 3.6, $\Delta_p(a/b) \neq 0$ for large p with $p^2 \nmid b$. Hence $Q_p(a/b) \neq 0$.
(c) By Lemma 3.5, p'^{2i} divides the i-th coefficient of Q_p for all primes $p < p' \leq 2p$.
(d) By Theorem 5.1, Proposition 3.4 and $\log(53N)! = o(N^2)$, for $|x-G| \leq e^{-\rho N}$ and large N,

$$|Q_N(x)| \leq \exp\bigl(-(D - \Lambda + o(1))N^2\bigr) \leq e^{-\eta N^2},$$

since $\eta = 54.68745 < D - \Lambda = 54.68853\ldots$ In the formal proof this is the inequality $\eta < D - \Lambda$ with $\int \kappa$ replaced by its lower bound 22.0607 and $D_0 - \Lambda_0$ by its lower bound 7.98775.

The numerical conditions hold as well. We have $\rho\lambda = 270$ and $0 < \eta < 2d = 106$. Next $\omega = \frac{3}{5} < \lambda$ and $\delta < \frac{1}{3}$. The third condition reads $53 \leq (\lambda-\omega)(\eta+106(1-3\delta)) = 59.43\ldots$ Further $2(\omega-\delta(2\lambda-\omega)) = 1.199999\ldots > 1$. Finally $d + 4d\delta\lambda = 53.0000596\ldots < \eta\lambda = 53.0419104\ldots$ So Proposition 2.2 gives $|G - a/b| > b^{-\rho\lambda} = b^{-270}$ for all integers a and all sufficiently large b. This proves Theorem 1.1.

9. Remarks on the method

9.1. **The three ingredients.** For fixed parameters the method gives the exponent $d\rho/\eta$, and its three ingredients come from different places. The degree $d = n - g$ comes from the construction. Of $\rho = 278.376313578$, the prescribed columns contribute $n_0 c \approx 228.3$ through the energy, and the weight and the Vandermonde factor of one column contribute $\psi \approx 50.1$. Of $n_0 c$, the part $n_0\sigma \approx 19.6$ comes from the bound of Section 7 for few prescribed columns. The number η is a small difference of two large numbers, $D \approx 6.4 \cdot 10^3$ and $\Lambda \approx 6.4 \cdot 10^3$. The corank bounds of Section 4 contribute about 22 to it, and the bound of Section 7 about 25. Without them η would be about 8. Since the exponent is inversely proportional to η, a further gain of 0.2 in η would lower it by about 1.

9.2. **The row determinant.**

Remark 9.1. The energy method bounds $|\det \mathsf{L}|$ by $e^{O(n \log n)}$, which is invisible at the scale e^{N^2}. At the extremal configuration, numerical evaluation for small N suggests that $\log|\det \mathsf{L}|$ is about $-28N^2$. So our bound captures most of its true size. The comparison of Lemma 7.2 loses only $e^{O(N)}$, and so do the Gram inequalities of Proposition 7.6. The loss is in the restriction to Gram measures on finitely many circles, and in the requirement that the bound hold for the maximum over all configurations. These numerical values are not part of the formal proof.

9.3. **Other parameters.** All the results of Section 4 are lower bounds for the corank. The closed formulas, the counting inequality, Theorem A.1 and the families derived from them are proved for all parameters. The rank data of Appendix A and the reference measures belong to $(58, 9, 3, 5, 3)$, and

Sections 5–7 are written and formalized for these parameters. Theorem 7.3 does not depend on the parameters. A different choice of parameters changes d, ρ and η together, and we make no claim that ours is optimal. Except for the pair (F4) and the intervals of Theorem A.1, the function κ uses no relation between different families. So at a prime where intervals overlap the exact corank may be larger than each bound. An upper bound for the corank is not needed. More independence between families would increase η without changing anything else.

10. Formalization

Theorem 1.1 and the results it depends on are formalized in Lean 4 [10] with Mathlib [47], on top of OpenAI's formalization of the irrationality of G [36]. The development is available at

`https://github.com/slitsey/catalan-irrationality-measure`.

Theorem 1.1 is the theorem

`CatalanIrrationalityMeasure.catalan_irrationalityMeasure_270`

in `Solution.lean`. Its statement, in `Challenge.lean`, uses only Mathlib. Catalan's constant is written as the series $\sum_j (-1)^j/(2j+1)^2$, and the conclusion is $|G - a/b| > b^{-270}$ for all integers a and all sufficiently large b. The modules of the proof contain no `sorry` and no `native_decide`. The command `#print axioms` shows that only the three standard axioms `propext`, `Classical.choice` and `Quot.sound` are used. The proof is kernel-checked: Lean's kernel accepts every declaration. It has not been run through the independent checker `lake comparator` [29]. The prime number theorem is taken from the PNT+ project [48]. The parts of OpenAI's formalization that are used were converted to Lean's module system and ported to Lean 4.35.0-rc2. The repository's `README` describes the changes and the layout of the modules.

Two lemmas of Section 5 on the weights of nodal polynomials come from Davis's formalization [9]. We use two of its Lean files, taken from commit `3c45153` under the Apache License 2.0. Their declarations and proofs are unchanged. We adapted the module paths and imports, and added comments and axiom checks. The repository's `NOTICE` file and a provenance record list these files with their source hashes.

The proof combines conventional formal proofs with numerical certificates. These are the inverses of the base matrices modulo 1009, the intervals and the chord table of Section 5, the trial sequences, the cell covers and the bounds of Appendix B, the reference measures and the bounds of Appendix C, and the rank data of Appendix A. The certificates are exact rational and integer computations checked by the Lean kernel without compiled code. The interval arithmetic used to generate them [23] is not trusted.

Disclosure of AI use

The mathematics, the Lean formalization, the certificate generators and the text of this paper were produced by AI systems at the request of the author. The author posed the problems, directed the work and is responsible for the paper. The systems were Claude Opus 5.5, Claude Sonnet 5.5 and Claude Haiku 5.5 (Anthropic), running in Claude Code, and GPT-6.1 Sol (OpenAI), running in Codex. The underlying irrationality proof for Catalan's constant and its Lean formalization are due to OpenAI [36].

Acknowledgements. We thank the authors of Mathlib and of the PNT+ project, OpenAI for making its formalization available, and Cinna Davis for making the formalization of [9] available.

Appendix A. Counting on every subspace

On 37 of the 113 intervals of Theorem 4.10 the corank bound comes from the counting inequality applied to all constant subspaces at once. These intervals lie in $(3N, 16.2N)$. No subspace is constructed. The proof bounds the maximum over subspaces of the right side of Theorem A.1 by forced intersections of row spaces, and the only finite data are upper bounds for ranks.

Let V be a finite-dimensional vector space over $\mathbb{F}_p$. Let c be a nonnegative integer. For j in a finite set $\mathcal{J}$ let $B_j\colon V \to W_j$ be a linear map and w_j a nonnegative integer, and let $\lambda_{j,k}$ $(k < w_j)$ be linear forms on $\mathbb{F}_p^c$. Let $\mathcal{C}$ be the common kernel of the maps $B_j \circ (\lambda_{j,k} \otimes \mathrm{id})$ on $\mathbb{F}_p^c \otimes V$. A *parametrization* is a linear map ι from $\mathbb{F}_p^c \otimes V$ to the data x of a decomposition modulo p (Definition 4.9) with three properties on $\mathcal{C}$. The value $\iota(x)$ is admissible, the map ι is injective on $\mathcal{C}$, and the row combination $c(\iota(x))$ satisfies the hypothesis of Theorem 4.1 with some polynomial of degree less than r. Nothing is required outside $\mathcal{C}$.

Theorem A.1. *Let $N \geq 1$ and let p be an odd prime with $H < p^2$. For a parametrization and every subspace $U \subset V$,*

$$\nu_N(p) \geq c \dim U - \sum_{j \in \mathcal{J}} w_j \operatorname{rank}(B_j|_U).$$

Proof. Since the coefficients $\lambda_{j,k}$ are scalars, $\lambda_{j,k} \otimes \mathrm{id}$ maps $\mathbb{F}_p^c \otimes U$ into U. So the restriction of ι to $\mathbb{F}_p^c \otimes U$ is a parametrization with V replaced by U and B_j by $B_j|_U$, and its common kernel is $\mathcal{C} \cap (\mathbb{F}_p^c \otimes U)$. Let Φ be the map on $\mathbb{F}_p^c \otimes U$ with components $B_j \circ (\lambda_{j,k} \otimes \mathrm{id})$. These components map into $B_j(U)$, so $\operatorname{rank}\Phi \leq \sum_j w_j \operatorname{rank}(B_j|_U)$. The hypotheses of Theorem 4.1 hold on $\ker\Phi$, with the injective map $x \mapsto c(\iota(x))$ on this kernel. The theorem gives $\nu_N(p) \geq \dim\ker\Phi \geq c\dim U - \sum_j w_j \operatorname{rank}(B_j|_U)$. □

In the applications c and the w_j are affine functions of N and p on an interval of primes, while V and the B_j are fixed. The space $\mathbb{F}_p^c$ is a space of polynomials of degree less than c in the variable of a decomposition modulo

p. The forms $\lambda_{j,k}$ are Taylor coefficients at $X = -1, 0, 1$ and coefficients at ∞. The map ι places the data block by block along the rows. Its three properties are proved once for the whole family of parametrizations. On $\mathcal{C}$ the moment columns take the values of a polynomial of degree less than r. Since the intervals lie in $(3N, 16.2N)$, where $p > q$, we have $a = 0$ in Theorem 4.1.

Lower bounds for the maximum. For a subspace S of the dual V^* let $U = S^\perp$. Let $R_j \subset V^*$ be the row space of B_j, and let $r_j \geq \dim R_j$ be real numbers, in practice upper bounds for the ranks. Since $\operatorname{rank}(B_j|_U) = \dim R_j - \dim(R_j \cap S)$, the right side of Theorem A.1 is

$$E(S) = c(\dim V - \dim S) - \sum_j w_j\big(\dim R_j - \dim(R_j \cap S)\big).$$

Let $e_0 = c \dim V - \sum_j w_j r_j$ and $\delta_j = r_j - \dim R_j \geq 0$. Then

$$E(S) = e_0 + \sum_j w_j \delta_j + \sum_j w_j \dim(R_j \cap S) - c \dim S. \tag{A.1}$$

A map may also be omitted from the sum if its weight is subtracted from c, because $\dim R_j - \dim(R_j \cap S) \leq \dim V - \dim S$. The following three lemmas give a subspace with $E(S) > e_0$, from relations among the row spaces. All sets of indices are subsets of $\mathcal{J}$, and all w_j are nonnegative.

Lemma A.2 (pair)**.** *Let $a \neq b$ lie in a set I such that $R_a \leq R_j$ or $R_b \leq R_j$ for all $j \in I$. Let $s \geq \dim(R_a + R_b)$ and $0 \leq h \leq \min\big(w_a, w_b, \sum_{j\in I} w_j - c\big)$. Then some S has $E(S) \geq e_0 + \max(0, r_a + r_b - s)\, h$.*

Proof. Take $S = R_a \cap R_b$ and let $k = \dim S$. Then $S \leq R_j$ for $j \in I$, so (A.1) gives

$$E(S) - e_0 \geq w_a\delta_a + w_b\delta_b + \Big(\sum_{j\in I} w_j - c\Big)k \geq h(\delta_a + \delta_b + k).$$

Since $\dim(R_a + R_b) + k = \dim R_a + \dim R_b$, we have $r_a + r_b - s \leq \delta_a + \delta_b + k$. Also $\delta_a + \delta_b + k \geq 0$. □

Lemma A.3 (several points)**.** *Let $c \geq 0$. Let $I_1, \dots, I_l$ be disjoint sets with $a_i \in I_i$, such that $R_{a_i} \leq R_j$ for $j \in I_i$. Put $W_i = \sum_{j\in I_i} w_j$ and suppose that $W_i \leq c$ for each i. Let $s \geq \dim \sum_i R_{a_i}$. Let $h \geq 0$ with $h \leq w_{a_i}$ for each i and $h \leq \sum_i W_i - (l-1)c$. Then some S has $E(S) \geq e_0 + \max(0, \sum_i r_{a_i} - s)\, h$.*

Proof. Let T be the kernel of the sum map $\bigoplus_i R_{a_i} \to V^*$, and let $k = \dim \mathsf{T}$. Let $K_i \leq R_{a_i}$ be the projection of T to the summand i, let $\kappa_i = \dim K_i$, and let $S = \sum_i K_i$. Then $\kappa_i \leq k$. The rank formula for the sum map gives $\sum_i \dim R_{a_i} \leq s + k$. The kernel of the sum map on $\bigoplus_i K_i$ is again T, so $\dim S = \sum_i \kappa_i - k$.

For $j \in I_i$ we have $K_i \leq R_j \cap S$, so $\dim(R_j \cap S) \geq \kappa_i$. Since the I_i are disjoint, the a_i are distinct, and (A.1) gives

$$E(S) - e_0 \geq \sum_i w_{a_i}\delta_{a_i} + \sum_i W_i\kappa_i - c\sum_i \kappa_i + ck \geq h\sum_i \delta_{a_i} + \sum_i (W_i - c)\kappa_i + ck.$$

Since $W_i - c \le 0$ and $\kappa_i \le k$, the last two terms are at least $k(\sum_i W_i - lc + c) \ge hk$. So $E(S) - e_0 \ge h(\sum_i \delta_{a_i} + k)$, and $\sum_i r_{a_i} - s \le \sum_i \delta_{a_i} + k$. □

Lemma A.4 (kernel)**.** *Let $R \subset V^*$ be a subspace with $\dim R \le s$, and let Z be a set with $R_j \le R$ for $j \in Z$. For every integer $d \ge 0$ with $d + s \le \dim V$ some S has $E(S) \ge cd - \sum_{j \notin Z} w_j \min(r_j, d)$.*

Proof. The space $R^\perp \subset V$ has dimension at least d. Choose $U \le R^\perp$ of dimension d and let $S = U^\perp$. Then $B_j|_U = 0$ for $j \in Z$, and $\operatorname{rank}(B_j|_U) \le \min(r_j, d)$ for the other j. □

The rank data. The space V is a space of profiles of the coefficient vectors of a block, with coordinates T_k. The conditions at the moment columns are linear combinations of a small number of basic rows

$$A[k] \mapsto T_k + T_{k+1}, \qquad C[k] \mapsto T_{k+1} - T_k,$$

together with the row of the centre coordinate. So the moment values themselves are not needed. Every row factors through the map G whose components are these basic rows. This gives $r_j \le \min(\operatorname{rank} F_j + \#\text{basic rows},\ \operatorname{rank}(G, F_j))$, where F_j is the fixed part of the map. Each rank bound is certified by an integer left inverse, checked by the Lean kernel. Integer left inverses are valid at every prime, so no prime enters this step.

The 37 intervals are treated as follows: 13 by Lemma A.4, 12 by Lemma A.3, 11 by Lemma A.2, and one by $E(V)$ itself. On each interval the bound is an affine function $\nu_N(p) \ge c_0 N + c_1 p - 2$ with rational c_0 and c_1. The rounding of the integer orders is absorbed by the shortening τ of the interval and the lowering of the bound by 10^{-6} in Section 4.10.

Appendix B. The bounds for the energy method

This appendix gives the trial sequences of Section 6, the certified bounds that they satisfy, and the bounds for the weights of Section 5.

Trial sequences. For the constrained case we use a trial sequence with 40 Chebyshev coefficients and a finite number of geometric terms, with the multiplier $\lambda_1 = 1.68493357$. For the case with doubled interaction we use trial sequences of the form (6.2), with 160 Chebyshev coefficients, which are dyadic rationals, and 75 geometric terms. Of these, 41 have complex z_j and complex r_j. We place the complex charges at the 41 equally spaced points $x_0 = -0.16 + 0.008j$ $(0 \le j \le 40)$, all with $\epsilon_0 = 10^{-2}$. The real terms sit at points $\pm z$ with z increasing from 0.8 to 0.999997. The numbers z_j are rationals with at most twelve decimals. The norms of the geometric terms have closed forms: two terms (z_j, r_j) and (z_l, r_l) contribute $-\frac{1}{2}\operatorname{Re}(r_j r_l \log(1 - z_j z_l) + r_j \bar{r}_l \log(1 - z_j \bar{z}_l))$ to $\|u\|^2$.

The suprema of the fields. We must bound two functions of one variable from above on their domains: the field X on $[-1, 1]$ and the field Y on $[0, 1]$. We locate the real critical points. The derivative is a rational function

with a known denominator, and an exact identity gives its numerator. The Descartes rule, after a Möbius transformation of the interval, bounds the number of real roots of the numerator. Sign evaluations at 81-bit rational points then isolate this number of roots. One global Lipschitz bound on each isolating interval and the values at the end points give the supremum.

The bounds. For the base trial sequences this gives

$$\sup X \le -0.96450499, \quad \sup Y \le -1.57499400.$$

So $M_X + M_Y = -2.53949899$, and the norms satisfy $2\|u\|^2 + \frac{1}{2}\|v\|^2 \le 0.7736206$. The first coefficient satisfies $u_1 \ge 9/50$, which is used near $x = 0$. For the constrained case, with the weights $(\alpha_A, \beta_h, \gamma_H - \frac{1}{2}, \lambda_1, 2g_0/n_0) = (11/29, 3/29, 49/58, \lambda_1, 5/29)$, the bounds $\sup X_1 \le -1.15854$ and $\sup Y \le -1.32981$ are obtained by isolating the zeros of the numerators of the derivatives, which are polynomials with integer coefficients. The norm bound is $n_1 = 0.67761$.

The increments. The increment of the case with doubled interaction has 40 Chebyshev coefficients, and that of the constrained case has 24. Both are dyadic. Their fields are bounded on the whole domains by the same methods. For the case with doubled interaction the norm bound at τ is $0.7736206 - 0.117647903\,\tau + 0.005215061\,\tau^2$, the field X gains at most $0.01484272\,\tau$, and the column field Y with γg is bounded by a piecewise linear function of γ with knots at $\alpha_1, \frac{1}{5}, \frac{3}{10}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$. The fields F_1 and F_2 contain the node term $\hat{\delta}/58$. They are bounded by covers by cells, with exact rational and integer arithmetic on each cell and the maximum of $\hat{\delta}$ on each cell. For F_1 on $|x| \ge 401/2500$ the cover has 150 cells. It uses a Taylor expansion of the polynomial part and an expansion to order 5 of the logarithmic terms. On the positive part of E_W below $401/2500$ six further cells suffice. For F_2 the covers of $[98749/10^6, 9995/10^4]$ have 17 and 14 cells at the two ends of the move. This gives

$$\sup F_1 \le 0.007, \qquad \sup F_2 \le 0.007 - \log 2$$

at $\tau = 0$, and -0.0268 and $0.0233 - \log 2$ at $\tau = 1$. For the constrained case the norm bound at τ is $0.67761 - 0.082834152\,\tau + 0.004096781\,\tau^2$, and the bound $\mathcal{M}_1$ of Section 6 is the chord of 0.35072 and 0.35295. The two numbers are maxima over the 865 segments of the chord table at the two end values of λ_1. With $a_0 = 41/58$, $a_1 = 76/29$ and $U_0 = 2.2631834$ these bounds give the inequalities of §6.4. Each check $V(\alpha) \le \sigma\alpha$ is a polynomial inequality on a closed interval and is proved by a cover in exact rational arithmetic.

The bound for the weights. Inequalities (5.6) and (5.7) concern the explicit function L of (5.3). On the circle $|z| = \frac{3}{4}$ write $z = \frac{3}{4}(s + i\sqrt{1 - s^2})$ with $s = \cos\theta$. Then $|z - t|^2 = \frac{9}{16} + t^2 - \frac{3}{2}ts$ is affine in s for every real t. So $\log|1 - z|$ is a concave function of s. Each $G_c(z)$ is arccosh of $(|z - a_c| + |z - b_c|)/w_c$, a sum of square roots of affine functions of s, and therefore concave in s as well. Hence the left side of (5.6) is a concave function of $s \in [-1, 1]$. Its maximum

is $3.25053\ldots$ near $s=-0.1741$. A concave function is bounded by its secants outside the points and by the chord between them, so three points at distance $\frac{1}{2000}$ around the maximum suffice. At these points the value is enclosed to within 10^{-17} by exact rational arithmetic with integer square roots, and the bound $D_\Gamma = 3.2506$ holds with a margin $6\cdot 10^{-5}$. On the real segments we use, for real x, the closed form $e^{G_c(x)}w_c/4 \le \frac12 \max(w_c, |2x-a_c-b_c|)$ together with $6\log|1-x| \le 6\log\frac14$ for $x \ge \frac34$ and $6\log|1-x| \le 6\log 2$ for $x \le -\frac34$. This gives (5.7) with values below -9, far below D_Γ. The knot values v_i of the chord table lie between -1.97 and 17.99.

Appendix C. The reference measures of Section 7

The bound of §7.6 uses reference measures q_+ and q_-, a row reference R and a column reference C. This appendix describes them and the certified bounds (7.23).

The measures. On the real line of the variable $Y = x + 1/x$ with $|Y| \ge 2$ we take measures G_+ and G_- that are piecewise uniform on finitely many cells in Y, with masses m_+ and m_- and total mass 1. They carry the positive and the negative free nodes, and $R = G_+ + G_-$ is the row reference. The column reference C has mass 1. The measures are completed by auxiliary measures on the 16 inner circles. Each auxiliary measure has a cubed Poisson density: its Fourier coefficients are $-q^k(1+Ak+Bk^2)/k$ with $q = e^{-\varepsilon/2}$ and $\varepsilon = 27/250$, and the constants A and B are rational functions of $r = q^2$. The auxiliary measures ν_+ and ν_- sit at the phases $\pi j/93$, with 616 and 935 entries respectively, and have total masses $1-m_+$ and $\frac{53}{58}-m_-$. Let X be the inverse of $x \mapsto x + 1/x$ with $|X| \le 1$. Then $q_+ = X_\# G_+ + \nu_+$ is the measure in the plane of z. In the plane of u we take $q_- = J_\#\tilde q_-$ with $\tilde q_- = X_\# G_- + \nu_-$.

The fields. Let $L(\mu) = \int \log|X|\,d\mu$. For $\sigma = \pm$ the free row field is

$$F_\sigma(x) = L(R) + L(G_\sigma) - \kappa_\sigma + U^{G_\sigma + R - 2_\# C}(x+1/x) + (m_\sigma + \tfrac{11}{29})\log|x| \\ - \tfrac{39}{29}\log(1+x^2) + \tfrac{3}{29}\log(1-x) + [\sigma = +](\mathsf{P}(x) + \mathsf{E}) + A_\sigma(x),$$

where $A_+ = U^{\nu_+}$ and $A_-(x) = \int(\log|x-w| + \log|1-xw|)\,d\nu_-(w)$. The polynomial $\mathsf{P}(x) = \sum_{k\le 900} m_k x^k/k$ with $m_k = \int X^k\,dG_+$ is the truncated reflection potential of the positive measure, and $\mathsf{E} \ge 0$ bounds its tail. The column field is

$$\mathsf{C}(s) = -L(R) + 2U^C(s) - U^R(2s) + \tfrac{9}{58}\log s + \tfrac{3}{58}\log(1-s).$$

A prescribed row outside S has the field H_σ obtained from F_σ by three changes. The references R and C are removed, the coefficients $\frac{11}{29}$ and $\frac{3}{29}$ are replaced by 0, and $\frac{39}{29}$ is replaced by $\frac{53}{58}$. Its domain is $x \le -19501/10^5$ and $x \ge 98749/10^6$, which contains E_x. A prescribed row in S has the field

$$\mathsf{H}_S(y) = \tfrac{48}{58}\log y - \tfrac{53}{58}\log(1+y^2) - \kappa_+ - \mathsf{P}(y) + \int \log|1-yw|\,d\nu_+(w)$$

for $0 < y < 1$. The angular fields are the same potentials on the 16 circles, together with the terms $(53\sigma_j - \tau_j/2)/58$.

The constant I. Let $E(\mu,\nu) = \iint \log|Y - Y'|\, d\mu\, d\nu$. Evaluating Q on the references gives $\mathsf{I}_0 = -\frac{1}{2}E(R,R) - E(C,C) + E(R, 2_{\#}C)$. Since $R = G_+ + G_-$, six pair energies of lists of cells occur: those of G_+, G_- and C with themselves, $E(G_+, G_-)$, $E(G_+, 2_{\#}C)$ and $E(G_-, 2_{\#}C)$. The energies $I(q_\pm)$ split into cell, cross and auxiliary parts. Their cell parts contain the same pair energies. For piecewise uniform measures a pair energy is a bilinear form in the jumps of the densities, with the kernel $t^2 \log|t|$, up to constants. We evaluate it on integer end points in fixed point arithmetic. Well separated blocks of end points are treated by a series in the offsets from their centres, using the moments of the blocks, and the other pairs by a recursion on the blocks. This encloses each pair energy to within 10^{-7}. The remaining terms of $I(q_\pm)$ are the logarithmic moments of the cells, the integrals of the cells against the rings, which are sums of primitives over the end points, and the quadratic forms of the auxiliary measures, which have closed forms on the rings. All of these are enclosed to within 10^{-9}. The reflection term $-\iint \log(1 - xx')$ of $I(q_+)$ enters only through a lower bound from its truncated series. Finally $\int_0^{53} \tau^* \geq 1425.3339$. Together these give $\mathsf{I}_0 + \mathsf{I}_{\mathrm{gas}} \leq 2.366997$.

The column field. The three bounds c_i^{C} of (7.23) on $(0,1)$ use a chain of 386 units $s = (M + \delta z)/10^{12}$ with $|z| \leq 1$, together with two end ranges. On a unit the field is at most an integer polynomial in z plus an explicit error. The far cells of the references enter through exact integer moments of contiguous clusters, 1152 clusters in all, and the Taylor series of the logarithmic kernel. The near cells of C enter through the closed form of their potential. The two logarithms enter by series or, on the end units, by monotonicity. The constant is $L(R) \geq -0.6905432$, from the primitive of $\log|X|$ in Y.

The angular bounds. The bounds $\kappa_\pm$ are proved for both signs on each of the 16 radii, by 32 certificates. The auxiliary part is cut at degree 384, and its tail is bounded. The gas is split into bins. On a bin the kernel has the form $\mathsf{Q}_u(x) = \mathsf{a} - 2\mathsf{b}ux + \mathsf{c}x^2$ with $u = \cos\theta$. Jensen's inequality and an odd truncation of the logarithmic series bound $\int \frac{1}{2}\log \mathsf{Q}_u$ by a polynomial in u. The negative sign has a direct and a reflected kernel. The sum over the bins is converted exactly to a Chebyshev series. Its bound by $\kappa_\pm$ on $[-1,1]$ is certified by an identity of the same form as the one for the free rows below.

The prescribed rows. The auxiliary series of $\mathsf{H}_\pm$ are cut at degree 384. The gas potential is expanded in Chebyshev polynomials on $[0,1]$. The first 64 terms are exact, and the tail is bounded by Bessel's inequality from a bound for the L^2 energy of the cell densities. The reflection series is dropped by sign for the positive gas and kept through exact moments for the negative gas. The radial field is a sum of 16 blocks, one for each radius. Its polynomial part goes exactly into the Chebyshev variable, and the part of degree at most 16 is bounded by a chord cover with the radii as nodes. This gives a polynomial majorant, and a sum of squares identity gives $\mathsf{H}_\pm \leq -0.9783$.

On $[\frac{1}{10}, 1]$ the field H_S is treated in the same way. There P is replaced by its partial sum of degree 32, since the moments are nonnegative, the auxiliary series is cut at degree 32 with an absolute tail, and $\log y$ and $\log(1+y^2)$ are expanded explicitly. This gives $\mathsf{H}_S \le -0.997$. On $(0, \frac{1}{10}]$ one has directly $\mathsf{H}_S(y) \le \frac{24}{29}\log y + \log(1+y) - \kappa_+ +$(a fixed tail), which is smaller. So $\mathsf{M} = -0.976$ is admissible.

The free rows. The free row fields are bounded on each sign as follows. Put $t = |x|$ and $z = 2t - 1$. The exact field is at most its truncation to degree 384 in the auxiliary terms, up to a tail below 10^{-8}. We bound the truncated field by one polynomial in z. The direct potential of the own measure, $\int \log|t - X(y)|$, is replaced by its Chebyshev series, with a tail bounded through the variation of the density. Thin pieces of nonnegative density are dropped, since $\log|t - v| \le 0$ there. All other potentials are bounded from above by finite sums of terms $w \log|1 - r(\mathsf{q})t|$ with rational q: tangent majorants for the reflection and opposite pieces, chord majorants for the column panels, and the same form for $\log t$ and $\log(1-t)$. Each such term has an explicit Chebyshev expansion with a geometric tail. The moments of these terms and the direct moments are computed by exact integer arithmetic with Kronecker packing. The radial field and the term $-\frac{39}{29}\log(1+t^2)$ are bounded by a chord cover and a Chebyshev polynomial. This gives a polynomial $\mathsf{p}(z)$ with $F_\sigma \le \mathsf{p}$ plus explicit small errors. Finally $\mathsf{p}(z) \le \mathsf{c}$ on $[-1, 1]$ is proved by an exact identity $\mathsf{d}(\mathsf{c} - \mathsf{p}) = \mathsf{r} + \mathsf{q}_0^2$ between Chebyshev series, with q_0 of degree 16383 and with r small in the ℓ^1 norm of its coefficients. This gives the bounds c_0^{R} and c_1^{R} of (7.23) at the two ends of the move, for both signs.

References

[1] Roger Apéry, *Irrationalité de $\zeta(2)$ et $\zeta(3)$*, Astérisque **61** (1979), 11–13.

[2] Frits Beukers, *A note on the irrationality of $\zeta(2)$ and $\zeta(3)$*, Bull. London Math. Soc. **11** (1979), no. 3, 268–272.

[3] Hrishik Sai Bojnal, *An irrationality exponent below 2402 for Catalan's constant*, GitHub repository, October 15, 2026, https://github.com/proof-factory/catalan-exponent; design search and certification by Claude Fable 5.1 (Anthropic), 2026.

[4] Frank Calegari, Vesselin Dimitrov, and Yunqing Tang, *The linear independence of 1, $\zeta(2)$, and $L(2, \chi_{-3})$*, 2024, arXiv:2408.15403.

[5] ———, *Arithmetic holonomy bounds and effective Diophantine approximation*, 2025, Theorem 5.1 gives $\mu(L(2, \chi_{-3})) \le 24781$, arXiv:2510.04156.

[6] Gregory V. Chudnovsky, *The irrationality exponent of the lemniscate period: A determinant proof with a Lean 4 formalization*, 2026, Preprint, 19 October 2026, arXiv:2610.26539.

[7] Marcos Costa Santos Carreira, *Determinant constructions for the Dirichlet beta value at four*, Preprint, 9 October 2026, doi:10.5281/zenodo.23326647, 2026.

[8] ———, *General designs and an irrationality exponent for Catalan's constant*, Preprint, 9 October 2026, revised 21 October and 7 November 2026, doi:10.5281/zenodo.24042950; sources at https://github.com/MarcosCarreira/catalan-designs, 2026.

[9] Cinna Davis, *Catalan's constant: irrationality exponent research*, GitHub repository, https://github.com/1337isnot1337/catalan-constant; Lean verification at https://github.com/bigospeedrun/verifier/issues/87, October 31, 2026.
[10] Leonardo de Moura and Sebastian Ullrich, *The Lean 4 theorem prover and programming language*, Automated Deduction – CADE 28, Lecture Notes in Computer Science, vol. 12699, Springer, 2021, pp. 625–635.
[11] Payman Eskandari, *Improved rational approximations to Catalan's constant*, 2026, Preprint, 22 September 2026, arXiv:2609.26354.
[12] Payman Eskandari, V. Kumar Murty, and Yusuke Nemoto, *Mixed motives and linear forms in the Catalan constant*, J. Lond. Math. Soc. (2) **114** (2026), no. 4, Paper No. e70708, arXiv:2510.20648.
[13] Azriel Fasten, *Irrationality of the Riemann zeta value at five*, Lean formalization, Palomar registry entry PALOMAR-2026-09-25-000018, https://github.com/realazthat/zeta5-lean2, commit `2fa4e82`, 2026.
[14] Aabir Fauzan, $\zeta(5)$ *is irrational*, Preprint v1, September 17, 2026, Zenodo, doi:10.5281/zenodo.22826419, 2026.
[15] Moritz Firsching, *Zeta5*, Lean formalization, https://github.com/mo271/Zeta5, commit `7fe7367`, 2026.
[16] Stéphane Fischler, *Irrationality of values of L-functions of Dirichlet characters*, J. Lond. Math. Soc. (2) **101** (2020), no. 2, 857–876.
[17] Peter J. Forrester, *Meet Andréief, Bordeaux 1886, and Andreev, Kharkov 1882–83*, 2018, arXiv:1806.10411.
[18] David Niedbala Giraudin, *A new upper bound for the irrationality exponent of* $\zeta(2)$, 2026, Preprint, 2 October 2026, arXiv:2610.02912.
[19] ———, *A new upper bound for the irrationality exponent of* $\zeta(3)$, 2026, Preprint, 22 September 2026, arXiv:2609.26980.
[20] ———, *A new upper bound for the irrationality exponent of* $\zeta(5)$, 2026, Preprint, 1 November 2026, arXiv:2611.01708.
[21] Masayoshi Hata, *Rational approximations to* π *and some other numbers*, Acta Arith. **63** (1993), no. 4, 335–349.
[22] Rohit Kumar Jha, *A finite irrationality measure for Catalan's constant*, Preprint v1, October 8, 2026, doi:10.5281/zenodo.23237269; Lean sources at https://github.com/rohitkrjha/catalan-measure, 2026.
[23] Fredrik Johansson, *Arb: efficient arbitrary-precision midpoint-radius interval arithmetic*, IEEE Trans. Comput. **66** (2017), no. 8, 1281–1292, arXiv:1611.02831.
[24] Christian Krattenthaler and Tanguy Rivoal, *On a linear form for Catalan's constant*, South East Asian J. Math. Math. Sci. **6** (2008), no. 2, 3–15, arXiv:0810.1927.
[25] Christian Krattenthaler and Wadim Zudilin, *Hypergeometry inspired by irrationality questions*, Kyushu J. Math. **73** (2019), no. 1, 189–203, arXiv:1802.08856.
[26] Shawn Krisman, *One of* $\beta(4), \beta(6), \ldots, \beta(36)$ *is irrational*, Hexagon preprint hexagon:2610.00285v2, October 16, 2026, https://hexagonmath.org/2610.00285; Lean sources at https://github.com/telaviv/beta-irrationality, 2026.
[27] Veekesh Kumar and Riccardo Tosi, *Irrationality of values of L-functions for real Dirichlet characters*, 2026, Preprint, 30 October 2026, arXiv:2610.41458.
[28] Li Lai, *The linear independence of* 1, $\zeta(3)$, *and* $\zeta(4)$, 2026, Preprint, 5 October 2026 (v2: 15 October 2026), arXiv:2610.06838.
[29] Lean FRO, *Comparator: checking a Lean proof against a trusted statement*, https://github.com/leanprover/comparator, 2026, Software; accessed October 2026.
[30] Elyasheev Leibtag, Hila Barkan, Rotem Kalisch, Shachar Weinbaum, Michael Shalyt, and Ido Kaminer, *New directions in irrationality*, 2026, Preprint, 20 October 2026, arXiv:2610.27799; supplement and code at https://github.com/RamanujanMachine/new-directions-in-irrationality.

[31] Sean Litsey, *Irrationality measures from Löwner determinants*, Hexagon preprint hexagon:2610.00401, https://hexagonmath.org/2610.00401, 2026.
[32] Jingwen Liu, Kai Jiang, and Pingwen Zhang, *Irrationality exponents of logarithms of positive rational numbers*, 2026, Preprint, 7 October 2026, revised 5 November 2026 (v3), arXiv:2610.10192.
[33] Raffaele Marcovecchio, *The Rhin–Viola method for* $\log 2$, Acta Arith. **139** (2009), no. 2, 147–184.
[34] MultiAgency swarm, *Catalan:* $\mu(G) \le 528.2315$, GitHub repository, https://github.com/darkhipo/catalan-constant-gen5; verifier submission https://github.com/bigospeedrun/verifier/issues/103, November 2, 2026.
[35] Yu. V. Nesterenko, *On Catalan's constant*, Proc. Steklov Inst. Math. **292** (2016), no. 1, 153–170.
[36] OpenAI, *Catalan's constant is irrational*, Preprint dated September 24, 2026, released October 6, 2026, with a Lean formalization, available at the original paper PDF and the Lean formalization (repository commit `fd4aeeb`, October 7, 2026), 2026.
[37] ———, *The irrationality exponent of* π *is 2*, Preprint dated September 24, 2026, released October 6, 2026, with a Lean formalization, paper and Lean formalization (repository commit `fd4aeeb`), 2026.
[38] Thomas Prellberg, *Three irrationality results for the dilogarithm: Rhin–Viola and Viola–Zudilin constructions at* $-1/4$, $1/5$, *and* $-1/3$, 2026, Preprint, 7 September 2026, arXiv:2609.06895.
[39] Georges Rhin and Carlo Viola, *On a permutation group related to* $\zeta(2)$, Acta Arith. **77** (1996), no. 1, 23–56.
[40] ———, *The group structure for* $\zeta(3)$, Acta Arith. **97** (2001), no. 3, 269–293.
[41] Tanguy Rivoal, *Nombres d'Euler, approximants de Padé et constante de Catalan*, Ramanujan J. **11** (2006), no. 2, 199–214.
[42] Tanguy Rivoal and Wadim Zudilin, *Diophantine properties of numbers related to Catalan's constant*, Math. Ann. **326** (2003), no. 4, 705–721.
[43] Dan Romik, $\zeta(5)$ *is irrational: a Lean 4 formalization of A. Fauzan's proof*, Lean formalization, https://github.com/danromik/zeta5-irrationality, commit `fc2fefa`, 2026.
[44] Edward B. Saff and Vilmos Totik, *Logarithmic potentials with external fields*, Grundlehren der mathematischen Wissenschaften, vol. 316, Springer, Berlin, 1997.
[45] V. Kh. Salikhov, *On the irrationality measure of the number* π, Russian Math. Surveys **63** (2008), no. 3, 570–572.
[46] Zhi-Wei Sun, *Catalan's constant is irrational*, 2026, arXiv:2609.04176.
[47] The mathlib Community, *The Lean mathematical library*, Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), 2020, pp. 367–381.
[48] The PrimeNumberTheoremAnd contributors, *The PNT+ project: a Lean formalization of the prime number theorem and related results*, https://github.com/AlexKontorovich/PrimeNumberTheoremAnd, The vendored compatibility additions used here are recorded in the repository's `NOTICE` and `PORTING.md`.
[49] Dane Wachs, *A note on a recent claimed proof of the irrationality of Catalan's constant*, 2026, arXiv:2609.22339.
[50] Ryan Xu, *Improved* $\kappa = 2$ *trial:* $\mu(G) \le 727.03$, GitHub issue, October 22, 2026, https://github.com/slitsey/catalan-irrationality-measure/issues/5; data and scripts at https://github.com/ry-xu/slightlysupervised, 2026.
[51] Doron Zeilberger and Wadim Zudilin, *The irrationality measure of* π *is at most* 7.103205334137..., Mosc. J. Comb. Number Theory **9** (2020), no. 4, 407–419, arXiv:1912.06345.
[52] Wadim Zudilin, *A few remarks on linear forms involving Catalan's constant*, Chebyshevskii Sb. **3** (2002), no. 2(4), 60–70, In Russian, arXiv:math/0210423.

[53] ———, *An Apéry-like difference equation for Catalan's constant*, Electron. J. Combin. **10** (2003), Research Paper 14, arXiv:math/0201024.
[54] ———, *Two hypergeometric tales and a new irrationality measure of* $\zeta(2)$, Ann. Math. Québec **38** (2014), no. 1, 101–117, arXiv:1310.1526.
[55] ———, *A determinantal approach to irrationality*, Constr. Approx. **45** (2017), no. 2, 301–310, arXiv:1507.05697.
[56] ———, *Arithmetic of Catalan's constant and its relatives*, Abh. Math. Semin. Univ. Hambg. **89** (2019), no. 1, 45–53, arXiv:1804.09922.

Email address: sean.litsey.math@gmail.com