hexagon:2610.00187v6

AN IRRATIONALITY MEASURE FOR CATALAN'S CONSTANT

SEAN LITSEY

ABSTRACT. Let $G = \sum_{j\geq 0}(-1)^j(2j+1)^{-2}$ be Catalan's constant. We prove that $|G - a/b| > b^{-410}$ for all integers a and all sufficiently large integers b. Hence the irrationality exponent of G is at most 410. 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 every prime in an explicit range, which lowers the denominators. On the analytic side we bound a determinant that the energy method treats trivially, using Pick matrices, Heine's formula and an inequality for the logarithmic energy of signed measures. A quadrature formula extends the upper bound to an exponentially small neighbourhood of G, as a measure requires. The proof is formalized in the Lean proof assistant.

CONTENTS

1. Introduction 2
2. The construction and the criterion 7
3. Denominators and nonvanishing 9
4. Corank modulo p 15
5. The neighbourhood of G 25
6. The energy method with prescribed columns 30
7. A signed bound for the row determinant 33
8. Proof of Theorem 1.1 40
9. Remarks on the method 41
10. Formalization 42
Disclosure of AI use 42
Appendix A. Explicit subspaces for the remaining intervals 43
Appendix B. The bounds for the energy method 44

Date: November 1, 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, Heine's formula, Gauss quadrature, Lean.

Appendix C. The trial measures of Section 7 45
References 46

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 $L(2, \chi_{-3})$ irrational in 2024 [5]. For G the classical methods produce linear forms that do not decay fast enough, and its irrationality remained open. In October 2026 OpenAI [28] 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,*

$$\left| G - \frac{a}{b} \right| > b^{-410}.$$

In particular $\mu(G) \leq 410$.

The proof uses the prime number theorem. The theorem and every result it depends on have been verified in the Lean 4 proof assistant on top of OpenAI's formalization. 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 [3] obtained Apéry's forms for $\zeta(2)$ and $\zeta(3)$ as integrals of Legendre-type polynomials. Hata [17] 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 [31, 32] 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 [27] applied it to $\log 2$, and Zeilberger and Zudilin [41] to π, using integrals of Salikhov [35]. For Catalan's constant the analogous linear forms exist but do not decay fast enough. Rivoal and Zudilin [33] 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 $G = \beta(2)$. Zudilin [43] 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 [42, 20, 21, 46]. Eskandari, Murty and Nemoto [12] construct a mixed motive with period G and obtain linear forms in 1 and G from it, and Eskandari [11] improves their approximation exponent.

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 [29]. Liu, Jiang and Zhang proved $\mu(\log r) = 2$ for every rational $r > 0$ with $r \neq 1$ [26]. Giraudin lowered the bound for $\mu(\pi^2)$ to 5.0193784 with two hypergeometric constructions of Zudilin [44] at new parameters [15], and the bound for $\mu(\zeta(3))$ to 5.5138800 [16]. Leibtag, Barkan, Kalisch, Weinbaum, Shalyt and Kaminer [24] 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 [30] proved that $\mathrm{Li}_2(-1/4)$, $\mathrm{Li}_2(1/5)$ and $\mathrm{Li}_2(-1/3)$ are irrational. Fauzan announced a proof that $\zeta(5)$ is irrational, using Hankel determinants [13]. Lai combined this method with simultaneous approximations to show that 1, $\zeta(3)$ and $\zeta(4)$ are linearly independent over $\mathbb{Q}$ [22]. For $L(2, \chi_{-3})$ the effective holonomy bounds of Calegari, Dimitrov and Tang give $\mu(L(2, \chi_{-3})) \leq 24781$ [6, Thm. 5.1].

Determinant methods. OpenAI's argument [28] 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 [14] 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 [45].

The first irrationality measure for G was obtained by Jha [18], 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 $\beta(4)$. Bojnal [4] and Xu [40] obtained $\mu(G) < 2402$ and $\mu(G) \le 727.03$ at other parameters. Xu's certificate is numerical and not formalized. Davis [9] and Arkhipov [2] reached $\mu(G) \le 528.24$ with new quadrature nodes, the zeros of $T_K(P(t))$ for a fixed polynomial P, and Davis's result is formalized in Lean. These constructions are independent of the ideas of the present paper. An earlier preprint of Sun [36] claimed the irrationality of G, but Wachs [39] showed that its main estimate is not justified.

1.3. **The companion paper.** Catalan's constant is a difference of two values of the trigamma function ψ'. With

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

one has $G = \frac{1}{16}\Psi(\frac{1}{4},\frac{3}{4})$. The companion paper [25] proves irrationality measures for $\Psi(s,1)$ with $s = \frac{k}{4}, \frac{k}{3}, \frac{k}{5}, \frac{k}{7}, \frac{1}{2}$, and for $\Psi(\frac{1}{3},\frac{2}{3}) = 9L(2,\chi_{-3})$. For $\Psi(\frac{1}{4},1)$ and $\Psi(\frac{3}{4},1)$, which are rational multiples of $G \pm 5\pi^2/48$, and for $s = \frac{k}{5}, \frac{k}{7}$, these are the first proofs of irrationality that we know of. Its determinants are Löwner matrices of digamma differences, which 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.* $Q_p(a/b) \ne 0$ 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, $D \approx 6431$ 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 8.30 through a signed bound for a determinant of row functions (Section 7). The result is $\eta \approx 38.3$, which is what makes $d\rho/\eta$ as small as 410.

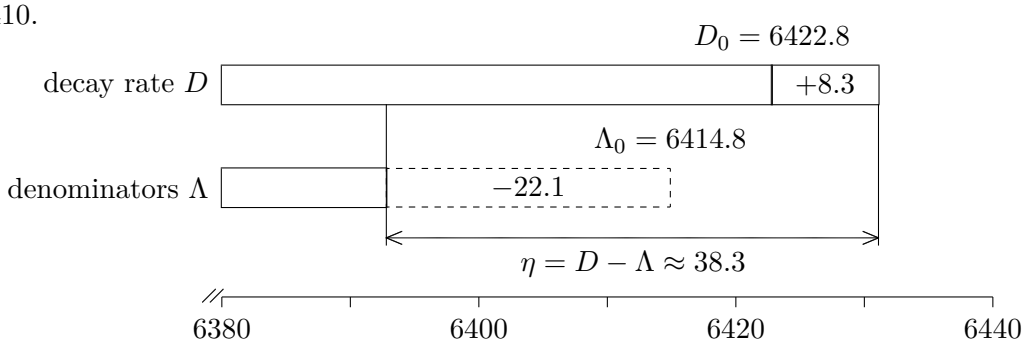


FIGURE 1. The exponent balance for the parameters $(58, 9, 3, 5, 3)$, per unit of N^2. The axis is broken at 6380. OpenAI's estimates give the decay rate D_0 and the denominator exponent Λ_0, which differ by about 8. The signed bound of Section 7 adds 8.3 to D_0, and the corank bounds of Section 4 remove 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 every prime in $(\frac{16}{7}N, \frac{53}{2}N]$. The kernel vectors are explicit. 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 22.05.

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. A form of the principle with a fixed rational subspace treats the remaining windows of primes, with explicit kernel vectors listed in the formalization.

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 perturbed determinant is then a sum of determinants in which some columns are replaced by values of the row functions at the nodes. We bound the weight of each replaced column and its generalized Vandermonde factor separately. The two bounds peak at different columns, which lowers ρ by about 5.

4. *A signed bound for the row determinant.* OpenAI's integrand contains a determinant $\det \mathsf{L}$ of size n that the energy method bounds by 6^n. We write it with a Pick matrix and reduce it, by a Schur complement and a contour bound, to the block of the negative nodes. Heine's formula writes that block as a ratio of two log-gas integrals. We bound the numerator from above and the denominator from below. An inequality for the logarithmic energy of signed measures (Theorem 7.8) keeps the combined estimate concave. This raises the decay rate by 8.3.

5. *Near-extremal configurations.* The signed bound needs hypotheses on the configuration. We show that they hold whenever the configuration comes close to the maximum of OpenAI's energy. On all other configurations OpenAI's bound already suffices.

The parameters of OpenAI's construction are kept symbolic throughout, so every statement except the final numerical verifications holds for an arbitrary choice of them. We use $(58, 9, 3, 5, 3)$.

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 signed

degree	d	53
neighbourhood exponent	$\rho = 58c_2 + W_0$	296.59
decay rate (OpenAI's estimate)	D_0	6422.84
decay rate (with the signed bound)	D	6431.13
denominator exponent (with the corank gain)	Λ	6392.79
difference	$\eta = D - \Lambda$	38.34
bound	$d\rho/\eta$	409.99…

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

bound, η would be about 8, and the method would give an exponent near 2000.

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 signed 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 [28] for the construction and refer to it for details. Citations of the form [28, 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 $U_{|r-g|-1}$ [28, (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 [28, 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 $i \ne j$, $H_i = \sum_{k\le i} 1/k$ [28, 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*

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

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. Every prime $p' \in (p, u_1]$ has $p'^2 \mid b$, and $p' \le u_1 < 2u_0 < 2p$ because $2\omega < \lambda$. Let Π be the product of these p'^2. Then $\Pi \mid b$ and $\log \Pi = 2(\vartheta(u_1) - \vartheta(p))$.

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^{dp} divides the integer. Since $p^2 \nmid b$, $Q_p(a/b) \ne 0$ by (b). So $b^{dp}Q_p(a/b)/\Pi^{dp}$ is a nonzero integer, and

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

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 $\eta p^2 \le dp\big(\log b - 2(\vartheta(u_1) - \vartheta(p))\big)$. With $\vartheta(u_1) - \vartheta(p) \ge u_1 - p - 2\delta u_1$ this is

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

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. □

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 second ensures that the larger primes of the interval lie below twice the chosen one, so that their squares can be recovered from the coefficients of Q_p. The last makes the final inequality contradictory. 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 them. Choosing any prime of the interval would lose a power of b. Taking the lower end $u_0 = (\lambda - \omega)\log b$ as a free parameter allows $\lambda < 2$, since the recovery needs only $u_1 < 2u_0$.

For Catalan's constant we take $d = 53$, $\omega = \frac{3}{5}$, $\delta = 2.9 \cdot 10^{-7}$ and $\lambda = 410/\rho$. The last hypothesis then reads $410 > B_*$ with $B_* = d\rho/(\eta - 4d\delta) = 409.99999\ldots$

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 [28, §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 $\sum_{v=0}^q(-1)^v\binom{q}{v}F[t+v] = [X^{t+q}]((X-1)^qF)$ for every polynomial F, Lemma 3.2 gives

$$p^2\sum_r c_r\mathcal{A}_N^0(r,k) \equiv \sum_{u\ge0}\Big(\alpha_{u,J}\,Q_c[up+\varsigma] + \beta_{u,J}\,\widehat{D}_c[up+\varsigma]\Big), \tag{3.1}$$

where

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

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 $\det\big(p^2(\mathcal{A}_N^0+z\mathcal{L}_N)\big) = p^{2n}\Delta_N(z)$. □

3.4. **Pole counts.** Let p be an odd prime with $\sqrt{H}<p\le H$. OpenAI's method [28, §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 $\ell_N(p) = N f(p/N)$ 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 [28, §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 $m, \ell \ge 0$, $m + \ell \le n$ and $C \ge n$, where

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

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

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

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

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

$$\Lambda = \frac{2673}{8} \log 2 + 116 + \frac{24269}{4} - \int_0^\infty \kappa(\theta)\, d\theta = 6392.7880\ldots$$

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 $i \geq 0$ 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

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

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

$$Q_N(X/p'^2) = (d_0N)!\,\frac{\mathcal{D}_N}{p'^{2n}}\det(p'^2\mathcal{A}_N^0 + X\mathcal{L}_N)$$

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, $\Delta_p(z) \neq 0$. In particular $\Delta_p(a/b) \neq 0$ 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 $H = H_0p < p^2/4$. 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 $q = q_0p$, $C = C_0p$, $h = h_0p$ and $g = g_0p$, 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 $s^{b_0p+k}(1-s)^{q_0p} = s^\ell(s^p)^{b_0+k_0}(1-s^p)^{q_0}$ 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 [28, 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 $p^2z\mathcal{L}_p = p\,(pz)\,\mathcal{L}_p$ 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 exceeds every 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 $\mathcal{A}_{N,p}^0$ 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 $c^{\mathsf{T}} \mathcal{A}_{N,p}^0$ 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 $\mathcal{A}_{N,p}^0$. 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. Recall $e = (p + 1)/2 - q - h$.

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* $2g < p$, $H \le 4p$ *and* $k_T = \min(g,\, n-1-g-2p,\, C-3p) \ge 1$*:*

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

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

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

(c) *(Theorem 4.6) If* $e \ge 0$, $n \le 4p+2g+1$, $L \le 6p$ *and* $b+q \le p+\varsigma_0$*, 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* $e \ge 0$ *and* $n \le 4p+2g+1$*:* $m_1 - m_0 - e$*, 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 $12N < p \le 26N$, 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* $2g < p$, $H \le 4p$ *and* $k_T = \min(g,\, n-1-g-2p,\, C-3p) \ge 1$. *Then* $\nu_N(p) \ge k_T - (n+q+g+h-3p)_+$.

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

Proof. For $1 \le d \le k_T$ 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 $d \le g$ and $g+2p+d \le n-1$. 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 $J \geq 3$ do not meet the support, since $3p+\ell \geq C-d \geq C-g = L$. Choose $x^* \neq 0$ 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 $\ell \geq b$, and the value there is $f_0(x^*) \sum_d s_d \varrho_\ell(d)$, the coefficient of $X^{3p+\ell}$ in $\sum_d s_d R_d$. 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,*

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

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 $2h \leq p+1$, $L \leq 3p \leq H$, $n+q+g+h \leq 3p$ and $2p \leq n+q$, and put $n^* = \min(n, 2p+2g+1)$. Then $\nu_N(p) \geq n^* - g - 2p - 1$.*

Proof. We restrict c to the rows $r < n^*$, so that $m_r \leq n^* - 1 - g \leq 2p+g$. 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 $|p-m| \leq p+g = p+C-L$. Multiplying by $(1-X)^h$ and summing, we get for every c supported on these rows

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

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 $(1-X^2)^p = 1 - X^{2p}$,

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

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 XWP_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 $Q_c = (-1)^q X(1-X)^{q+h+(p-1)/2}(1+X)^{(p-1)/2} S_c$, and $q+h+(p-1)/2+e = p$. Since $(1-X)^p = 1-X^p$,

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

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 $u \geq 4$ of $\widehat{D}_c$ and $u \geq 5$ 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 $b+q \leq \varsigma < p$ 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, $(1-X)^e Q_c = (1-X^p)Y_c$ with Y_c a polynomial. Put $\sigma_c(\ell) = \sum_{u\geq 0} Q_c[up+\ell]$ and sum the coefficients of this identity over the residue class of $\varsigma \geq e$ 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 $\varsigma - e \geq m_0$ 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

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

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 $\dim \mathcal{F} \leq 4w-(w-e)$, 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 $x = (x_1, \ldots, x_t)$;
(b) polynomials $\pi_1, \ldots, \pi_t \in \mathbb{F}_p[X]$ 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 $h \geq p$ 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 $h \geq p$, and on $(3N, \frac{16}{5}N]$.

For the intervals of primes below $12N$ that these families do not cover, we use the counting inequality in a special form. The space of row combinations is the tensor product of a space of polynomials with a fixed rational subspace $V \subset \mathbb{Q}^s$, and the conditions are Taylor coefficients composed with fixed linear maps. The subspaces V are explicit but have no closed form. Appendix A states this form (Theorem A.1) 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* 102 *intervals* $(a_i, b_i]$, *with* $\frac{16}{7} = a_1 < b_1 \leq a_2 < \cdots < b_{102} = \frac{53}{2}$, *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* $a_i N < p \leq b_i N$ *for some* i,

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

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

$$\int_0^\infty \kappa(\theta)\, d\theta = \frac{362355158746019386015965 98}{16426096632711159925781 25} = 22.05972\ldots \;\geq\; 22.05325.$$

The contributions by ranges of $\theta = p/N$ are as follows. The third column counts the intervals on which the bound is that of Theorem A.1 with an explicit subspace.

range of θ	intervals	with explicit subspace	area
(16/7, 3]	11	0	0.06259
(3, 6]	46	23	0.25136
(6, 12]	27	9	0.71965
(12, 53/2]	18	2	21.02610
total	102	34	22.05972

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

interval of θ	$\kappa(\theta)$	bound
$(12, 13]$	$\theta/2 - 5$	Theorem 4.2(c), $(p-1)/2 - g$
$(13,\ 146/11]$	$73 - 11\theta/2$	Theorem 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)
$(809/50,\ 178/11]$	$73 - 9\theta/2$	Appendix A
$(178/11,\ 1619/100]$	$\theta - 16$	Appendix A
$(1619/100,\ 118/7]$	$\theta - 16$	(F3)
$(118/7,\ 69/4]$	$9\theta/2 - 75$	Theorem 4.2(d), $(9p - 2C)/2$
$(69/4,\ 35/2]$	$\theta/2 - 6$	Theorem 4.2(d), $(p-1)/2 + h - b$
$(35/2,\ 128/7]$	$64 - 7\theta/2$	Theorem 4.2(d), $(2(n+q+h) - 7p)/2$
$(64/3,\ 23]$	$3\theta - 64$	Theorem 4.2(a), $3p - (n+q+h)$
$(23,\ 24]$	5	Theorem 4.2(a) up to $\theta = 70/3$, then (b)
$(24,\ 26]$	$53 - 2\theta$	Theorem 4.2(b), $n - g - 2p$
$(26,\ 53/2]$	$53 - 2\theta$	(F1)

The areas of the rows of this table add up to $21.02610\ldots$, and the 84 intervals below 12 contribute $1.03361\ldots$, so the total is $22.05972\ldots$. The exact sum is computed in the formalization from the list of intervals. Figure 2 shows κ on $(12, \frac{53}{2}]$.

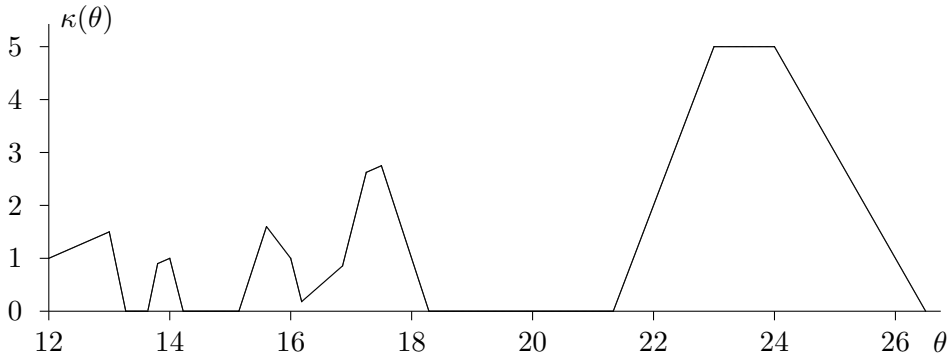


FIGURE 2. The corank profile $\kappa(\theta)$ on $(12, \frac{53}{2}]$, where $\nu_N(p) \geq N\kappa(p/N) - 2$ for $p = \theta N$. The bound on the pole count is $f(\theta) = 116$ on this whole range, so κ is the gain over OpenAI's denominators.

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 cap $g = 5N$ 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

$$\Gamma_0 = 8.29824, \qquad U_0 = U + \Gamma_0/n_0^2 = 2.2583256\ldots,$$
$$D = n_0^2(U_0 - \tfrac{1}{2}\log 2) = 6431.1338\ldots$$

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

(5.1)
$$\rho = n_0 c_2 + W_0, \qquad c_2 = 3.596893, \quad W_0 = 87.9663, \quad \rho = 296.586094.$$

Here c_2 is the increase of the energy bound per prescribed column (Section 6), and W_0 bounds the logarithmic size of the weight and the binomial 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 $t_1, \dots, t_J$ be distinct points, with J at least the number of degrees of freedom of the rows. Lagrange interpolation gives

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

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 row functions that define $\mathcal{A}_N$. So the perturbed determinant can be treated by the energy method. The price is the size of the weights.

We take $J = d_0 N = 53N$ nodes. Put $g_* = 401/2500$, $a_* = 2g_*/(1+g_*^2) = 2005000/6410801$ and $J' = \lceil J/2 \rceil$. The nodes are the points

$$t_{j,\pm} = \pm\sqrt{a_*^2 + (1-a_*^2)(1+s_j)/2}, \qquad T_{J'}(s_j) = 0,$$

that is, the roots of $T_{J'}\big((2t^2-(1+a_*^2))/(1-a_*^2)\big)$. We keep all J' positive roots and the first $J - J'$ negative ones, so for odd J exactly one root is dropped. The nodes lie in $S = (-1,-a_*) \cup (a_*,1)$, and in the coordinate x with $t = 2x/(1+x^2)$ every node has $|x| > g_*$. The nodes avoid a neighbourhood of $x = 0$, where the fields of the energy method are largest, and they are Chebyshev nodes for S. Dropping a root changes the nodal polynomial by a linear factor, so it changes the derivative bound and the size of the Lagrange polynomials only by factors polynomial in N. So an odd number of nodes does not change the exponent. This is needed, since $J = 53N$ is odd for odd N.

Let g_S be the Green function of $\mathbb{C}\setminus\bar S$ with pole at infinity. In closed form

$$g_S(z)=\tfrac12\operatorname{arccosh}\Big(\frac{|z^2-1|+|z^2-a_*^2|}{1-a_*^2}\Big).$$

The Lagrange polynomials satisfy $|\ell_a(z)|\le e^{53Ng_S(z)+o(N)}$ at distance at least c/N from the nodes, by the derivative identity for Chebyshev polynomials at their roots. 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 by a contour that depends on k.

Lemma 5.2 (keyhole bound)**.** *Let P be a real polynomial, $E\ge0$, and $F(t)=t^{-E}P(t)$. Let $0<r\le1$, and suppose that $|F|\le M$ on the circle $|t|=r$ and $|F|\le M'$ on the segments $r\le|t|\le1$ of the real axis. Then*

$$\Big|4\sum_{l\ge0}c_l\,[t^{E+2l}]P\Big|\le 4(1+E)\max(M,M').$$

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 $c_l=\frac1\pi\int_{-1}^{1}t^{2l}(1-t^2)^{-1/2}dt$, its value is $\frac4\pi\int_{-1}^{1}\Pi(t)(1-t^2)^{-1/2}dt$, which is at most $4\max_{[-1,1]}|\Pi|$. By Cauchy's estimate $|\pi_i|\le Mr^i$, so $|R(t)|\le EM$ for $r\le|t|\le1$, and $|\Pi|\le M'+EM$ there. For $|t|<r$ the maximum principle gives $|\Pi(t)|\le\max_{|t'|=r}|\Pi|\le(1+E)M$. □

The formal proof replaces the integral by the finite Gauss–Chebyshev identity $\frac1M\sum_{j<M}\cos^{2l}\frac{(2j+1)\pi}{2M}=c_l$ for $l<M$, and Cauchy's estimate by the finite Cauchy formula at roots of unity. So no integration is needed. For $0\le y\le1$ and $z\in\mathbb{C}$ let

$$\Phi_y(z)=(10-58y)\log|z|+6\log|1-z|+53g_S(z).$$

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

$$|\beta'_k|\le e^{(w(y)+o(1))N},\qquad w(y)=\min_r\max_{z\in\Gamma_r}\Phi_y(z),\tag{5.3}$$

where Γ_r is the circle $|z|=r$ together with the segments $r\le|z|\le1$ of the real axis. For y near 1 the factor $|z|^{10-58y}$ favours a large circle, and for small y a small one. The circle $|z|=1.001$ alone gives $w(y)\le52.79$ for every y. The keyhole contour improves this only slightly at the worst column. Its use is elsewhere: it lowers $w(y)$ at the columns where the binomial factor of Lemma 5.5 is largest.

5.3. **Column expansion and Andréief's identity.** Write $\mathcal{P}(r,a) = A_r(t_a)$. 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 $\prod_{f(k)\neq\varnothing} \varepsilon\beta_{f(k),k} \det M_f$. 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 $(A_r(t_{f(k)}))_r$ otherwise.

Lemma 5.3. *Suppose that* $|\det M_f| \le M_0 \prod_{f(k)\neq\varnothing} C_k$ *for every* f*, and that* $|\beta_{a,k}| \le B_k$*. Then*

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

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 5.4. *Let* m_1, m_2 *be finite measures, and let* F_r $(r < n' + m)$*,* F_i' $(i < n')$ *and* k *be integrable functions. 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* $B(r, n'+j) = F_r(t_j')$ *for* $j < m$*. Then*

$$(n'!)^2 \det B = \iint \det\big[F_r((t\|t')_c)\big]_{r,c} \cdot \det[k(t_i,s_j)]_{i,j<n'} \det[F_i'(s_j)]_{j,i<n'}\, dm_1^{n'}(t)\, dm_2^{n'}(s),$$

where $t\|t'$ *is the list of the* n' *variables followed by the* m *prescribed points.*

Proof. For $m = 0$ this is Andréief's identity applied twice [14]. 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. □

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 5.5. *Let* $s_1, \dots, s_{n'} \in [0,1]$*, 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*

$$|\det[s_l^{E_i}]| \le \prod_{j=1}^{m} \mathcal{B}(k_j) \prod_{i<l} |s_i - s_l|, \qquad \mathcal{B}(k) = \sum_{i=0}^{m-1} \binom{n'}{n'-k+i},$$

with $\binom{n'}{r} = 0$ *for* $r \notin [0, n']$*.*

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

$|Q_{j,i}| = e_{n'-k_j+i}(s) \le \binom{n'}{n'-k_j+i}$, since $s \in [0,1]^{n'}$. Hadamard's inequality, applied to the rows of Q in the ℓ^1 norm, gives the claim. □

5.4. **The weight of one column.** The bound of Lemma 5.5 does not depend on the points s_l. Since $\binom{n'}{r} \le \binom{n}{r}$,

$$\log \mathcal{B}(k) \le n_0 \tilde{h}_2(k/n) N + O(\log N),$$

where $\tilde{h}_2(y)$ is the binary entropy $h_2(y) = -y \log y - (1-y)\log(1-y)$ for $y \ge \frac{1}{2}$ and $\log 2$ for $y < \frac{1}{2}$. So the replaced column k contributes at most $e^{(w(y)+n_0\tilde{h}_2(y)+o(1))N}$, with $y = k/n$. Put

$$W_0 \ge \sup_{0 \le y \le 1} \big(w(y) + 58\tilde{h}_2(y)\big).$$

Appendix B proves $W_0 = 87.9663$.

The two factors peak at different columns. The weight $w(y)$ increases up to $y \approx 0.84$, while $\tilde{h}_2(y)$ is largest on $[0, \frac{1}{2}]$. So the maximum of $w + 58\tilde{h}_2$ is about 87.96, at $y \approx 0.62$, well below the sum $52.79 + 40.20$ of the separate maxima (Figure 3). This saving of about 5 in ρ is the reason for treating the columns one at a time.

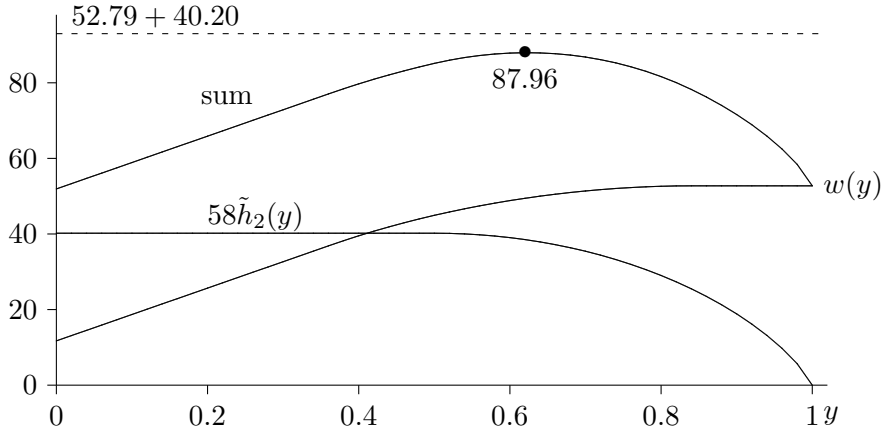


FIGURE 3. The weight $w(y)$ of column $k = yn$, the binomial factor $58\tilde{h}_2(y)$, and their sum, per unit of N. The values were computed numerically on a grid of step 0.02 and are not part of the proof. The certified bound for the sum is $W_0 = 87.9663$. The dotted line is the sum of the separate maxima.

The remaining factor of ρ is $n_0 c_2$, the increase of the energy bound caused by the prescribed columns. It comes from Section 6, and with it $\rho = 58 \cdot 3.596893 + 87.9663 = 296.586094$.

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.

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 $t = 2x/(1+x^2)$ [28, §5]. It has the form

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

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 [28, Props. 5.2, 5.3]. 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 6^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

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

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 [34]. 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, which satisfy $|x| > g_*$, and points $y \in (0,1)$. 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_2 enters ρ.

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

$$F_1(x) = -\gamma_d \log(1+x^2) - 2\mathsf{T}(u,x), \qquad g_* \le |x| \le 1.$$

For those of the second kind it is

$$F_2(y) = -\gamma_d \log(1+y^2) + \frac{n_0 - 2g_0 - 1}{n_0} \log y - u_1 y, \qquad 0 < y < 1,$$

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 $(n-1-2g)/n$. By Lemma 5.5 the generalized Vandermonde determinant contributes only the factor $\prod_{i<l} |s_i - s_l|$ to the energy, and the binomial factors are assigned to the columns. The energy bound then reduces to a quadratic inequality in α and α_2.

Lemma 6.1 (energy inequality with prescribed points)**.** *Let $\ell = \log 2$, and let $\mathcal{N}, \mathcal{F}, \mathcal{R}, \mathcal{F}_2$, c_2 and $m_0 \ge 0$ be real numbers with*

$$\mathcal{N} + \mathcal{F} + U - a_0\ell \le -m_0, \qquad \ell - \mathcal{R} + \mathcal{F}_2 \le 0,$$
$$\big((a_1-1)\ell - \mathcal{F} + \mathcal{R} - c_2\big)^2 \le 4\ell m_0.$$

Then for all $\alpha, \alpha_2 \ge 0$

$$\big(-(a_0 - \tfrac12) + (a_1-1)\alpha - \alpha^2 + \alpha_2 - \alpha_2^2\big)\ell + \mathcal{N} + (1-\alpha)\mathcal{F} + (\alpha - \alpha_2)\mathcal{R} + \alpha_2\mathcal{F}_2 \le -U + \tfrac12\ell + c_2\alpha.$$

Proof. The left side minus the right side equals

$$(\mathcal{N} + \mathcal{F} + U - a_0\ell) + \alpha\big((a_1-1)\ell - \mathcal{F} + \mathcal{R} - c_2\big) - \ell\alpha^2 + \alpha_2(\ell - \mathcal{R} + \mathcal{F}_2) - \ell\alpha_2^2.$$

The last two terms are nonpositive by the second hypothesis. The first three form a concave quadratic in α. It is nonpositive for all α, since $a + s\alpha - \ell\alpha^2 \le 0$ for all α as soon as $s^2 + 4\ell a \le 0$. □

The lemma is applied with $\mathcal{N}$ a bound for $2\|u\|^2 + \frac12\|v\|^2$, $\mathcal{F} = M_X + M_Y$, $\mathcal{R}$ a bound for the supremum of F_1 and $\mathcal{F}_2$ a bound for the supremum of F_2. The analogue for the constrained case has no prescribed points of the second kind. Its polynomial in α is

$$\big(-(a_0 - \tfrac12) + (a_1-1)\alpha - \alpha^2\big)\ell + n_1 + (1-\alpha)\mathcal{F}_1 + \alpha\lambda_1\tfrac{2g_0}{n_0} \le -U + \tfrac12\ell + c_1\alpha,$$

with n_1 a bound for the norms of the trial sequences and $\mathcal{F}_1 = M_{X_1} + M_Y$. It holds for all real α when its discriminant is nonpositive and $c_1 \le (a_1 - 1)\ell - \mathcal{F}_1 + \lambda_1 \frac{2g_0}{n_0}$.

For our parameters the trial sequences of Appendix B give, with $U = 2.255858836$, the hypotheses of Lemma 6.1 with $m_0 = 2.986 \cdot 10^{-6}$ and $c_2 = 3.596893$. The constrained case admits $c_1 = 3.549542 < c_2$.

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

$$u_k = c_k + \operatorname{Re} \sum_j r_j z_j^k \qquad (k \ge 1;\ c_k = 0 \text{ for } k > 160), \tag{6.1}$$

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 $|x| \le g_*$. Appendix B gives the charges 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. In the constrained case the inequality of Lemma 6.1 still holds at U_0 for every α, with the slope c_2. In the case with doubled interaction it holds at U_0 only for $\alpha \ge \frac{1}{16}$, because the extra term $U_0 - U$ is absorbed by the term $c_2\alpha$ once α is not small. So a new argument is needed only in the case with doubled interaction with fewer than $n/16$ prescribed columns. Section 7 supplies it, with the same constant c_2. Precisely, it proves the following.

Proposition 6.2. *In the case with doubled interaction, for every configuration with $m \le n/16$ prescribed columns, the integrand of Lemma 5.4, with the generalized Vandermonde determinant replaced by the bound of Lemma 5.5, is at most*

$$\exp\Big(n^2\big(-U_0 + \tfrac{1}{2}\log 2 + c_2\alpha + o(1)\big)\Big)\, C_0^m \prod_j \mathcal{B}(k_j), \qquad \alpha = m/n,$$

with an absolute constant C_0.

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 $k_1, \dots, k_m$ is at most $C_0^m \prod_j \mathcal{B}(k_j) e^{n_0 c_2 N m} \exp\big(n_0^2 N^2(-U_0 + \frac{1}{2}\log 2 + o(1))\big)$. The factor C_0^m comes from the expansion of the row functions at the prescribed points. By (5.3) and the bound for $\mathcal{B}(k)$, the constants of Lemma 5.3

satisfy

$$\log(B_k C_k) \le \bigl(n_0 c_2 + w(k/n) + n_0 \tilde{h}_2(k/n) + o(1)\bigr) N \le (\rho + o(1)) N.$$

For $|\varepsilon| \le e^{-\rho N}$ Lemma 5.3 gives a factor $(1 + O(N^5))^n = e^{o(N^2)}$. With $D = n_0^2 (U_0 - \frac{1}{2}\log 2)$ this is the claim. □

7. A signed 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 6^n. We prove a bound for it that is exponentially small in N^2 on the near-extremal configurations. The argument is the chain

(7.1)

$$|\det \mathsf{L}| \le 11^{k_+} \det \mathsf{L}_{--} = 11^{k_+} \prod_{b<c} (1 - x_b x_c) \prod_{b,p} (1 - x_b x_p)^{-1} \cdot \frac{Z_s(\Pi_- \Pi_+)}{Z_s(\Pi_+)}.$$

The inequality comes from a Pick matrix, a Schur complement and a contour bound (§7.1–7.3). The identity is Heine's formula (§7.4). The last factor is a ratio of two log-gas integrals. We bound the numerator from above and the denominator from below (§7.5). Positive nodes close to 1 need a separate argument (§7.6). An inequality for the logarithmic energy of signed measures keeps the combined estimate concave (§7.7). Prescribed columns add prescribed rows (§7.8). The hypotheses of the reduction hold on all near-extremal configurations (§7.9). The section ends with the trial measures (§7.10).

7.1. The determinant of the row functions.

In the case with doubled interaction OpenAI's integrand contains the determinant of the row functions at points $x_1, \dots, x_n \in (-1, 1) \setminus \{0\}$, which we call the nodes [28, Prop. 5.3]. With $\mathsf{e} = n - 1 - 2g$ it factors as

$$\det \varphi(x) = C_{n,g}\, V(x) \prod_a |x_a|^{-(n-1-g)} \det \mathsf{L}(x), \tag{7.2}$$

where $C_{n,g}$ depends only on n and g, and $V(x) = \prod_{a<b}(x_b - x_a)$. To describe L let

$$B(z) = \prod_a \frac{z - x_a}{1 - x_a z}, \qquad \mathcal{K}_{ab} = \frac{1}{1 - x_a x_b},$$

be the Blaschke product of the nodes and the Pick matrix of the Szegő kernels $k_a(z) = 1/(1 - x_a z)$. Put $w_a = 1$ at a negative node and $w_a = -5$ at a positive node. The value $-5 = 1 - 6$ comes from the factor $\frac{3}{2}$ in the row functions A_r. With $E = \operatorname{diag}(w_a B'(x_a) x_a^{\mathsf{e}})$ one has $\det \mathsf{L} = \det(\mathcal{K} + E)/\det \mathcal{K}$. In the basis k_a the operator L acts on the model space $\mathcal{K}_B = H^2 \ominus B H^2$, which has dimension n. OpenAI's argument uses only $\|\mathsf{L}\| \le 6$, that is $|\det \mathsf{L}| \le 6^n$, 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). We now prove a bound of this order.

Let k_- and k_+ be the numbers of negative and positive nodes. For a weight $w \in \mathbb{R}$ let $F(w) = \mathcal{K} + \operatorname{diag}(E_-, wE_+)$, where E_- and E_+ are the diagonal blocks of $\operatorname{diag}(B'(x_a)x_a^{\mathsf{e}})$ at the negative and the positive nodes. So $\det \mathsf{L} = \det F(-5)/\det \mathcal{K}$. The case $w = 1$ is a Pick matrix.

Lemma 7.1. *For every configuration* $0 \le F(1) \le 2\mathcal{K}$.

Proof. For a real vector c put $\phi(z) = B(z)\sum_a c_a/(z - x_a)$. This is the function of $\mathcal{K}_B$ with node values $\phi(x_a) = c_a B'(x_a)$, and $\|\phi\|^2_{H^2} = c^{\mathsf{T}}\mathcal{K}c$. By the residue theorem on the unit disc

$$c^{\mathsf{T}}(F(1) - \mathcal{K})c = \sum_a \frac{x_a^{\mathsf{e}}\phi(x_a)^2}{B'(x_a)} = \frac{1}{2\pi i}\oint_{|z|=1} \frac{z^{\mathsf{e}}\phi(z)^2}{B(z)}\,dz.$$

On the unit circle $|z^{\mathsf{e}}\phi^2/B| = |\phi|^2$, so the right side lies in $[-c^{\mathsf{T}}\mathcal{K}c,\ c^{\mathsf{T}}\mathcal{K}c]$. □

7.2. **A Schur complement bound.** The determinant $\det F(-5) = \det(F(1) - 6\operatorname{diag}(0, E_+))$ is a perturbation of a Pick matrix in the positive block only. The following inequality bounds it by the negative block.

Theorem 7.2. *Let* $M = \begin{pmatrix} M_1 & R^* \\ R & M_2 \end{pmatrix}$ *be a Hermitian block matrix with* $0 \le M \le \beta$ *and* $M_1 > 0$. *Let* T *be Hermitian of the size of* M_2, *with* $-\tau \le T \le \tau$, *and let* $\gamma \ge 0$. *Then*

$$\left|\det \begin{pmatrix} M_1 & R^* \\ R & M_2 - \gamma T \end{pmatrix}\right| \le (\beta + \gamma\tau)^{\dim M_2} \det M_1.$$

Proof. The determinant equals $\det M_1 \cdot \det(S - \gamma T)$, where $S = M_2 - RM_1^{-1}R^*$ is the Schur complement. Since $M \ge 0$, we have $S \ge 0$, and $S \le M_2 \le \beta$. So $-\gamma\tau \le S - \gamma T \le \beta + \gamma\tau$, and every eigenvalue of $S - \gamma T$ has modulus at most $\beta + \gamma\tau$. □

We apply this in the metric of $\mathcal{K}$, that is, after the congruence by a Cholesky factor of $\mathcal{K}$. By Lemma 7.1 the hypotheses hold with $\beta = 2$ and $\gamma = 6$, provided that $\pm E_+ \le \tau\mathcal{K}$ as quadratic forms. The result is

$$|\det \mathsf{L}| \le (2 + 6\tau)^{k_+} \det \mathsf{L}_{--}, \qquad \det \mathsf{L}_{--} = \frac{\det(\mathcal{K}_{--} + E_-)}{\det \mathcal{K}_{--}}, \tag{7.3}$$

where $\mathcal{K}_{--}$ is the block of $\mathcal{K}$ at the negative nodes. The factor $(2 + 6\tau)^{k_+}$ is $e^{O(N)}$. All the information of order N^2 is in $\det \mathsf{L}_{--}$, which depends on the positive nodes only through the values $B_+(x_b)$ of their Blaschke product at the negative nodes.

7.3. **A contour bound.** It remains to bound τ. The quadratic form of E_+ is the sum of the residues of $z^{\mathsf{e}}\phi^2/B$ at the positive nodes, which lie in the right half of the disc. So we integrate over the boundary of the right half-disc instead of the circle.

Lemma 7.3. *Let $M_* = \sup_{-1<y<1} |y|^{\mathsf{e}}/|B(iy)|$. For every real vector c*

$$\Big|\sum_{x_a>0} B'(x_a) x_a^{\mathsf{e}} c_a^2\Big| \le \Big(1+\frac{M_*}{2}\Big) c^{\mathsf{T}} \mathcal{K} c.$$

If $\frac{\mathsf{e}}{n} - \frac{1}{n}\sum_a \Big(\frac{t^2}{t^2+x_a^2} - \frac{x_a^2 t^2}{1+x_a^2 t^2}\Big) > 0$ for every $t \in (0,1]$, then $M_* = 1$, and the bound is $\frac{3}{2} c^{\mathsf{T}} \mathcal{K} c$.

Proof. With ϕ as in the proof of Lemma 7.1, the sum is the contour integral of $z^{\mathsf{e}}\phi^2/B$ over the boundary of the right half-disc, divided by $2\pi i$. No node lies on this boundary. On the arc the integrand has modulus $|\phi|^2$, so the arc contributes at most $\|\phi\|^2$. On the diameter $z = iy$ the modulus is at most $M_*|\phi(iy)|^2$. By the Fejér–Riesz inequality on a diameter, $\int_{-1}^{1} |\phi(iy)|^2 dy \le \pi\|\phi\|^2$. So the diameter contributes at most $\frac{1}{2} M_* \|\phi\|^2$.

For the second statement write $\log(|y|^{\mathsf{e}}/|B(iy)|) = nv(y)$. Then $v(y) = \frac{\mathsf{e}}{n}\log|y| + \frac{1}{2n}\sum_a \log\frac{1+x_a^2 y^2}{y^2+x_a^2}$, and $v(1) = 0$ because $|B(i)| = 1$. Differentiation gives $tv'(t) = \frac{\mathsf{e}}{n} - \frac{1}{n}\sum_a (\frac{t^2}{t^2+x_a^2} - \frac{x_a^2 t^2}{1+x_a^2 t^2})$. If this is positive on $(0,1]$, then $v < 0$ on $(0,1)$. Since v is even, $M_* = 1$. □

The formal proof avoids the residue theorem. It subtracts the poles with divided differences, integrates the remaining holomorphic function over the half-disc with primitives, and proves the Fejér–Riesz step by the same device on the upper half-disc. With $\tau = \frac{3}{2}$, (7.3) becomes

$$|\det \mathsf{L}| \le 11^{k_+} \det \mathsf{L}_{--}. \tag{7.4}$$

The condition on t asks that the nodes are not too concentrated near 0 at any scale. At the extremal measure its left side is at least 0.33 for every t.

7.4. **Heine's formula for the negative block.** For $s \ge 1$ and a positive continuous function f on $[-1,1]$ put

$$Z_s(f) = \int_{[0,\pi]^s} \prod_{l<l'} (\cos\theta_l - \cos\theta_{l'})^2 \prod_l f(\cos\theta_l)\, d\theta.$$

By Heine's formula $Z_s(f)$ is $s!$ times the Hankel determinant of the moments of $f(u)(1-u^2)^{-1/2}du$. Let

$$\Pi_-(u) = \prod_{x_b<0} (1 + x_b^2 + 2x_b u), \qquad \Pi_+(u) = \prod_{x_p>0} (1 + x_p^2 + 2x_p u)^{-1}.$$

Theorem 7.4. *Let $s = k_+ - g$. If $s \ge 1$ and $k_- \le \mathsf{e}+1$, then*

$$\det \mathsf{L}_{--} = \prod_{b<c}(1-x_b x_c) \prod_{b,p}(1-x_b x_p)^{-1} \cdot \frac{Z_s(\Pi_-\Pi_+)}{Z_s(\Pi_+)},$$

where b, c run over the negative nodes and p over the positive nodes.

Sketch. The full proof is formalized. We sketch it for orientation. One first writes $\det \mathsf{L}_{--} = \det[x_b^j + B_+(x_b)x_b^{m'-j}]_{b,j<k_-}/V(x_-)$ with $m' = \mathsf{e} + k_- - 1$, by Lagrange interpolation at the negative nodes. Since $1 + x^2 + 2xu = x(x + x^{-1} + 2u)$, the factor Π_- is a polynomial in u with roots $-\frac{1}{2}(x_b + x_b^{-1})$. By Christoffel's formula, $Z_s(\Pi_-\Pi_+)/Z_s(\Pi_+)$ is a determinant of the monic orthogonal polynomials of the weight $\Pi_+(u)(1-u^2)^{-1/2}$ at these roots, divided by a Vandermonde determinant. That weight is of Bernstein–Szegő type: Π_+ is the reciprocal of $|P_+(e^{i\theta})|^2$ for the polynomial $P_+(z) = \prod_p(1 + x_p z)$. Its orthogonal polynomials of degree at least k_+ are explicit. At $u = -\frac{1}{2}(x + x^{-1})$ they are Laurent expressions in x built from P_+. Since $P_+(z)/z^{k_+}\overline{P_+}(1/z)$ evaluated at $-x_b$ is $B_+(x_b)$ up to sign, the two determinants agree after the products of Cauchy type are collected. □

The formal proof works with the angle variables throughout. It needs Heine's formula, Andréief's identity and the Bernstein–Szegő polynomials, but no integration over orthogonal groups. Both integrals in Theorem 7.4 are log-gas integrals. There are s particles $u_l = \cos\theta_l$ on $[-1, 1]$ with logarithmic repulsion, in the external field $\log\Pi_- + \log\Pi_+$ or $\log\Pi_+$. Each field is a sum of logarithmic potentials $l(x, u) = \log(1 + x^2 + 2xu)$ of the nodes.

7.5. **Bounds for the two integrals.** The numerator is bounded from above by the method of OpenAI's energy estimate, now with one population. The key identity is $\log|u - v| = -\log 2 - \sum_{k\geq 1}\frac{2}{k}T_k(u)T_k(v)$ for $u \neq v$ in $[-1, 1]$. With the moments $M_k = \sum_l T_k(u_l)$ the pair sum is $-\binom{s}{2}\log 2 - \sum_k \frac{1}{k}(M_k^2 - \sum_l T_k(u_l)^2)$. For any sequence a and any $\bar{s} > 0$ one has $-M_k^2 \leq -2\bar{s}a_k M_k + \bar{s}^2 a_k^2$, which completes the square against a.

Lemma 7.5. *Let $\bar{s} \geq s$ and let a be a summable sequence. With $\mathsf{T}(a, u) = \sum_k a_k T_k(u)/k$ and $\|a\|^2 = \sum_k a_k^2/k$,*

$$\log Z_s(f) \leq -s^2\log 2 + 2\bar{s}^2\|a\|^2 + s\sup_{[-1,1]}\big(\log f - 4\bar{s}\,\mathsf{T}(a, \cdot)\big) + o(\bar{s}^2).$$

The diagonal terms $\sum_l T_k(u_l)^2/k$ diverge when summed over k. They are regularized by a damping factor $(1-\varepsilon_N)^k$ as in [28, Prop. 6.1], which changes the bound by $o(\bar{s}^2)$. The completed squares $(M_k - \bar{s}a_k)^2/k$ are dropped in this lemma. They are kept in §7.9.

The denominator is bounded from below by Jensen's inequality. Let m be a probability measure on $[0, \pi]$ of the form $\mathsf{m} = (1 + 2\sum_{k\leq K}\gamma_k\cos k\theta)\frac{d\theta}{\pi}$, with a positive density. In the variable $u = \cos\theta$ it is a polynomial multiple of the arcsine law.

Lemma 7.6. *For every bounded measurable Q,*

$$\log Z_s(e^Q) \geq s(s-1)\Big(-\log 2 - \sum_k \frac{2\gamma_k^2}{k}\Big) + s\int Q(\cos\theta)\,d\mathsf{m}(\theta) + s\,\mathrm{Ent}(\mathsf{m}),$$

where $\mathrm{Ent}(\mathsf{m})$ *is the entropy of the density of* m *with respect to* $d\theta$.

Proof. Apply Jensen's inequality to the product density $\prod_l \frac{d\mathsf{m}}{d\theta}(\theta_l)$. The logarithmic energy of m in the variable u is $-\log 2 - \sum_k 2\gamma_k^2/k$, by the expansion of $\log|u-v|$ above. □

The fields of the nodes integrate against m in closed form. For $|x|<1$ one has $\int l(x,\cos\theta)\,d\mathsf{m} = -2\sum_k (-x)^k\gamma_k/k$, a polynomial in x. The arcsine part contributes 0, because the logarithmic potential of the arcsine law is constant on $[-1,1]$.

7.6. **Positive nodes near 1.** A positive node $x_p = 1-\epsilon$ puts into Π_+ a factor $(1+x_p^2+2x_pu)^{-1}$ of size ϵ^{-2} near $u=-1$. Lemma 7.5 applies the supremum of the field to all s particles. So a single such node would change the bound by about $N\log(1/\epsilon)$, while the true change of $\det \mathsf{L}_{--}$ is bounded. The following reduction removes such nodes from the log gas.

Theorem 7.7. *Let f be positive and nonincreasing on $[-1,1]$, and let $w = w_r\prod_{p\in P}k_p$ with $w_r>0$, $k_p(u) = (1+2y_pu+y_p^2)^{-1}$ and $y_p\in(0,1)$. If $|P| = 2m'$ or $|P| = 2m'-1$, then*

$$\log\frac{Z_s(fw)}{Z_s(w)} \le m'\log f(-1) + \log\frac{Z_{s-m'}(fw_r)}{Z_{s-m'}(w_r)}.$$

Sketch. The full proof is formalized. We sketch it for orientation. For an odd number of factors add one more factor, at a point to be chosen. Write both integrals as Hankel determinants. Moving all y_p to a common value changes the ratio monotonically. This uses a one-point comparison for the weight $(1+2yu+y^2)^{-1}$, whose Chebyshev moments form the geometric sequence $(-y)^k$. For a common value y the $2m'$ factors behave like m' particles pinned near $u=-1$. Fischer's inequality bounds the numerator by $f(-1)^{m'}$ times the determinant of the remaining block. As $y\to 1$ the Schur complement of the pinned block in the denominator converges to the Hankel matrix of the remaining $s-m'$ particles. □

We apply the theorem to the positive nodes with $1-x_p<10^{-3}$. Since $\Pi_-(-1) = \prod_b(1-x_b)^2 \le 4^{k_-}$, each pair of such nodes changes the bound by a factor $e^{O(N)}$. In OpenAI's energy each such node contributes $-\beta_h n\log 10^3$, which is much more. Hence a near-extremal configuration has only $O(N)$ such nodes, and their total effect is $e^{O(N^2)}$ with a constant smaller than the margin of the final inequality (§7.10).

7.7. **Concavity: an inequality for signed logarithmic energy.** We now combine OpenAI's energy with (7.4), Theorem 7.4 and Lemmas 7.5 and 7.6. The total is a functional of the node measure σ, split into its negative part σ_- and its positive part σ_+, and of the column measure. Its quadratic part is not obviously negative. Theorem 7.4 contributes the term $-\iint\log(1-xy)\,d\sigma_-(x)d\sigma_+(y)$, and Lemma 7.5 applied with a trial sequence

that depends on σ contributes a convex term. Write

$$I(\zeta) = -\iint \log|x-y|\, d\zeta(x)d\zeta(y), \qquad L(\zeta) = -\iint \log(1-xy)\, d\zeta(x)d\zeta(y)$$

for real measures on $[-1,1]$ and on $[0,1]$ respectively. The second is a sum of squares of moments, $L(\zeta) = \sum_k \frac{1}{k}\big(\int x^k d\zeta\big)^2$. The quadratic part of the combined functional is $-\frac{1}{2}I - \frac{1}{2}\sum_k (B_k - 2C_k)^2/k + \frac{1}{2}L(\sigma_+)$, where B_k and C_k are the power moments of the nodes and the columns. The term $\frac{1}{2}L(\sigma_+)$ is absorbed by the following theorem.

Theorem 7.8. *Let $\zeta = \zeta_- + \zeta_+$ be a real measure of finite logarithmic energy and total mass 0, with ζ_- on $[-1,0]$ and ζ_+ on $[0,1]$. Let R be the reflection $x \mapsto -x$. Then*

$$I(\zeta) \geq L(\zeta_+) + L(\mathsf{R}\zeta_-) \geq L(\zeta_+).$$

Proof. Put $\zeta_1 = \zeta_+$ and $\zeta_2 = \mathsf{R}\zeta_-$, both on $[0,1]$, and let $K_1(a,b) = -\log|a-b|$ and $K_2(a,b) = -\log(a+b)$. Since $|x-y| = a+b$ for points of opposite signs, $I(\zeta) = K_1(\zeta_1) + K_1(\zeta_2) + 2K_2(\zeta_1,\zeta_2)$, which equals $\frac{1}{2}(K_1+K_2)(\zeta_1+\zeta_2) + \frac{1}{2}(K_1-K_2)(\zeta_1-\zeta_2)$.

First, $\zeta_1+\zeta_2$ has mass 0. One has $K_1 = G_{\mathbb{D}} + L$, where $G_{\mathbb{D}}(a,b) = \log|(1-ab)/(a-b)|$ is the Green function of the unit disc, whose energy is nonnegative. By Frullani's integral $-\log(a+b) = \int_0^\infty (e^{-(a+b)t} - e^{-t})\frac{dt}{t}$, so $K_2(\zeta') = \int_0^\infty \big(\int e^{-at} d\zeta'(a)\big)^2 \frac{dt}{t} \geq 0$ for ζ' of mass 0. Hence $(K_1+K_2)(\zeta_1+\zeta_2) \geq L(\zeta_1+\zeta_2)$.

Second, $K_1 - K_2 = \log\frac{a+b}{|a-b|}$ satisfies $(K_1-K_2)(\zeta') \geq L(\zeta')$ for every real ζ' on $[0,1]$. Indeed, with $a = e^{-x}$ and $b = e^{-y}$ one has $\log\frac{a+b}{|a-b|} = \log\coth\frac{|x-y|}{2} = \sum_{j\ \mathrm{odd}} \frac{2}{j} e^{-j|x-y|}$. For $x,y \geq 0$, $e^{-j|x-y|} = e^{-j(x+y)} + 2j\int_0^{\min(x,y)} e^{-j(x+y-2v)}dv$. Put $\Phi_j(v) = \int_{x\geq v} e^{-j(x-v)}d\zeta'$. Then $(K_1-K_2)(\zeta') = \sum_{j\ \mathrm{odd}}\big(\frac{2}{j}\Phi_j(0)^2 + 4\int_0^\infty \Phi_j(v)^2 dv\big)$. In $L(\zeta') = \sum_k \frac{1}{k}\big(\int e^{-kx}d\zeta'\big)^2$ the terms with odd k are $\Phi_k(0)^2/k$. For even $k = j+1$, integration by parts gives $\int e^{-(j+1)x}d\zeta' = \Phi_j(0) - \int_0^\infty e^{-(j+1)v}\Phi_j(v)\,dv$. The Cauchy–Schwarz inequality with weights $1+\frac{1}{j}$ and $1+j$ gives $\big(\int e^{-(j+1)x}d\zeta'\big)^2 \leq (1+\frac{1}{j})\Phi_j(0)^2 + \frac{1}{2}\int \Phi_j^2$. Summing, $L(\zeta') \leq (K_1-K_2)(\zeta')$.

Together, $I(\zeta) \geq \frac{1}{2}L(\zeta_1+\zeta_2) + \frac{1}{2}L(\zeta_1-\zeta_2) = L(\zeta_1) + L(\zeta_2)$. □

By Theorem 7.8 the quadratic part is at most $-\frac{1}{4}I - \frac{1}{2}\sum_k (B_k - 2C_k)^2/k \leq 0$ on differences of admissible measures. So the combined functional is concave. Its maximum is bounded by the value at a trial measure plus the supremum of a linear functional, that is, by a constant plus the suprema of fields of one variable. The formal proof uses a discrete form of Theorem 7.8 for weighted point configurations against a reference measure with bounded density. The prescribed rows of the next subsection change the masses of the two kinds of nodes. To compensate, we add further points distributed by the arcsine law, whose potential is constant on $[-1,1]$.

7.8. **Prescribed rows.** With m prescribed columns the row determinant is evaluated at the free points and at the prescribed points. The prescribed points enter $\det \mathsf{L}$ as nodes but are not integrated over. Lemma 7.3 must then be modified. Its constant $\frac{3}{2}$ need not hold when the sum runs only over the free positive nodes while the Blaschke product and the Gram matrix contain all nodes. Already two positive nodes can require a larger constant. We therefore keep the residues at the prescribed positive nodes as a separate term

$$\mathsf{W} = \sum_{a \text{ prescribed},\, x_a>0} \frac{x_a^{\mathsf{e}}}{(1-x_a^2)|B'(x_a)|},$$

and bound each of its terms by the reproducing kernel estimate $(c_a B'(x_a))^2(1-x_a^2) \le c^{\mathsf{T}}\mathcal{K}c$.

Instead of absorbing W into τ, we expand. Write $\det(F(1)+\operatorname{diag}(\cdot))$ as a sum of principal minors over subsets $\mathcal{S}$ of the prescribed positive nodes. This gives

$$|\det \mathsf{L}| \le 11^{k_+} \sum_{\mathcal{S}} \Big(\frac{6}{11}\Big)^{|\mathcal{S}|} \det \mathsf{L}_{--}^{(\mathcal{S})}\, \mathsf{W}_{\mathcal{S}}, \tag{7.5}$$

where in the term for $\mathcal{S}$ the nodes of $\mathcal{S}$ interact with the other nodes through $\log(1-xy)$ instead of $\log|x-y|$. The energy then has three prescribed populations: prescribed negative nodes, prescribed positive nodes, and the nodes of $\mathcal{S}$ in (7.5). Their fields are bounded on their domains like the field F_1 of Section 6. The prescribed nodes are images of quadrature nodes, so the mass of their measure σ' near 1 is limited by the spacing of the quadrature nodes: $\sigma'([x,1]) \le C_1(1-x)$ for $x \ge x_c = 0.999$. For such a measure of mass α and a field V with $V \le M_1$ on $[0,x_c]$ and $V(x) \le c_0 + r\log\frac{1}{1-x^2}$ on $[x_c,1)$,

$$\int V\, d\sigma' \le \alpha M_1 + C_1 \int_0^{1-x_c} \big(c_0 + r\log(1/u) - M_1\big)_+\, du.$$

The integral has a closed form. This bounds the prescribed rows without a finite optimization over their positions.

7.9. **Near-extremal configurations.** The reduction above needs three hypotheses on the configuration. These are the condition of Lemma 7.3, a bound $O(N)$ on the number of positive nodes within 10^{-3} of 1, and a fixed range for the proportion of negative nodes. We do not need them everywhere. If OpenAI's bound at the rate U_0 holds for a configuration, there is nothing to prove. Otherwise the energy of the configuration is within $\Gamma_0 N^2$ of its maximum, and we show that such configurations satisfy the three hypotheses.

For such a configuration, keep the completed squares and the nonnegative differences between the fields and their suprema in the proof of OpenAI's estimate. Their sum is at most the deficit, a constant times $\Gamma_0 N^2$. This bounds the fraction of nodes on which a field is far below its supremum, by Chebyshev's inequality. It also bounds every linear statistic $\sum_k F_k(M_k - \bar{s}a_k)/k$ of the moments, by the Cauchy–Schwarz inequality. The condition of

Lemma 7.3 is a family of such linear statistics in t. It is checked on a grid of 24 values of t, with a polynomial lower bound between the grid points. The count of nodes near 1 follows from the field of the positive nodes near $x = 1$, which lies below its supremum by a fixed amount. The range for the proportion of negative nodes follows from polynomial minorants of the indicator functions of the two signs.

7.10. **The trial measures.** The concave majorant of §7.7 is bounded at a trial measure. This consists of a reference measure for the free nodes, with a piecewise polynomial density, trial sequences for the rows, the columns and the numerator of Theorem 7.4, and the measure m of Lemma 7.6. The bound reads

$$\begin{aligned}\text{constant} + \alpha \sup(\text{fields of the prescribed rows}) \\ + \sup(\text{fields of the free nodes}) \\ + \sup(\text{column field}) \le -U_0 + \tfrac{1}{2}\log 2 + c_2\alpha.\end{aligned}$$

The constant is the value at the trial measure. It consists of logarithmic energies and moments of the reference measure, the norms of the trial sequences and the entropy of m, all of which have closed forms.

The fraction α of prescribed columns ranges over $[0, \frac{1}{16}]$. We use trial measures at $\alpha = 0$, $\frac{1}{200}$ and $\frac{1}{16}$, and interpolate linearly between consecutive ones. Every field is affine in the interpolation parameter, so its supremum on an interval is at most the chord through its suprema at the ends. The constant term is a quadratic polynomial in the parameter, whose curvature is an explicit combination of the cross energies of the two end measures. It remains to bound, at each of the three trial measures, the suprema of six fields of one variable. Appendix C lists them and describes the verification. The smallest margin of the resulting inequality is 0.037 per unit of N^2.

This proves Proposition 6.2, and with §6.4 it proves Theorem 5.1.

8. Proof of Theorem 1.1

OpenAI's energy method gives the decay rate $D_0 = n_0^2(U - \frac{1}{2}\log 2) = 6422.8355\ldots$ for our parameters. Section 7 raises it to $D = 6431.1338\ldots$ With $\Lambda = 6392.7880\ldots$ the polynomial $Q_N = (d_0 N)!\,\mathcal{D}_N\Delta_N$ has integer coefficients, degree at most $53N$, and

$$|Q_N(x)| \le e^{-\eta N^2} \qquad (|x - G| \le e^{-\rho N}), \qquad \eta = 38.33924.$$

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

$$\rho = 296.586094, \quad \eta = 38.33924, \quad \omega = \tfrac{3}{5}, \quad \delta = 2.9 \cdot 10^{-7},$$
$$\lambda = \frac{410}{\rho} = 1.382397\ldots$$

(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 $p < p' \le 2p$.
(d) By Theorem 5.1, Proposition 3.4 and $\log(53N)! = o(N^2)$, for $|x-G| \le e^{-\rho N}$ and large N,

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

since $\eta = 38.33924 < D - \Lambda = 38.34580\ldots$ In the formal proof this is the inequality $\eta < D - \Lambda$ with $\int \kappa$ replaced by its lower bound 22.05325.

The numerical conditions hold as well. We have $\rho\lambda = 410$ and $0 < \eta < 2d = 106$. Next $\omega = \frac{3}{5} \le 1$ and $2\omega = 1.2 < \lambda$. Further $2(\omega - \delta(2\lambda - \omega)) = 1.199998\ldots > 1$. Finally $d + 4d\delta\lambda = 53.0000849\ldots < \eta\lambda = 53.0000857\ldots$ So Proposition 2.2 gives $|G - a/b| > b^{-\rho\lambda} = b^{-410}$ 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 $\rho = 296.586094$, the prescribed columns contribute $n_0 c_2 \approx 208.6$ through the energy, and the weight and the binomial factor of one column contribute $W_0 \approx 88.0$. Bounding these two factors separately, by their maxima, would add about 5 to ρ. 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 signed bound of Section 7 about 8. Without them η would be about 8. Since the exponent is inversely proportional to η, a further gain of 0.1 in η would lower it by about 1.

9.2. **The row determinant.**

Remark 9.1. The energy method bounds $|\det \mathsf{L}|$ by 6^n, which is $e^{O(N)}$ and invisible at the scale N^2. At the extremal configuration, numerical evaluation for small N suggests that $\log|\det \mathsf{L}|$ is about $-28N^2$. So most of its true size is not captured by our bound either. The bound (7.4) loses only $e^{O(N)}$, and Theorem 7.4 is an identity. The loss is in the log-gas bounds of §7.5, which use fixed trial measures, and in the requirement that the combined functional be bounded by its maximum over all configurations. A sharper treatment

of the denominator $Z_s(\Pi_+)$ would raise Γ_0. 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, and the proofs give the kernel vectors explicitly. The closed formulas, the counting inequality and the families derived from it are proved for all parameters. So are the other arguments of the paper, except the explicit subspaces of Appendix A and the trial measures, which belong to $(58, 9, 3, 5, 3)$. 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), 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, or larger explicit subspaces, would increase η without changing anything else.

10. Formalization

All results of this paper are formalized in Lean 4 [10] with Mathlib [37], on top of OpenAI's formalization of the irrationality of G [28]. The development is available at

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

Theorem 1.1 is the theorem

`CatalanIrrationalityMeasure.catalan_irrationalityMeasure_410`

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 $|G - a/b| > b^{-410}$ for all integers a and all sufficiently large b. The modules 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 repository also contains a configuration `comparator.json` for `lake comparator` [23], which checks that the solution file proves the challenge statement with these axioms only. The prime number theorem is taken from the PNT+ project [38]. 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.

The proof combines conventional formal proofs with numerical certificates. These are the inverses of the base matrices modulo 1009, the trial sequences and the covers of Appendices B and C, and the subspaces, kernel vectors and interval conditions 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 [19] 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 [28].

Acknowledgements. We thank the authors of Mathlib and of the PNT+ project, and OpenAI for making its formalization available.

Appendix A. Explicit subspaces for the remaining intervals

On 34 of the 102 intervals of Theorem 4.10 the corank bound comes from the counting inequality with an explicit rational subspace. These intervals lie in $(3N, 16.2N)$. The subspaces were found numerically, as optimal constant subspaces for the bound below on each interval. They are dense rational subspaces with no evident closed form, so they are listed in the formalization. Only the integer identities and the interval conditions described below enter the proof.

Theorem A.1. *Let $v_1, \dots, v_t \in \mathbb{Z}^s$ span a subspace $V \subset \mathbb{Q}^s$ of dimension t. For j in a finite set $\mathcal{J}$ let B_j be an integer matrix with s columns and $r_j \le t$ an integer such that there are $t - r_j$ linearly independent vectors $k \in \mathbb{Z}^t$ with $B_j(\sum_a k_a v_a) = 0$. Let $\mathcal{E} \neq 0$ be the product of a nonzero $t \times t$ minor of $(v_1, \dots, v_t)$ and, for each j, of a nonzero maximal minor of these kernel vectors. Let $p \nmid \mathcal{E}$ be an odd prime with $H < p^2$, write $q = ap + r$ with $0 \le r < p$, let $V_p = \mathbb{F}_p^t$ be the space of coordinates of V modulo p, and let m and w_j $(j \in \mathcal{J})$ be nonnegative integers. Suppose that there are linear maps $\tau_{j,k} \colon V_p^m \to V_p$ $(k < w_j)$ and $\iota \colon V_p^m \to \mathbb{F}_p^n$ such that, on the common kernel of the maps $B_j \circ \tau_{j,k}$, the map ι is injective and satisfies the hypothesis of Theorem 4.1. Then*

$$\nu_N(p) \ge m\,t - \sum_{j \in \mathcal{J}} w_j\, r_j.$$

Proof. Since $p \nmid \mathcal{E}$, the reductions of $v_1, \dots, v_t$ are linearly independent, so V_p is the reduction of V in the coordinates of this basis. The kernel vectors of B_j satisfy an integer identity, so their reductions lie in the kernel of $B_j|_{V_p}$. They remain independent because their minor is a unit modulo p. Hence $\operatorname{rank}(B_j|_{V_p}) \le r_j$. The map $\Phi = (B_j \circ \tau_{j,k})_{j,k}$ therefore has rank at most $\sum_j w_j r_j$, and its kernel has dimension at least $mt - \sum_j w_j r_j$. Theorem 4.1, applied to $\ker \Phi$ with no further conditions, gives the bound. □

In the applications m and the w_j are affine functions of N and p on an interval of primes, while V and the B_j are fixed. So the bound is linear on the interval. The row combinations are given by decompositions modulo p (Definition 4.9) whose data x_d range over V_p. The maps $\tau_{j,k}$ are Taylor coefficients at $X = -1, 0, 1$ and coefficients at ∞. Admissibility of the data is required only on the common kernel of these conditions.

The integer identities. For each of the 34 intervals the formalization lists integer vectors $v_1, \dots, v_t$ spanning V, a set of t pivot columns with the adjugate of the corresponding minor, and for each map B_j its integer matrix, the bound r_j, $t - r_j$ kernel vectors in the coordinates of V, and again a pivot set with an adjugate. A single Boolean function of these data, evaluated by the Lean kernel, checks the shapes, the identities $\mathrm{adj} \cdot v_{\mathrm{piv}} = \delta \cdot I$ and the products $B_j(\sum_a k_a v_a) = 0$ in integers. Its soundness theorem gives the hypotheses of Theorem A.1 on V and the B_j, with $\mathcal{E}$ the product of the minors. No prime enters this check.

The interval conditions. The hypotheses on the conditions and on ι depend on N and p. An interval is described by finitely many linear inequalities in N and $(p-1)/2$, and on each of finitely many pieces of the interval the numbers m and w_j are affine in the same variables. On each piece a verified procedure checks four things. The orders of the Taylor conditions are nonnegative and fit the interval. The prime exceeds 59. The map ι is injective on the common kernel and lands in the admissible data. The values at the moment columns are given there by a polynomial in the column of degree less than $q \bmod p$, so that Theorem 4.1 applies. Each required inequality is derived from the inequalities of the interval by nonnegative combinations and integer rounding, that is, by Chvátal–Gomory cuts. The soundness theorem of the procedure gives the hypotheses of Theorem A.1 at every odd prime $p > 59$ of the interval with $p \nmid \mathcal{E}$. For large N all primes of the interval are of this kind. For 33 of the 34 intervals the hypotheses are established in this way, and for the remaining one by a direct proof of the same statement. Since the intervals lie in $(3N, 16.2N)$, where $p > q$, we have $a = 0$ in Theorem 4.1.

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 bound for W_0.

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 (6.1), 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 $-\frac{1}{2}\,\mathrm{Re}(r_j r_l \log(1 - z_j z_l) + r_j \bar{r}_l \log(1 - z_j \bar{z}_l))$ to $\|u\|^2$.

Covers by cells. We must bound three functions of one variable from above on their whole domains: the field X on $[-1, 1]$, the field Y on $[0, 1]$, and the field F_1 on $g_* \le |x| \le 1$. With $x = \cos\theta$, each is a constant plus a cosine series with 160 terms plus finitely many terms $\mathrm{Re}\, w \log(1 - a e^{im\theta})$.

We cover each domain by cells $[x_c - h', x_c + h']$ whose endpoints are multiples of 2^{-40}. On each cell we bound the function by its Taylor expansion at x_c. The cells are grouped in blocks. For each block the derivatives of the polynomial part are computed exactly at the centre of the block up to order 70, and the remainder beyond order 70 is bounded by a single constant for the block, multiplied by $h'^{70}/70!$. The Taylor expansion at the centre of a cell is obtained from that of the block by an exact shift. The logarithmic terms are expanded to order 5, with remainders bounded from the exact minimum of $|1 - ae^{im\theta}|$ on the cell. A few cells at $x = 0$ and $x = \pm 1$ treat the singular terms separately. The covers of X, Y and F_1 have 47, 38 and 23 blocks and 373, 302 and 176 cells. Each cell is checked by exact rational and integer arithmetic.

The bounds. The covers give

$$\sup X \le -0.96450499, \quad \sup Y \le -1.57499400, \quad \sup_{g_* \le |x| \le 1} F_1 \le -0.0631062.$$

So $M_X + M_Y = -2.53949899$, and the norms satisfy $2\|u\|^2 + \frac{1}{2}\|v\|^2 \le 0.7737$. The first coefficient satisfies $u_1 \ge 9/50$, which is used near $x = 0$. For the points of the second kind a one-variable estimate gives $\sup_{0<y<1} F_2(y) \le -77/100$ for $u_1 \ge 9/50$. 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 $n_1 = 0.67761$. With $a_0 = 41/58$, $a_1 = 76/29$ and $U = 2.255858836$ these give the constants stated after Lemma 6.1. At the rate U_0 the inequality of the constrained case still holds with the slope c_2, and that of the case with doubled interaction holds for $\alpha \ge \frac{1}{16}$ with the slope c_2.

The bound for W_0. For a fixed radius the function $y \mapsto \max_{\Gamma_r} \Phi_y$ is a maximum of functions affine in y, hence convex. The entropy h_2 is concave. On a cell $[y_1, y_2] \subset [\frac{1}{2}, 1]$ we replace h_2 by its tangent at the midpoint. The sum is then convex, and its supremum on the cell is attained at an endpoint. On $[0, \frac{1}{2}]$ the term $\tilde{h}_2 = \log 2$ is constant. We cover $[0, 1]$ by 14 cells, each with a rational radius r. At each endpoint the maximum of Φ_y on the circle and on the segments is bounded by a cover of the contour by small arcs and intervals, in exact arithmetic. This proves

$$\sup_{0 \le y \le 1} \big(w(y) + 58\tilde{h}_2(y)\big) \le W_0 = 87.9663.$$

The supremum is attained near $y = 0.62$, with radius about 0.6.

Appendix C. The trial measures of Section 7

A trial measure for the concave majorant of §7.7 consists of a reference measure for the free nodes with a piecewise polynomial density, trial sequences for the rows, the columns and the numerator of Theorem 7.4, and the measure m of Lemma 7.6. The trial sequences have 120 Chebyshev coefficients and

geometric tails with complex poles, as in (6.1). We use three trial measures, at $\alpha = 0$, $\frac{1}{200}$ and $\frac{1}{16}$. The constant of Theorem 7.7 interpolates between them in the same way as the fields.

At each of the three trial measures we bound the suprema of six fields: the fields of the negative and the positive free nodes, the column field, and the fields of the prescribed negative nodes, the prescribed positive nodes and the nodes of $\mathcal{S}$ in (7.5). Each field is a polynomial plus finitely many terms $P(x)\log|x-c|$ and $\operatorname{Re}\varrho\log(x-z)$ with complex z and ϱ, plus a constant.

The suprema are verified on covers by cells, as in Appendix B. The endpoints of the domains are treated by Laurent expansions and logarithmic end intervals. The polynomial part of a field has degree about 120. On a block of cells we shift its Taylor expansion exactly to a fixed order and bound the remaining coefficients by one table shared by the block, which keeps the integers small. The smallest margin of the inequality of §7.10 is 0.037 per unit of N^2.

References

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

[2] Dmitri I. Arkhipov, *Catalan:* $\mu(G) \le 528.2327$, GitHub repository, https://github.com/darkhipo/catalan-constant-gen5; verifier submission https://github.com/bigospeedrun/verifier/issues/95, November 1, 2026.

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

[4] 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.

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

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

[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 2026, doi:10.5281/zenodo.23532239; 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, Kumar Murty, and Yusuke Nemoto, *Mixed motives and linear forms in the Catalan constant*, 2025, Preprint, October 2025 (v3: 29 August 2026), arXiv:2510.20648.

[13] Aabir Fauzan, $\zeta(5)$ *is irrational*, Preprint v1, September 17, 2026, Zenodo, doi:10.5281/zenodo.22826419, 2026.

[14] Peter J. Forrester, *Meet Andréief, Bordeaux 1886, and Andreev, Kharkov 1882–83*, 2018, arXiv:1806.10411.
[15] David Niedbala Giraudin, *A new upper bound for the irrationality exponent of* $\zeta(2)$, 2026, Preprint, 2 October 2026, arXiv:2610.02912.
[16] ———, *A new upper bound for the irrationality exponent of* $\zeta(3)$, 2026, Preprint, 22 September 2026, arXiv:2609.26980.
[17] Masayoshi Hata, *Rational approximations to* π *and some other numbers*, Acta Arith. **63** (1993), no. 4, 335–349.
[18] R. K. 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.
[19] Fredrik Johansson, *Arb: efficient arbitrary-precision midpoint-radius interval arithmetic*, IEEE Trans. Comput. **66** (2017), no. 8, 1281–1292, arXiv:1611.02831.
[20] 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.
[21] Christian Krattenthaler and Wadim Zudilin, *Hypergeometry inspired by irrationality questions*, Kyushu J. Math. **73** (2019), no. 1, 189–203, arXiv:1802.08856.
[22] Li Lai, *The linear independence of* 1, $\zeta(3)$, *and* $\zeta(4)$, 2026, Preprint, 5 October 2026 (v2: 15 October 2026), arXiv:2610.06838.
[23] Lean FRO, *Comparator: checking a Lean proof against a trusted statement*, `https://github.com/leanprover/comparator`, 2026, Software; accessed October 2026.
[24] 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`.
[25] Sean Litsey, *Irrationality measures for* $L(2,\chi_{-3})$ *and other differences of trigamma values*, Hexagon preprint hexagon:2610.00401, `https://hexagonmath.org/2610.00401`, 2026.
[26] Jingwen Liu, Kai Jiang, and Pingwen Zhang, *Irrationality exponents of logarithms of positive rational numbers*, 2026, Preprint, 7 October 2026, revised 8 October 2026 (v2), arXiv:2610.10192.
[27] Raffaele Marcovecchio, *The Rhin–Viola method for* $\log 2$, Acta Arith. **139** (2009), no. 2, 147–184.
[28] 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.
[29] ———, *The irrationality exponent of* π *is 2*, Preprint dated September 24, 2026, released October 6, 2026, `https://github.com/openai/math/tree/main/preprints/The-irrationality-exponent-of-pi-is-2-September-24-2026`, 2026.
[30] 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.
[31] Georges Rhin and Carlo Viola, *On a permutation group related to* $\zeta(2)$, Acta Arith. **77** (1996), no. 1, 23–56.
[32] ———, *The group structure for* $\zeta(3)$, Acta Arith. **97** (2001), no. 3, 269–293.
[33] Tanguy Rivoal and Wadim Zudilin, *Diophantine properties of numbers related to Catalan's constant*, Math. Ann. **326** (2003), no. 4, 705–721.
[34] Edward B. Saff and Vilmos Totik, *Logarithmic potentials with external fields*, Grundlehren der mathematischen Wissenschaften, vol. 316, Springer, Berlin, 1997.
[35] V. Kh. Salikhov, *On the irrationality measure of the number* π, Russian Math. Surveys **63** (2008), no. 3, 570–572.
[36] Zhi-Wei Sun, *Catalan's constant is irrational*, 2026, arXiv:2609.04176.

[37] 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.

[38] 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`.

[39] Dane Wachs, *A note on a recent claimed proof of the irrationality of Catalan's constant*, 2026, arXiv:2609.22339.

[40] Ryan Xu, *Improved* $\kappa = 2$ *trial:* $\mu(G) \leq 727.03$, 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.

[41] 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.

[42] 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.

[43] ———, *An Apéry-like difference equation for Catalan's constant*, Electron. J. Combin. **10** (2003), Research Paper 14, arXiv:math/0201024.

[44] ———, *Two hypergeometric tales and a new irrationality measure of* $\zeta(2)$, Ann. Math. Québec **38** (2014), no. 1, 101–117, arXiv:1310.1526.

[45] ———, *A determinantal approach to irrationality*, Constr. Approx. **45** (2017), no. 2, 301–310, arXiv:1507.05697.

[46] ———, *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