

AN IRRATIONALITY MEASURE FOR CATALAN'S CONSTANT

SEAN LITSEY

ABSTRACT. Building on OpenAI's proof that Catalan's constant $G = \sum_{j \geq 0} (-1)^j / (2j+1)^2$ is irrational, we prove $|G - a/b| > b^{-1.281}$ for every integer a and every sufficiently large integer b . Thus $\mu(G) \leq 1.281$; replacing the prime number theorem in the final approximation criterion by Chebyshev's estimates gives 3.662. We change the determinant parameters, obtaining a $20N$ matrix and a polynomial of degree at most $18N$. Uniform physical prime-column pools and sparse residue relations sharpen its common denominator. Four exact energy certificates and a separate treatment of free and fixed populations give a uniform mixed-integrand bound. Restricted quadrature and joint atom weights then control the determinant polynomial on $|x - G| \leq e^{-120N}$. Prime-index nonvanishing when $p^2 \nmid b$ completes the measure criterion. We retain the earlier $48N$ construction and its 1.981 bound to explain the methods. Every ingredient of the new global theorem is formalized in Lean 4 with only the standard axioms and no unproved steps.

1. INTRODUCTION

OpenAI [7] recently proved that Catalan's constant

$$G = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)^2} = 0.91596\,55941\dots$$

is irrational, and formalized the proof in Lean. (An earlier claimed proof [9] was found to contain a gap [10].) The argument is by contradiction: under the hypothesis $G \in \mathbb{Q}$ an explicit determinant Δ_N of size $n = 48N$ is a nonzero rational number whose denominator is controlled prime by prime, so that $\log |\Delta_N| \geq -(A + o(1))n^2$; on the other hand a real-variable estimate gives $\log |\Delta_N| \leq -(U - o(1))n^2$ with $U > A$. The margin is small: $U - A \approx 1.3 \cdot 10^{-4}$. As OpenAI remarks, the argument proves qualitative irrationality; an irrationality measure requires additional control of rational approximations.

Recall that the *irrationality exponent* $\mu(\xi)$ of a real number ξ is the infimum of the μ such that $|\xi - a/b| > b^{-\mu}$ for all integers a and all sufficiently large b .

Date: October 13, 2026.

2020 Mathematics Subject Classification. Primary 11J82; Secondary 11J72, 68V20.

Key words and phrases. Catalan's constant, irrationality measure, irrationality exponent, determinants, Andréief identity, formal verification, Lean.

Theorem 1.1. *For every integer a and every sufficiently large integer b ,*

$$\left| G - \frac{a}{b} \right| > b^{-1281}.$$

In particular $\mu(G) \leq 1281$. If in the last step the prime number theorem is replaced by Chebyshev's bounds, the same argument gives $|G - a/b| > b^{-3662}$.

Theorem 1.1, together with every result in this paper, has been verified in the Lean 4 proof assistant on top of OpenAI's formalization, with no unproved steps and only the standard axioms; see Section 8. The formalization is available at <https://github.com/slitsey/catalan-irrationality-measure>. The use of AI in this work is disclosed at the end of the paper.

1.1. Previous results. The first finite irrationality measure for G was obtained by Jha [4], who proved $\mu(G) \leq 10^8$, also with a Lean formalization built on OpenAI's. The present work was carried out independently; we found Jha's preprint in a literature search after our proof was complete. Both arguments have the shape that OpenAI's proof suggests: one replaces the determinant by a polynomial in the constant, clears denominators, and needs a perturbation estimate near G . They differ in the perturbation estimate. Jha bounds the complementary minors of OpenAI's matrix in every dimension, by an energy estimate with node padding, with a codimension loss of the form $\exp(1000n(n-m))$. Here the perturbation of the matrix is written, by a quadrature formula, as a combination of point evaluations of OpenAI's row functions, and the new parameter family has perturbation exponent $\rho = 120$ (Section 7). The improvement also uses actual local arithmetic and certified energy fields for that family. For other recent work on rational approximations to G see Eskandari [2]; for earlier work on the arithmetic of G and related values see [8, 11]; for an effective approach to linear independence of related L -values see Calegari, Dimitrov and Tang [1].

1.2. Outline of the proof. Sections 2–6 develop the $48N$ determinant and the earlier 1981 bound. Section 7 changes the parameters and proves the current result. The steps for the new family are:

- (i) Its determinant polynomial has degree at most $18N$. Actual prime valuations supply a common denominator of normalized growth $8795/4800 + (21/200) \log 2$. Multiplication by $(18N)!$ gives an integer polynomial, nonzero at a/b for large prime indices p with $p^2 \nmid b$.
- (ii) Four certified endpoint trials and a uniform interpolation between free and fixed populations give physical real-place constant $2.256085 - \frac{1}{2} \log 2$. Joint quadrature weights give perturbation exponent 120; the arithmetic gap permits decay rate $\eta = 1.77$ after normalization.
- (iii) The prime-window criterion of Section 4 applied with $d = 18$, $\eta = 1.77$ and $\rho = 120$ gives exponents less than 1281 and 3662.

1.3. Notation and conventions. We use the notation of OpenAI's paper [7] and refer to it for the constructions; we restate only what we use. Citations of the form [7, Prop. 2.1] point to numbered statements, sections and equations of OpenAI's paper. We write $\log$ for the natural logarithm.

2. OPENAI'S CONSTRUCTION AND THE FACTS WE USE

For a positive integer N put [7, (2)]

$$\begin{aligned} n &= 48N, & b &= 7N, & q &= g = 4N, & h &= 2N, \\ L &= 59N, & C &= 63N, & H &= 65N, & A &= 19N, \end{aligned}$$

and $D_* = n - 1 - 2g = 40N - 1$. (The letter A for $19N$ is OpenAI's; the constant A below is different.) The row polynomials $P_r, D_r \in \mathbb{Z}[t]$ ($0 \leq r < n$) are those of [7, (3)]; we only need that

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

where T_d is the Chebyshev polynomial. In particular P_r depends only on $|r-g|$. Let $f(t) = \sqrt{1-t^2}$, $d\mu(t) = |t| dt/f(t)$ on $(-1, 1)$ (a measure of mass 2),

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

and let $\mathcal{A}_N$ be the real $n \times n$ matrix

$$(2.2) \quad \mathcal{A}_N(r, k) = \int_{-1}^1 \int_0^1 A_r(t) \frac{\psi_k(s)}{1-ts} ds d\mu(t) \quad (0 \leq r, k < n),$$

so that OpenAI's determinant is $\det \mathcal{A}_N$ [7, (7) and §5.1]. Put $c_l = 4^{-l} \binom{2l}{l}$ for integers $l \geq 0$ and $c_z = 0$ otherwise.

Proposition 2.1 (OpenAI).

(a) (Rational structure [7, Prop. 2.1 and (20)].) *There are rational matrices $\mathcal{A}_N^0, \mathcal{L}_N \in \mathbb{Q}^{n \times n}$ with $\mathcal{A}_N = \mathcal{A}_N^0 + G \mathcal{L}_N$ and*

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

(b) (Real place [7, Prop. 7.1].) *With $U_{\text{OAI}} = 2.290939875 - \frac{1}{2} \log 2$,*

$$\limsup_{N \rightarrow \infty} \frac{\log |\det \mathcal{A}_N|}{n^2} \leq -U_{\text{OAI}}.$$

(c) (Local estimates [7, Prop. 2.2, §3, Prop. 3.4].) *For every rational z OpenAI bounds $v_\ell(\det(\mathcal{A}_N^0 + z \mathcal{L}_N))$ from below at every prime ℓ , by quantities depending on ℓ , N and the denominator of z . The sum of these bounds, weighted by $\log \ell$, is at most $(A + o(1))n^2$, where*

$$A = \frac{8609}{4608} + \frac{505}{4608} \log 2.$$

- (d) (Nonvanishing [7, Prop. 4.1].) *There is an absolute constant $P_* \geq 260$ such that, for every prime $p > P_*$ and every $z \in \mathbb{Q}$ whose denominator is not divisible by p , $v_p(\det(\mathcal{A}_p^0 + z\mathcal{L}_p)) = -96p$; in particular this determinant is nonzero.*

Two remarks on (c) and (d). OpenAI states them under the hypothesis $G = z \in \mathbb{Q}$, but the proofs only use the arithmetic of the rational matrix $\mathcal{A}_N^0 + z\mathcal{L}_N$. In (d) OpenAI assumes $p > \max(P_*, \text{den } z)$, but the proof only uses $p \nmid \text{den } z$; this uniformity in z is essential for an irrationality measure, and is how we use the result.

We also use OpenAI's real-place machinery: the double integral, the two sheet estimates [7, Props. 5.2, 5.3], and the certified energy dual [7, Prop. 6.1, §7]. They are recalled where needed in Section 5.

OpenAI's irrationality proof is the inequality $U_{\text{OAI}} > A$, and $U_{\text{OAI}} - A = 1.3031 \dots \cdot 10^{-4}$. With a new certificate for one case of OpenAI's energy dual, Proposition 5.11 below shows that (b) holds with

$$U = 2.29180925 - \frac{1}{2} \log 2$$

in place of U_{OAI} . We use this constant throughout.

Definition 2.2. $\Delta_N(X) = \det(\mathcal{A}_N^0 + X\mathcal{L}_N) \in \mathbb{Q}[X]$. Thus $\Delta_N(G) = \det \mathcal{A}_N$, and $\Delta_N(z)$ for $z \in \mathbb{Q}$ is the rational determinant OpenAI studies under the hypothesis $G = z$. We write $\gamma_0 = U - A$.

Numerically $U = 1.9452356 \dots$, $A = 1.9442359 \dots$ and

$$(2.4) \quad 9.996945 \cdot 10^{-4} < \gamma_0 = 9.996948 \dots \cdot 10^{-4}.$$

3. ARITHMETIC OF $\Delta_N(X)$

Lemma 3.1 (Degree). $\deg \Delta_N \leq 44N$.

Proof. By (2.3) row r of $\mathcal{L}_N$ depends only on P_r , hence only on $|r - g|$ by (2.1). So rows r and $2g - r$ of $\mathcal{L}_N$ agree for $0 \leq r < g$. Subtracting row r from row $2g - r$ for each $r < g$ is a unipotent change of rows, which does not change the determinant, and it makes the g rows of index in $(g, 2g]$ of the new X -coefficient vanish. A determinant $\det(B + XL')$ in which L' has only $n - g$ nonzero rows has degree at most $n - g = 44N$ in X . $\square$

Proposition 3.2 (Common denominator). *There are positive integers D_N such that $D_N \Delta_N(i) \in \mathbb{Z}$ for every $i \in \mathbb{N}$, and for every $\varepsilon > 0$, $\log D_N \leq (A + \varepsilon)n^2$ for all large N .*

Proof. At an integer $z = i$ the denominator of z is 1, and OpenAI's local lower bounds for $v_\ell(\Delta_N(i))$ become independent of i : at odd primes they depend only on ℓ and N (and vanish for $\ell > H$), and at $\ell = 2$ they depend on z only through a quantity that, over the integers, is largest at $z = 0$. Let $e_\ell(N) \geq 0$ be the resulting bound, and put $D_N = \prod_{\ell \leq H} \ell^{e_\ell(N)}$. Then $v_\ell(D_N \Delta_N(i)) \geq 0$ for every prime ℓ , and summing the bounds $e_\ell(N) \log \ell$

exactly as in OpenAI's proof of its finite-place bound [7, Prop. 3.4] gives $\log D_N \leq (A + o(1))n^2$. $\square$

Proposition 3.3 (Sparse prime-row dependencies). *Let $N \geq 1$ and let p be a sufficiently large prime such that $58N \leq 3p$, $p \leq 21N$ and $65N < p^2$. Put $e = \min(4N, 63N - 3p)$. For any rational z such that pz has denominator coprime to p ,*

$$v_p(\Delta_N(z)) \geq -96N + e$$

when $\Delta_N(z) \neq 0$. The corresponding lower bound with the usual formal convention $v_p(0) = 0$ also holds at zero.

Proof. Work over $\mathbb{F}_p$ and put $W = (1 - X^2)^{(p-1)/2}$. For $1 \leq d \leq e$, let $\delta_k P = P_{g+kp+d} - P_{g+kp-d}$ and define $\delta_k D$ similarly, for $k = 0, 1, 2$. Chebyshev's identities in characteristic p give $T_p(Y) = Y^p$ and $U_{p-1}(Y) = (Y^2 - 1)^{(p-1)/2}$. With

$$Q_d = 2(1 - X)^h X^{C-p-1} U_{d-1}(X^{-1}), \quad Q_d^{\text{lo}} = X^{-p} Q_d,$$

both expressions are polynomials, and direct substitution gives

$$\begin{aligned} \delta_0 D &= X^p Q_d, & \delta_1 D &= Q_d, & \delta_2 D &= 2Q_d^{\text{lo}} - X^p Q_d, \\ XW\delta_0 P &= 0, & XW\delta_1 P &= (1 - X^{2p})Q_d, & XW\delta_2 P &= 2(1 - X^{2p})Q_d^{\text{lo}}. \end{aligned}$$

There is no parity sign: substituting $Y = X^{-1}$ in U_{p-1} cancels the sign that arises from reversal. The support of Q_d is contained in $[C - p - d, H - p - 1] \subseteq [2p, 3p)$, and $Q_d = X^p Q_d^{\text{lo}}$.

The leading residues of p^2 times a raw matrix entry are obtained by grouping the coefficients of $XW P_r$ and D_r by their base- p digit. At a raw column $b \leq j < L$, a nonzero coefficient $(Q_d)_{2p+j \bmod p}$ can contribute only when j/p (integer division) is 1 or 2. Indeed, $3p \geq H - b$ excludes quotient zero, and $d \leq g$ excludes quotient three. Write $k = \lfloor j/p \rfloor$ and let $M_{u,k}$ and $Z_{u,k}$ denote OpenAI's rational moment and zeta coefficients. The three row differences have the common coefficient $(Q_d)_{2p+j \bmod p}$ times, respectively,

$$\begin{aligned} A_k &= -\frac{3}{2}Z_{3,k}, \\ B_k &= M_{1,k} - M_{3,k} - \frac{3}{2}Z_{2,k}, \\ C_k &= 2(M_{0,k} - M_{2,k}) - 3Z_{1,k} + \frac{3}{2}Z_{3,k}. \end{aligned}$$

For $k = 1, 2$, the exact rational identity

$$C_k = \frac{292}{59}B_k - \frac{3281}{177}A_k$$

is checked directly from the moment formulas. Its reduction modulo p proves the same identity for the leading row residues. The column filter is an integer linear combination of raw columns in $[b, L)$, so the identity persists after filtering.

For each d , replace row $g + 2p + d$ by $\delta_2 - \frac{292}{59}\delta_1 + \frac{3281}{177}\delta_0$. Every source row is strictly below this target. The resulting row transformation is lower triangular with diagonal one and therefore preserves the determinant. There

are exactly e distinct target rows, all below $48N$. All entries of the scaled matrix have denominator coprime to p , and each target row now has residue zero. Dividing these e rows by p preserves that denominator property, so $v_p(p^{96N}\Delta_N(z)) \geq e$. This proves the claim. A fixed cutoff $p \geq 10^6$ suffices for the finite rational denominator checks used in the formal proof. $\square$

Proposition 3.4 (An earlier residue window). *Let $N \geq 1$ and let $p \geq 10^6$ be prime, with $18N \leq p$, $3p \leq 58N$ and $65N < p^2$. For any rational z whose product pz has denominator coprime to p ,*

$$v_p(\Delta_N(z)) \geq -96N + (3p - 54N).$$

The same formal convention at zero applies.

Proof. Use all $g = 4N$ modes from the preceding proof. Each Q_d is still supported in $[2p, 3p)$: $3p + g \leq C$ and $H \leq 4p$. The obstruction to the digit identity is now quotient zero, contributed by coefficients in $[2p + b, H - p)$. For $c \in \mathbb{F}_p^g$ put $Q_c = \sum_{d=1}^g c_d Q_d$ and let

$$B : \mathbb{F}_p^g \longrightarrow \mathbb{F}_p^r, \quad r = H - 3p - b = 58N - 3p,$$

record precisely these coefficients of Q_c . Rank-nullity gives $\dim \ker B \geq g - r = 3p - 54N$. When $c \in \ker B$, its combined polynomial has support below $2p + b$ and above $L - p$, so only column quotients one and two contribute. The same rational digit identity therefore makes

$$J(c) = \sum_{d=1}^g c_d \left(\delta_2 - \frac{292}{59} \delta_1 + \frac{3281}{177} \delta_0 \right)$$

a left null vector of the leading residue matrix. The map J is injective: its coefficient at the distinct target row $g + 2p + d$ is exactly c_d , and all source rows lie below $g + 2p$ since $g < p$. Consequently the residue rank is at most $n - (3p - 54N)$.

For completeness, the rank-to-valuation step does not require a particular basis of $\ker B$. If a rational matrix A has entries with denominator coprime to p , and its reduction has rank at most $n - e$, factor that reduction through $\mathbb{F}_p^{n-e}$ and lift the factors to rational matrices with integer entries. Write $A = UV + pE$, with E also having denominators coprime to p . Multilinearity of the determinant kills every term with more than $n - e$ columns from UV ; all remaining terms contain at least e factors p . Thus $v_p(\det A) \geq e$. Apply this to $A = p^2 M$ and $e = 3p - 54N$. $\square$

Proposition 3.5 (Refined common denominator). *There are positive integers D_N^* such that $D_N^* \Delta_N(i) \in \mathbb{Z}$ for every $i \in \mathbb{N}$ and, for every $\varepsilon > 0$ and all sufficiently large N ,*

$$\log D_N^* \leq (A_* + \varepsilon)n^2, \quad A_* = A - \frac{5}{1728}.$$

Proof. Take the minimum of the old denominator exponent and the new bounds from Propositions 3.3 and 3.4 wherever each applies. This defines

D_N^* and proves integrality at all integer points. For large N the old exponent is exactly $96N$ throughout these windows. The resulting continuous saving is

$$f(x) = 3(x - 18)_+ - 3(x - 58/3)_+ - 3(x - 59/3)_+ + 3(x - 21)_+,$$

where $y_+ = \max(y, 0)$. It rises from zero to four on $[18, 58/3]$, is constant on $[58/3, 59/3]$, then decreases to zero at 21. It is nonnegative, supported in $[18, 21]$, and $\int_1^{65} f(x) dx = 20/3$. The weighted prime number theorem gives

$$\frac{N}{n^2} \sum_{p \leq 65N} f(p/N) \log p \longrightarrow \frac{5}{1728}.$$

Subtract this saving from Proposition 3.2 to obtain the result. $\square$

Lemma 3.6 (Integer-valued polynomials). *If $f \in \mathbb{Q}[X]$ has degree at most d and $f(\mathbb{N}) \subseteq \mathbb{Z}$, then $d! f \in \mathbb{Z}[X]$.*

Proof. The d -th forward difference of f is $d!$ times its coefficient of X^d , and it is an integer combination of $f(0), \dots, f(d)$. Subtract that coefficient times the integer-valued polynomial $X(X-1)\cdots(X-d+1)$ and induct on d . $\square$

Proposition 3.7 (Square-denominator nonvanishing). *For every prime $p > P_*$ and rational z such that pz has denominator coprime to p , $\Delta_p(z) \neq 0$. In particular, $\Delta_p(a/b) \neq 0$ if $a \in \mathbb{Z}$, $b \geq 1$ and $p^2 \nmid b$.*

Proof. The scaled residue matrix in OpenAI's nonvanishing proof is obtained from the rational matrix by multiplying each entry by p^2 and permuting its rows and columns. Each filtered entry is affine in z , so its change from the entry at zero is $p^2 z \mathcal{L}_p(r, k)$. The central-binomial coefficients c_l have denominators that are powers of 2, and the filter coefficients are integers. Thus the denominator of $\mathcal{L}_p(r, k)$ is coprime to the odd prime p . If pz also has denominator coprime to p , the change $p(pz) \mathcal{L}_p(r, k)$ has denominator coprime to p and residue zero. The scaled residue matrix is therefore exactly the matrix at $z = 0$. OpenAI's proof shows that this matrix has nonzero determinant for $p > P_*$, proving $\Delta_p(z) \neq 0$.

If $p \nmid b$, then $p(a/b)$ has denominator coprime to p . Otherwise write $b = pc$; $p^2 \nmid b$ implies $p \nmid c$, and $p(a/b) = a/c$ again has denominator coprime to p . $\square$

Corollary 3.8. $Q_N = (44N)! D_N^* \Delta_N(X)$ lies in $\mathbb{Z}[X]$ and has degree at most $44N$. If $p > P_*$ is prime, $a \in \mathbb{Z}$, $b \geq 1$ and $p^2 \nmid b$, then $Q_p(a/b) \neq 0$ and hence $b^{44p} |Q_p(a/b)| \geq 1$.

Proof. The first claims follow from Lemmas 3.1 and 3.6 and Proposition 3.5. Proposition 3.7 gives $\Delta_p(a/b) \neq 0$. Finally $b^{44p} Q_p(a/b)$ is a nonzero integer. $\square$

4. THE IRRATIONALITY-MEASURE CRITERION

Proposition 4.1. *Let $\xi \in \mathbb{R}$, $d \in \mathbb{N}$ and $\eta, \rho > 0$. Let $Q_p \in \mathbb{Z}[X]$ be defined for primes p and suppose that, for all primes $p \geq P_0$:*

- (i) $\deg Q_p \leq dp$;
- (ii) $Q_p(a/b) \neq 0$ for all $a \in \mathbb{Z}$ and $b \geq 1$ with $p \nmid b$;
- (iii) $|Q_p(x)| \leq e^{-\eta p^2}$ whenever $|x - \xi| \leq e^{-\rho p}$.

Put $\kappa = \max(d/\eta, 4)$. Then $|\xi - a/b| > b^{-3\rho\kappa}$ for all $a \in \mathbb{Z}$ and all sufficiently large b . Assuming the prime number theorem, for every $\varepsilon > 0$ and every $c > 1$ the same holds with the exponent $(1 + \varepsilon)\rho \max(d/\eta, c/\varepsilon)$. If (ii) holds whenever $p^2 \nmid b$, this PNT exponent improves to $(1 + \varepsilon)\rho \max(d/\eta, c/(2\varepsilon))$.

Proof. Suppose $|\xi - a/b| \leq b^{-3\rho\kappa}$ with b large, and put $X = \kappa \log b$. If every prime in $(X, 3X]$ divided b , then $\vartheta(3X) - \vartheta(X) \leq \log b \leq X/4$, where ϑ is Chebyshev's function. But Chebyshev's bounds $\vartheta(x) \leq x \log 4$ and $\vartheta(x) \geq (x - 1) \log 2 - O(\sqrt{x} \log x)$ give $\vartheta(3X) - \vartheta(X) \geq X \log 2 - o(X)$. Hence for large b there is a prime $p \in (X, 3X]$ with $p \nmid b$, and $p \geq P_0$. Since $\rho p \leq 3\rho\kappa \log b$, we get $|a/b - \xi| \leq e^{-\rho p}$, so by (iii) $|Q_p(a/b)| \leq e^{-\eta p^2}$. By (i) and (ii), $b^{dp} Q_p(a/b)$ is a nonzero integer, so $e^{-\eta p^2} \geq b^{-dp}$, that is $p \leq (d/\eta) \log b \leq X$. This contradicts $p > X$.

For the second statement use the window $(X, (1 + \varepsilon)X]$ with $X = \max(d/\eta, c/\varepsilon) \log b$. Then $\log b \leq \varepsilon X/c$, while the prime number theorem gives $\vartheta((1 + \varepsilon)X) - \vartheta(X) = \varepsilon X + o(X)$, which exceeds $\varepsilon X/c$ for large X .

For the stronger nonvanishing condition, put $X = \max(d/\eta, c/(2\varepsilon)) \log b$. If every prime in $(X, (1 + \varepsilon)X]$ had its square dividing b , their distinct prime squares would have product dividing b . Taking logarithms gives $2(\vartheta((1 + \varepsilon)X) - \vartheta(X)) \leq \log b \leq 2\varepsilon X/c$. The PNT contradicts this for large X . There is therefore a prime with $p^2 \nmid b$ in the window, and the polynomial argument proceeds unchanged. $\square$

The factor 3 in the first version, or $1 + \varepsilon$ in the second, is the price of finding a prime $p \nmid b$ in a window of the right size.

5. THE PERTURBATION ESTIMATE

This section proves the main new ingredient. Recall $\mathbf{U} = 2.29180925 - \frac{1}{2} \log 2$.

Theorem 5.1 (Perturbation estimate). *Let $\rho = 366.3$. For every $\varepsilon > 0$ there is N_0 such that for all $N \geq N_0$ and all real x with $|x - G| \leq e^{-\rho N}$,*

$$|\Delta_N(x) - \Delta_N(G)| \leq \exp(n^2(-\mathbf{U} + \varepsilon)).$$

Let $\delta = x - G$, so that $\Delta_N(x) = \det(\mathcal{A}_N + \delta \mathcal{L}_N)$. The plan: write $\mathcal{L}_N = \Phi \beta$, where the columns of Φ are values of OpenAI's row functions A_r at fixed nodes (§5.1); expand the determinant column by column (§5.2); write each resulting mixed determinant as an integral (§5.3); and bound the integrand by OpenAI's method (§§5.4–5.6). Throughout, $J = H = 65N$.

5.1. A quadrature formula for $\mathcal{L}_N$. Let $x_a = \cos((2a+1)\pi/(2J))$ for $0 \leq a < J$ be the roots of the Chebyshev polynomial T_J , let $\tau_a = (x_a - 1)/2 \in (-1, 0)$, and let $\ell_a(t) = \prod_{c \neq a} (t - \tau_c)/(\tau_a - \tau_c)$ be the Lagrange basis.

Lemma 5.2. *For $0 \leq r, k < n$,*

$$\mathcal{L}_N(r, k) = \sum_{a < J} A_r(\tau_a) \beta_{a,k}, \quad \beta_{a,k} = \sum_{v=0}^q (-1)^v \binom{q}{v} \sum_{i < H} [t^i] \ell_a \cdot 4 c_{(i-b-k-v)/2},$$

and $|\beta_{a,k}| \leq B := 8J^2 K^N$, where $K = 95 \cdot 10^{46}$.

Proof. Since $\deg P_r < H = J$, Lagrange interpolation gives $[t^i] P_r = \sum_a P_r(\tau_a) [t^i] \ell_a$. Substitute this into (2.3) and use $A_r(\tau) = P_r(\tau)$ for $\tau \leq 0$.

For the bound, put $g_a(t) = \ell_a(t)(t-1)^q$. Since $[t^i] \ell_a = 0$ for $i \geq H$ and $c_z = 0$ unless $z \geq 0$, the substitution $i = b+k+v+e$ and $\sum_v (-1)^v \binom{q}{v} [t^{j+v}] \ell_a = [t^{j+q}] g_a$ give

$$\beta_{a,k} = 4 \sum_{0 \leq e < H} c_{e/2} [t^{b+k+e+q}] g_a.$$

Every coefficient of a polynomial is bounded by its maximum modulus on the unit circle (average $g(\omega^j) \omega^{-jm}$ over the M -th roots of unity ω^j , $M > \deg g$). Since $|c_z| \leq 1$, it suffices to show $|g_a(t)| \leq 2J K^N$ for $|t| = 1$.

We have $T_J = 2^{J-1} \prod_c (X - x_c)$, so $\prod_c (t - \tau_c) = T_J(2t+1)/(2^J 2^{J-1})$. For the denominator of ℓ_a , $\prod_{c \neq a} |\tau_a - \tau_c| = 2^{1-J} \prod_{c \neq a} |x_a - x_c|$ and $\prod_{c \neq a} (x_a - x_c) = 2^{1-J} T_J'(x_a) = 2^{1-J} J U_{J-1}(x_a)$. Since $U_{J-1}(\cos \theta) \sin \theta = \sin(J\theta)$ and $\sin(J\theta_a) = \pm 1$ at $\theta_a = (2a+1)\pi/(2J)$, we get $|U_{J-1}(x_a)| \geq 1$, so the denominator is at least $J 2^{2-2J}$. Finally, for $|t| = 1$, $|t - \tau_a| \geq 1 + \tau_a = (1 + x_a)/2 \geq \sin^2(\pi/(4J)) \geq 1/(4J^2)$. Hence

$$|\ell_a(t)| \leq \frac{|T_J(2t+1)|}{2^{2J-1}} \cdot \frac{2^{2J-2}}{J} \cdot 4J^2 = 2J |T_J(2t+1)|.$$

Writing $\zeta = \cosh w$, one has $T_J(\zeta) = \cosh(Jw)$, $|\zeta - 1| + |\zeta + 1| = 2 \cosh \operatorname{Re} w$ and $|\cosh(Jw)| \leq e^{J|\operatorname{Re} w|}$, so

$$|T_J(\zeta)| \leq (S + \sqrt{S^2 - 1})^J, \quad S = \frac{1}{2}(|\zeta - 1| + |\zeta + 1|).$$

For $\zeta = 2t + 1$ with $|t| = 1$, let $s = |t + 1| \in [0, 2]$. Then $|\zeta - 1| = 2$, $|\zeta + 1| = 2s$, $S = 1 + s$, and $|t - 1|^2 = 4 - s^2$ by the parallelogram law. With $J = 65N$ and $q = 4N$,

$$|g_a(t)| \leq 2J \left((1 + s + \sqrt{s^2 + 2s})^{65} (4 - s^2)^2 \right)^N \leq 2J K^N,$$

where the last inequality holds for all $s \in [0, 2]$. (The function $(1 + s + \sqrt{s^2 + 2s})^{65} (4 - s^2)^2$ is largest near $s = 1.917$, where its logarithm is $110.394\dots$, while $\log K = 110.4728\dots$. We check it by splitting $[0, 2]$ into 21 intervals; on each, the first factor is increasing and the second decreasing in s , so it suffices to compare rational upper bounds at the endpoints with K .) $\square$

Remark 5.3. The cruder bound $\sum_i |[t^i]\ell_a| \leq \prod_{c \neq a} (1 - \tau_c) / \prod_{c \neq a} |\tau_a - \tau_c| \leq (3 + 2\sqrt{2})^J$, together with $\sum_v \binom{q}{v} = 2^q$, gives $|\beta_{a,k}| \leq 2^{q+2}(3 + 2\sqrt{2})^J$, whose logarithm is $(4 \log 2 + 65 \log(3 + 2\sqrt{2}))N \approx 117.35N$ instead of $\log K \cdot N \approx 110.47N$. The unit circle captures the cancellation in $(t - 1)^q$ between $t = 1$, where $|T_J(2t + 1)|$ is largest, and $t = -1$, where it is smallest.

Let Φ be the $n \times J$ matrix $\Phi(r, a) = A_r(\tau_a)$. Then $\Delta_N(x) = \det(\mathcal{A}_N + \delta \Phi \beta)$.

5.2. Column expansion. For a map $f: \{0, \dots, n-1\} \rightarrow \{\emptyset\} \cup \{0, \dots, J-1\}$ let M_f be the matrix whose column k is column k of $\mathcal{A}_N$ if $f(k) = \emptyset$ and the column $\Phi(\cdot, f(k))$ otherwise, and let $m(f) = \#\{k : f(k) \neq \emptyset\}$.

Lemma 5.4. *If $|\beta_{a,k}| \leq B$ and $|\det M_f| \leq M c^{m(f)}$ for every f , with $M, c \geq 0$, then*

$$|\det(\mathcal{A}_N + \delta \Phi \beta) - \det \mathcal{A}_N| \leq M((1 + J|\delta|Bc)^n - 1).$$

Proof. Column k of $\mathcal{A}_N + \delta \Phi \beta$ is $\mathcal{A}_N(\cdot, k) + \sum_a \delta \beta_{a,k} \Phi(\cdot, a)$. Multilinearity gives $\det(\mathcal{A}_N + \delta \Phi \beta) = \sum_f \prod_{k: f(k) \neq \emptyset} \delta \beta_{f(k), k} \det M_f$. The term $f \equiv \emptyset$ is $\det \mathcal{A}_N$, and the others sum in absolute value to at most $M \sum_{f \neq \emptyset} (|\delta|Bc)^{m(f)} = M((1 + J|\delta|Bc)^n - 1)$. $\square$

5.3. A mixed Andréief identity. Fix f , and let $p = n - m(f)$. Permuting columns (which only changes the sign of the determinant), M_f has p integral columns $\mathcal{A}_N(\cdot, E_i)$, $0 \leq E_1 < \dots < E_p < n$, followed by $m = n - p$ point columns $(A_r(\tau'_j))_r$, where $\tau'_j = \tau_{f(k_j)} \in (-1, 0)$.

Lemma 5.5. *Let μ, ν be finite measures, F_r ($r < p + m$), G_i ($i < p$) and K functions with the required integrability, and B the $(p + m) \times (p + m)$ matrix with $B(r, i) = \iint F_r(t) K(t, s) G_i(s) d\mu(t) d\nu(s)$ for $i < p$ and $B(r, p + j) = F_r(\tau'_j)$. Then*

$$(p!)^2 \det B = \int \int \det[F_r((t \parallel \tau')_c)]_{r,c} \times \det[K(t_i, s_j)]_{i,j < p} \det[G_i(s_j)]_{j,i < p} d\mu^p(t) d\nu^p(s),$$

where $t \parallel \tau'$ is the concatenated list of length $p + m$. Consequently $|\det B|$ is at most $(p!)^{-2}$ times the total mass of $\mu^p \otimes \nu^p$ times the supremum of the absolute value of the integrand.

Proof. For $m = 0$ this is Andréief's identity applied twice [3], as used by OpenAI for [7, (62)]. In general expand $\det B$ by multilinearity in the first p columns, which integrates t and s variables column by column, and then antisymmetrize first in the t variables and then in the s variables; the last m columns are constant throughout. $\square$

We apply this with $F_r = A_r$, $G_i = \psi_{E_i}$, $K(t, s) = 1/(1 - ts)$, μ as in §2 and ν Lebesgue measure on $(0, 1)$; the integrability is OpenAI's [7, §5.1].

The mass of $\mu^p \otimes \nu^p$ is $2^p \leq 2^n$, so

$$(5.1) \quad |\det M_f| \leq 2^n \sup_{t,s} \left| \det[A_r((t \parallel \tau')_c)] \cdot \det\left[\frac{1}{1 - t_i s_j}\right] \cdot \det[\psi_{E_i}(s_j)] \right|,$$

the supremum over $t \in (-1, 1)^p$ and $s \in (0, 1)^p$ with distinct nonzero coordinates (a set of full measure). We have bounded the factor $(p!)^{-2}$ by 1.

5.4. A generalized Vandermonde bound. The column determinant is $\det[\psi_{E_i}(s_j)] = \prod_j s_j^b (1 - s_j)^q \cdot \det[s_j^{E_i}]$. Write $V(y) = \prod_{i < j} (y_j - y_i)$ and $\mathcal{V}(y) = |V(y)|$.

Lemma 5.6. *Let $s_0, \dots, s_{p-1} \in [0, 1]$ and let $E_1, \dots, E_p < n$ be distinct. Then*

$$|\det[s_j^{E_i}]| \leq \left(\prod_j (1 + s_j) \right)^{n-p} \mathcal{V}(s) \leq 2^{p(n-p)} \mathcal{V}(s).$$

Proof. Let $m = n - p$, let $k_1, \dots, k_m$ be the integers in $[0, n)$ that are not among the E_i , and let $P(X) = \prod_j (X - s_j)$, a monic polynomial of degree p . Consider two $n \times n$ matrices. The rows of Ω are the linear functionals $f \mapsto f(s_j)$ ($j < p$) and $f \mapsto [X^{k_l}]f$ ($l \leq m$), evaluated on the monomials X^a , $a < n$. The columns of Γ are the coefficient vectors of the polynomials X^b ($b < p$) and $X^i P$ ($i < m$), all of degree less than n .

- Ordering the columns of Ω as $E_1, \dots, E_p, k_1, \dots, k_m$ makes it block triangular, with diagonal blocks $[s_j^{E_i}]$ and the identity. So $|\det \Omega| = |\det[s_j^{E_i}]|$.
- In the natural order of the monomials, Γ is block triangular with diagonal blocks the identity and the matrix $([X^{p+l}]X^i P)_{l,i}$, which is unitriangular because P is monic of degree p . So $|\det \Gamma| = 1$.
- The entries of $\Omega\Gamma$ are the values of the functionals on the polynomials. Since $P(s_j) = 0$, the block of evaluations on the $X^i P$ vanishes, so $\Omega\Gamma$ is block triangular with diagonal blocks the Vandermonde matrix $[s_j^b]_{j,b < p}$ and $Q = ([X^{k_l}]X^i P)_{l,i}$.

Hence $|\det[s_j^{E_i}]| = \mathcal{V}(s) |\det Q|$. A determinant is bounded by the product of the ℓ^1 norms of its columns, and each column of Q consists of distinct coefficients of $X^i P$. Its ℓ^1 norm is therefore at most the sum of the absolute values of the coefficients of P , which is $\prod_j (1 + s_j) \leq 2^p$. $\square$

Remark 5.7. The bound is close to optimal. The ratio $\det[s_j^{E_i}]/V(s)$ is a Schur polynomial with nonnegative coefficients, and its value at $s = (1, \dots, 1)$, for a single missing exponent k , is $\binom{n-1}{k}$, of size 2^{n-1} up to a polynomial factor.

5.5. The integrand. Use OpenAI's coordinate $t = 2x/(1+x^2)$, $x \in (-1, 1)$ [7, (64)], write x_i for the coordinates of the free nodes t_i and $y_j \in (-1, 0)$ for those of the fixed nodes τ'_j , and $z = x \parallel y \in (-1, 1)^n$. For a list z of n distinct nonzero numbers in $(-1, 1)$ OpenAI's column factorization [7, §5.1, before (65)] reads

$$(5.2) \quad \det[A_r(t_c)]_{r,c} = \mathcal{R}_n(z) \prod_c t_c^{C-1} (1-t_c)^h,$$

where $\mathcal{R}_n$ is OpenAI's sheet determinant. OpenAI proves two bounds for it [7, Props. 5.2, 5.3], with $K_0 = 10e^{12}$:

$$(5.3) \quad |\mathcal{R}_n(z)| \leq (1 + K_0 n)^n \mathcal{V}(1/z) \prod_c |z_c|^g \quad \text{if } \sum_c \frac{1-z_c^2}{1+z_c^2} \leq D_*,$$

$$(5.4) \quad |\mathcal{R}_n(z)| \leq (3/2)^n n^{n/2} 2^{-\binom{n}{2}} \prod_c (1+z_c^2)^{\frac{n-1}{2}} |z_c|^{g-(n-1)} \quad \text{always.}$$

Following OpenAI, $\kappa = 2$ refers to (5.3) and $\kappa = 1$ to (5.4). Let $w(x) = |t|^{C-1}(1-t)^h = 2^{C-1}|x|^{C-1}(1-x)^{2h}(1+x^2)^{-(C-1+h)}$.

The p -node majorants and their energy. For $x \in (-1, 1)^p$ and $s \in (0, 1)^p$ with distinct nonzero entries, put $t_i = 2x_i/(1+x_i^2)$ and

$$\begin{aligned} \mathcal{J}^{(p)}(x, s) &= \prod_i w(x_i) \cdot \frac{\mathcal{V}(t)\mathcal{V}(s)^2}{\prod_{i,j}(1-t_i s_j)} \prod_j s_j^b (1-s_j)^q, \\ \mathcal{I}_2^{(p)}(x, s) &= \mathcal{J}^{(p)}(x, s) \mathcal{V}(1/x) \prod_i |x_i|^{g-(n-p)}, \\ \mathcal{I}_1^{(p)}(x, s) &= \mathcal{J}^{(p)}(x, s) \prod_i (1+x_i^2)^{\frac{n-1}{2}} |x_i|^{g-(n-1)}. \end{aligned}$$

For $p = n$ these are OpenAI's majorants $\mathcal{I}_{2,n}$ and $2^{\binom{n}{2}} \mathcal{I}_{1,n}$ [7, (74)]. Let OpenAI's one-point fields be [7, (76)]

$$\begin{aligned} W_\kappa(x) &= \frac{19}{48} \log |x| + \frac{1}{12} \log(1-x) - \left(\frac{\kappa}{2} + \frac{17}{48}\right) \log(1+x^2), \\ W_s(s) &= \frac{7}{48} \log s + \frac{1}{12} \log(1-s), \quad D(x) = \frac{1}{6} - \frac{2x^2}{1+x^2}, \end{aligned}$$

and define the normalized energy

$$\begin{aligned} E_\kappa^{(p)}(x, s) = & \left(\frac{C}{n} + \frac{\kappa - 2}{2} - \frac{\kappa + 1}{2n} \right) \log 2 + \frac{1}{n} \sum_j W_s(s_j) \\ & + \frac{1}{n} \sum_i \left(W_\kappa(x_i) + \frac{\kappa + 2}{2n} \log(1 + x_i^2) \right) \\ & + \frac{1}{n^2} \left(\kappa \sum_{i < j} \log |x_i - x_j| + \sum_{i < j} \log(1 - x_i x_j) \right) \\ & + \frac{2}{n^2} \sum_{i < j} \log |s_i - s_j| - \frac{1}{n^2} \sum_{i, j} \log(1 - 2x_i s_j + x_i^2), \end{aligned}$$

all sums over indices $< p$; note the normalization by n , not p . For $p = n$ this is the normalized logarithm $n^{-2} \log \mathcal{I}_{\kappa, n} - \frac{1}{2} \log 2$ of OpenAI's energy argument. Expanding the coordinate change $t = 2x/(1 + x^2)$ gives, for every $p \leq n$,

$$(5.5) \quad \log \mathcal{I}_\kappa^{(p)}(x, s) = n^2 \left(E_\kappa^{(p)}(x, s) + \frac{1}{2} \log 2 \right) + e_\kappa(p) \log 2,$$

with $e_1(p) = p(C - 1) + \binom{p}{2} - n(C - 1)$ and $e_2(p) = e_1(p) - \binom{n}{2}$.

Retain the product in Lemma 5.6, rather than replacing each factor by 2. For $a \in [0, 1]$, put

$$E_{\kappa, a}^{(p)} = E_\kappa^{(p)} + \frac{a}{n} \sum_{j < p} \log(1 + s_j), \quad t = \frac{m}{n}.$$

The retained product contributes exactly $n^2(E_{\kappa, t}^{(p)} - E_\kappa^{(p)})$ to the logarithm of the majorant.

Proposition 5.8 (Pointwise bound). *In the setting of (5.1), let $K_n = \max\{(1 + K_0 n)^n, (3/2)^n n^{n/2}\}$, and let $\kappa = 2$ if $\sum_c \frac{1 - z_c^2}{1 + z_c^2} \leq D_*$ and $\kappa = 1$ otherwise. Then the integrand in (5.1) is at most*

$$K_n 2^{-\omega_\kappa N m} \exp \left(n^2 (E_{\kappa, t}^{(p)}(x, s) + \frac{1}{2} \log 2) \right), \quad \omega_1 = 28, \quad \omega_2 = 36.$$

Moreover, if $\kappa = 1$ then $\sum_{i < p} \frac{1 - x_i^2}{1 + x_i^2} > D_* - m$.

Proof. By (5.2), Cauchy's evaluation

$$\det \left[\frac{1}{1 - t_i s_j} \right] = \frac{V(t)V(s)}{\prod_{i, j} (1 - t_i s_j)}$$

[7, (63)] and Lemma 5.6, the integrand is at most

$$\left(\prod_j (1 + s_j) \right)^m |\mathcal{R}_n(z)| \prod_j w(y_j) \cdot \mathcal{J}^{(p)}(x, s).$$

For a fixed node $-1 < y < 0$, put $u = |y|$ and

$$R(u) = \frac{u^{19}(1 + u)^4}{(1 + u^2)^{40}}.$$

The identity

$$27(1+u^2)^4 - 256u^2 = (3u^2 - 1)^2(3u^4 + 14u^2 + 27)$$

gives $u^2/(1+u^2)^4 \leq 27/256$. Also $(1+u)^4 \leq 4(1+u^2)^2$ and $2u \leq (1+u^2)^2$, so

$$R(u) \leq 2(27/256)^9 \leq 2^{-28}.$$

The fixed-node numerator has powers $A = 19N$ and $2h = 4N$. The weaker of the two denominator exponents below is $C - 1 + h - (n-1)/2 = 41N - 1/2 \geq 40N$ for $N \geq 1$. For the second case, put $S(u) = u^{19}(1+u)^4/(1+u^2)^{64}$; its denominator exponent is $C - 1 + h = 65N - 1 \geq 64N$. Consequently

$$(5.6) \quad \begin{aligned} w(y)(1+y^2)^{\frac{n-1}{2}}|y|^{g-(n-1)} &\leq 2^{C-1}R(u)^N \leq 2^{C-1}2^{-28N}, \\ w(y)|y|^{g-(n-1)} &\leq 2^{C-1}S(u)^N \leq 2^{C-1}2^{-36N}. \end{aligned}$$

The stronger bound for S follows from two positive-polynomial identities. With $v = 1 + u^2$,

$$\begin{aligned} 3125v^6 - 46656u^2 &= (5u^2 - 1)^2(125u^8 + 800u^6 + 2190u^4 + 3344u^2 + 3125), \\ v^{10} - 2u(1+u)^4 &= (1-u)^2 + u^2(1-6u)^2 + u^4(1-u)^2 \\ &\quad + 119u^6 + \sum_{j=4}^{10} \binom{10}{j} u^{2j}. \end{aligned}$$

Thus $u^2/v^6 \leq 3125/46656$ and $u(1+u)^4/v^{10} \leq 1/2$, giving $S(u) \leq \frac{1}{2}(3125/46656)^9 \leq 2^{-36}$.

We use $\binom{n}{2} = \binom{p}{2} + pm + \binom{m}{2}$ repeatedly.

Case $\kappa = 1$. By (5.4), the first inequality in (5.6) and (5.5), the integrand is at most

$$K_n 2^{-28Nm} \exp(n^2(E_{1,t}^{(p)} + \frac{1}{2} \log 2)) \cdot 2^{-\binom{n}{2} + m(C-1) + e_1(p)},$$

and $-\binom{n}{2} + m(C-1) + e_1(p) = \binom{p}{2} - \binom{n}{2} \leq 0$. The case condition fails for z and each fixed node contributes at most 1 to its sum, which gives the last assertion.

Case $\kappa = 2$. By (5.3) we must bound $\mathcal{V}(1/z) \prod_c |z_c|^g$. Split $\mathcal{V}(1/z) = \mathcal{V}(1/x) \prod_{i,j} |1/y_j - 1/x_i| \mathcal{V}(1/y)$. Since $|1/y - 1/x| = |x - y|/|xy| \leq 2/|xy|$ and $\mathcal{V}(1/y) = \mathcal{V}(y) \prod_j |y_j|^{-(m-1)} \leq \prod_j |y_j|^{-(m-1)}$,

$$\mathcal{V}(1/z) \prod_c |z_c|^g \leq 2^{pm} \mathcal{V}(1/x) \prod_i |x_i|^{g-m} \prod_j |y_j|^{g-(n-1)}.$$

With the second inequality in (5.6) and (5.5), the integrand is at most $K_n 2^{-36Nm} \exp(n^2(E_{2,t}^{(p)} + \frac{1}{2} \log 2)) \cdot 2^{pm+m(C-1)+e_2(p)}$, and $pm + m(C-1) + e_2(p) = -\binom{m}{2} \leq 0$. $\square$

Remark 5.9. Each fixed node brings a factor 2^{C-1} , exactly compensated by the difference between the n -node and p -node normalizations. The remaining factors are 2^{-28N} and 2^{-36N} in the two cases. Retaining the numerator

and denominator together saves $32N \log 2$ per fixed node compared with the earlier separate bound $2^{2h} = 2^{4N}$.

5.6. The energy bound for $p \leq n$ nodes. OpenAI's certificate [7, §7] supplies, for $\kappa = 1, 2$, a parameter $\lambda_\kappa \geq 0$ ($\lambda_2 = 0$, $\lambda_1 = 2.47405979$), trial sequences $u^{(\kappa)}, v^{(\kappa)}$ (OpenAI's p, v), and rational upper bounds

$$(5.7) \quad \sup_{-1 \leq x < 1, x \neq 0} \{W_\kappa + \lambda_\kappa D - 2\kappa T(u^{(\kappa)}, \cdot) - S(v^{(\kappa)}, \cdot)\} \leq B_x^\kappa, \\ \sup_{0 < s < 1} \{W_s + 2T(v^{(\kappa)}, \cdot)\} \leq B_s^\kappa,$$

in the notation of [7, (77)–(78)]. For $\kappa = 1$ we use OpenAI's certificate, which gives $B_x^1 = -1.324376$ and $B_s^1 = -1.427976$. For $\kappa = 2$ we use the trial sequences of Proposition 5.11 in place of OpenAI's, with

$$B_x^2 = -0.98572217, \quad B_s^2 = -1.60942387.$$

With these, the dual value

$$U_\kappa = -\frac{11}{16} \log 2 + \kappa \|u^{(\kappa)}\|_*^2 + \frac{1}{2} \|v^{(\kappa)}\|_*^2 + B_x^\kappa + B_s^\kappa$$

satisfies $U_1 < -2.296789875$ [7, Prop. 7.1] and $U_2 < -2.29180925$ (Proposition 5.11). Put

$$c_\kappa = (\kappa + 2) \log 2 + |B_x^\kappa| + |B_s^\kappa| + \lambda_\kappa/6,$$

so that $c_1 = 5.24413\dots$ and $c_2 = 5.36773\dots$, and $c^* = 53678/10000 > \max(c_1, c_2)$.

Proposition 5.10 (Energy bound with missing points). *For every $d > 0$ there is N_0 such that for all $N \geq N_0$, all $p \leq n$ (with $m = n - p$) and all $x \in (-1, 1)^p$, $s \in (0, 1)^p$ with distinct nonzero coordinates,*

$$E_\kappa^{(p)}(x, s) \leq U_\kappa + d + \frac{m}{n} c_\kappa,$$

for $\kappa = 2$ unconditionally and for $\kappa = 1$ under $\sum_{i < p} \frac{1-x_i^2}{1+x_i^2} > D_* - m$.

Proof. OpenAI's proof of [7, Prop. 6.1] regularizes the logarithmic interactions by matched damping, controls the diagonal and endpoint corrections, and completes two squares with the trial sequences. All of these steps are estimates for sums over the points, normalized by n , and they remain valid, with the same error terms, uniformly for any number $p \leq n$ of points. The quadratic terms are bounded by $\kappa \|u\|_*^2 + \frac{1}{2} \|v\|_*^2$ as before. Only three places change.

- (1) The constant term. For $p = n$ the powers of 2 combine to $-\frac{11}{16} \log 2$. For $p < n$ the same computation gives $-\frac{11}{16} \log 2 + (\frac{\kappa}{2} + 1)(1 - p^2/n^2) \log 2$, and $1 - p^2/n^2 \leq 2m/n$.
- (2) The one-point terms. They are $\frac{1}{n} \sum_{i < p} X(x_i) \leq \frac{p}{n} B_x = B_x + \frac{m}{n} |B_x|$, since $B_x < 0$, and likewise for the column field with B_s .

- (3) The case field ($\kappa = 1$ only). OpenAI uses $\lambda \frac{1}{n} \sum_i D(x_i) \geq -\lambda/n$ when (5.3)'s condition fails. Since $D(x) = \frac{1-x^2}{1+x^2} - \frac{5}{6}$ and $D_* = \frac{5}{6}n - 1$, the shifted condition gives $\sum_{i < p} D(x_i) > -1 - \frac{m}{6}$, which costs $\lambda(1 + m/6)/n$.

The damping parameter is chosen in terms of d alone, and the remaining errors are $O_d(1/n)$, uniformly in p and the configuration. $\square$

5.7. A new certificate for the second case of the energy dual. For a finite list $c_1, \dots, c_K$ of reals and finitely many pairs (z_j, r_j) of complex numbers with $|z_j| < 1$, consider the trial sequence

$$(5.8) \quad u_k = c_k + \operatorname{Re} \sum_j r_j z_j^k \quad (k \geq 1; \ c_k = 0 \text{ for } k > K).$$

This is the form of OpenAI's trial sequences [7, §7].

Proposition 5.11. *There are trial sequences u, v of the form (5.8), with $K = 40$, coefficients c_k in $10^{-8}\mathbb{Z}$, twelve pairs (z_j, r_j) for u and fifteen for v , all z_j and r_j in $\mathbb{Q} + \mathbb{Q}i$ (they are listed in the file `Dual/Data.lean` of the formalization), such that (5.7) holds for $\kappa = 2$ with $B_x^2 = -0.98572217$ and $B_s^2 = -1.60942387$, and $\|u\|_*^2 \leq 0.066723974$, $\|v\|_*^2 \leq 1.292855037$. Hence*

$$U_2 \leq -\frac{11}{16} \log 2 + 2 \cdot 0.066723974 + \frac{1}{2} \cdot 1.292855037 \\ - 0.98572217 - 1.60942387 < -2.29180925,$$

and Proposition 2.1(b) holds with $U = 2.29180925 - \frac{1}{2} \log 2$ in place of U_{OAI} .

The sequences were found numerically, by maximizing the dual value $-U_2$ over sequences of the form (5.8) with prescribed bases z_j (OpenAI's, together with ± 0.9 , ± 0.97 , 0.96 , ± 0.99 and a few others) by a cutting-plane method, and then rounded. At the optimum the two fields of (5.7) nearly equioscillate, with about fifty and thirty local maxima.

Proof. All computations below are in exact integer and rational arithmetic and are carried out by Lean's kernel. We follow the structure of OpenAI's certificate, but make each step a decidable check of general lemmas.

Closed forms. For $|x| \leq 1$, OpenAI's formulas give

$$T(u, x) = \sum_{k \leq K} \frac{c_k T_k(x)}{k} - \frac{1}{2} \operatorname{Re} \sum_j r_j \log(1 - 2xz_j + z_j^2)$$

and $S(u, x) = \sum_{k \leq K} c_k x^k / k - \operatorname{Re} \sum_j r_j \log(1 - xz_j)$. Expanding the squares and using $\sum_{k \geq 1} w^k / k = -\log(1 - w)$,

$$\|u\|_*^2 = \sum_{k \leq K} \frac{c_k^2 + 2c_k t_k}{k} - \frac{1}{2} \operatorname{Re} \sum_{j, l} (r_j r_l \log(1 - z_j z_l) + r_j \bar{r}_l \log(1 - z_j \bar{z}_l)),$$

$$t_k = \operatorname{Re} \sum_j r_j z_j^k.$$

So the norms and the values of the fields at rational points are rational numbers plus finite sums of terms $\operatorname{Re}((c+di) \log w)$ with c, d, w rational. Each such term is approximated to 10^{-12} by OpenAI's rational approximations of the logarithm and the arctangent [7, §7]; the reduction of w to the range of these approximations is computed and checked rationally.

Stationary points. The derivative of each field F is a sum of rational functions with rational coefficients, hence $F' = A/Q$ with explicit $A, Q \in \mathbb{Q}[x]$ (of degrees 78 and 39 for the row field, 60 and 21 for the column field), where Q does not vanish on the domain. The domain is cut into 154 (respectively 94) intervals. On 104 (respectively 63) of them A has no root: after the substitution $x = a + (b-a)/(1+y)$, all coefficients of $(1+y)^{\deg A} A(x)$ have the same sign, so that A has no root in (a, b) (the case of Descartes's rule of signs without sign changes). The remaining 50 (respectively 31) intervals have length $2 \cdot 10^{-7}$ and contain the roots of A ; a breakpoint between two intervals of the first kind is not a root of A .

Brackets. On a bracket $[l, l+w]$ write $A(l+wt) = \sum a_i t^i$ and $Q(l+wt) = \sum q_i t^i$. For $0 \leq t \leq 1$, $|F'| \leq \sum |a_i|/(|q_0| - \sum_{i \geq 1} |q_i|)$; since A has a root in the bracket, a_0 and a_1 are small, and this gives $|F(x) - F(l)| \leq 10^{-8}$ on the bracket. The value $F(l)$ is computed as above.

Conclusion. The fields tend to $-\infty$ at 0 and 1 (the row field at $0^\pm$ and 1^-), so their suprema are attained at stationary points or at $x = -1$, and every stationary point lies in a bracket. Comparing the certified values with B_x^2 and B_s^2 proves (5.7). The bound for U_2 follows, with $\log 2 > 0.6931471803$. OpenAI's proof of [7, Prop. 7.1] applies to any trial sequences in its two cases and gives $\limsup_N \log |\det \mathcal{A}_N|/n^2 \leq \max(U_1, U_2) + \frac{1}{2} \log 2$, and $U_1 < -2.296789875$. $\square$

For efficient replay, the implementation clears polynomial denominators once and performs the root-cover and bracket substitutions over $\mathbb{Z}$, retaining rational arithmetic for the final scalar ratio and field approximations. Lean proves the coefficient conversions and the soundness of both substitutions. This changes the computation of the certificates, while preserving their trial sequences, covers and certified bounds.

Remark 5.12. OpenAI's own certificate for $\kappa = 2$ records its bounds to five digits, as $B_x^2 \leq -0.98399 + 2.4 \cdot 10^{-5}$, $B_s^2 \leq -1.60890 + 2.4 \cdot 10^{-5}$ and a norm term at most 0.77844, while its computations show that the largest bracket values of its two fields are $-0.98403338\dots$ and $-1.60894641\dots$ and its norm term is at most 0.778423472. Read to seven digits, OpenAI's certificate already gives $U_2 < -2.2910468$, and hence the exponent 34 000 by the argument of Section 6. This is also formalized (**SharpBarrier**).

5.8. A weighted endpoint certificate. For each fixed $a \in \{0, 1\}$, repeat the proof of Proposition 5.10 with the column field $W_s + 2T(v, \cdot) + a \log(1+s)$, retaining the quadratic mass correction instead of using $1 - p^2/n^2 \leq 2m/n$. Write $K_a = \kappa \|u_a\|_*^2 + \frac{1}{2} \|v_a\|_*^2$ and $b_a = B_{x,a} + B_{s,a}$. The resulting bound,

uniform in the same configurations, is

$$E_{\kappa,a}^{(p)} \leq -\frac{11}{16} \log 2 + K_a + (1-t)b_a + (\frac{\kappa}{2} + 1)(2t - t^2) \log 2 + \lambda_a t/6 + d.$$

Its error is uniform because $0 \leq a \leq 1$ and the added field is bounded and smooth at the damping endpoint. One may weaken it before damping at a cost $a \log 2/n$, which tends uniformly to zero. Since $E_{\kappa,t}^{(p)} = (1-t)E_{\kappa,0}^{(p)} + tE_{\kappa,1}^{(p)}$, the two fixed endpoint inequalities give

$$(5.9) \quad E_{\kappa,t}^{(p)} \leq V_0 + tC + t^2Q + d,$$

where

$$\begin{aligned} V_0 &= -\frac{11}{16} \log 2 + K_0 + b_0, \\ C &= K_1 - K_0 + b_1 - 2b_0 + (\kappa + 2) \log 2 + \lambda_0/6, \\ Q &= b_0 - b_1 - (\frac{\kappa}{2} + 1) \log 2 + (\lambda_1 - \lambda_0)/6. \end{aligned}$$

Thus only two fixed trials need certification; no interpolated family of norms is assumed.

For $\kappa = 2$, the $a = 0$ endpoint is Proposition 5.11. A second rational trial, listed in `WeightedDualTwo/Data.lean`, certifies

$$\begin{aligned} B_{x,1} &= -0.9498627, & B_{s,1} &= -1.2145624, \\ \|u_1\|_*^2 &\leq 0.0713397, & \|v_1\|_*^2 &\leq 1.3114478, & \lambda_0 &= \lambda_1 = 0. \end{aligned}$$

Here $B_{s,1}$ bounds the column field including $\log(1+s)$. The same exact rational checks used in Proposition 5.11 apply: the derivative numerators have degrees 81 and 62, and the covers have 159 and 93 pieces, including 51 and 31 stationary-point brackets. These caps give $C \leq 5.817$, $Q \leq 0$, and $V_0 < -2.29180925$.

For $\kappa = 1$, keep OpenAI's original trial at both endpoints, with $K_0 = K_1 = 0.9321$ as a common norm cap and $\lambda_0 = \lambda_1 = 2.47405979$. The new column check is

$$W_s + 2T(v^{(1)}, s) + \log(1+s) \leq -0.7586255.$$

Its derivative is rational with numerator degree 9 and denominator degree 3; a cover of five pieces has one stationary-point bracket. Combined with the old caps, it gives $C \leq 5.913488$, $Q \leq -1.70907$, and $V_0 \leq -2.29180925 - 0.004980625$. The extra constant margin absorbs the larger linear coefficient: for $t \geq 0$,

$$0.183488t - 1.70907t^2 \leq \frac{0.183488^2}{4 \cdot 1.70907} < 0.004980625.$$

Consequently the two cases give

$$(5.10) \quad E_{1,t}^{(p)} \leq -2.29180925 + d + 5.73t, \quad E_{2,t}^{(p)} \leq -2.29180925 + d + 5.817t.$$

All the new field, norm, cover and algebraic inequalities are checked by Lean's kernel.

The additional case-two decay absorbs its larger energy cost: since $8 \log 2 > 48(5.817 - 5.73)$ and $n = 48N$, its product of fixed-node and energy factors is bounded by the common factor $2^{-28Nm} \exp(nm \cdot 5.73)$. This uses the two cases together and does not change their trial sequences.

5.9. Completion of the proof.

Proof of Theorem 5.1. Fix $d > 0$ and let N be large. By (5.1), Proposition 5.8, (5.10), with $c^* = 5.73$ and $-2.29180925 = -U - \frac{1}{2} \log 2$, every f with $m = m(f)$ satisfies

$$|\det M_f| \leq 2^n K_n 2^{-28Nm} \exp(n^2(-U + d) + nm c^*) = M c^m,$$

with

$$M = 2^n K_n e^{n^2(-U+d)}, \quad c = 2^{-28N} e^{nc^*}.$$

(The degenerate configurations, such as two fixed nodes being equal or a vanishing coordinate, have measure zero or make the determinant vanish.) By Lemma 5.2, $|\beta_{a,k}| \leq B = 8J^2 K^N$. Now

$$\begin{aligned} \log(nJBc) &= \log(8nJ^3) - 28N \log 2 + N \log K + nc^* \\ &= -28N \log 2 + N \log K + 48Nc^* + \log(8 \cdot 48 \cdot 65^3 N^4), \end{aligned}$$

and $-28 \log 2 + \log K + 48c^* < 366.107$; so $nJBc \leq e^{366.3N}$ for $N \geq 4000$. If $|\delta| \leq e^{-366.3N}$ then $J|\delta|Bc \leq 1/n$ and $(1 + J|\delta|Bc)^n - 1 \leq e - 1 < 2$. Lemma 5.4 gives

$$|\Delta_N(x) - \Delta_N(G)| \leq 2M = 2^{n+1} K_n e^{n^2(-U+d)}.$$

Since $\log K_n = O(n \log n)$, taking $d = \varepsilon/2$ proves the theorem. $\square$

Remark 5.13 (The value of ρ). Per unit of N , the quadrature weights contribute $\log K < 110.475$, the fixed nodes contribute $-28 \log 2$, and the weighted energy contributes $48 \cdot 5.73 = 275.04$. The polynomial prefactor is absorbed by the remaining margin to $\rho = 366.3$. Retaining the Vandermonde product replaces the former combined charge $48 \log 2 + 48 \cdot 5.3678$ by the weighted energy charge, then the separate fixed-node caps and the quadratic margin lower it to $48 \cdot 5.73$.

6. THE $48N$ BOUND

The earlier 1981 bound. Let $\gamma_* = U - A_* = \gamma_0 + 5/1728$ and $\eta = 8969/1000$. The certified old margin (2.4) implies $2304\gamma_* > 2304 \cdot 9.996945 \cdot 10^{-4} + 20/3 > 8.969$. For sufficiently small $\varepsilon > 0$ and large primes p , the real bound and perturbation estimate give $|\Delta_p(x)| \leq 2e^{n^2(-U+\varepsilon)}$ on $|x - G| \leq e^{-366.3p}$. The factorial and the factor two are $e^{o(n^2)}$, while Proposition 3.5 bounds D_p^* by $e^{n^2(A_*+\varepsilon)}$. Thus $Q_p = (44p)! D_p^* \Delta_p$ satisfies $|Q_p(x)| \leq e^{-\eta p^2}$ for large p . Corollary 3.8 gives its degree and square-denominator nonvanishing.

Apply Proposition 4.1 with $d = 44$, $\rho = 366.3$ and $R = d/\eta = 44000/8969$. Choose $c = 10001/10000$ and $\varepsilon = c/(2R) = 89698969/880000000$. The PNT exponent is

$$(1 + \varepsilon)\rho R = 366.3 \left(\frac{44000}{8969} + \frac{10001}{20000} \right) = 1980.157945 \dots < 1981.$$

With the Chebyshev window the exponent is $3\rho R = 5390.968892 \dots < 5391$. $\square$

Remark 6.1 (Further reductions). The sparse-row mechanism raises the denominator gap without changing the determinant's degree or the perturbation estimate. Further improvements can come from additional residue dependencies in other prime windows, more efficient determinant parameters, or a smaller perturbation exponent. The contribution $\rho/2$ of the square-prime window alone exceeds two even for the improved $\rho = 120$, so reaching irrationality exponent two requires a further structural improvement; the present proof does not establish such a bound.

7. A TWENTY-ROW PARAMETER FAMILY

The preceding sections explain the $48N$ construction and its 1981 bound. We now change the determinant parameters and prove the stronger bound in Theorem 1.1. This change requires new arithmetic and real-place estimates; substituting parameters into the earlier numerical constants would not suffice.

7.1. The determinant and its arithmetic. Use the same moment and contact construction with

$$\begin{aligned} n &= 20N, & b &= 3N, & q &= N, & g &= 2N, & h &= N, \\ L &= 24N, & C &= 26N, & H &= 27N, & A &= 8N. \end{aligned}$$

In the definitions of the row polynomials, replace the previous parameter values by these values: the offset is $r - 2N$, the reversal length is $26N$, and the common row factor is $(1 - t)^N$. The filtered columns are $t^{3N+k}(1 - t)^N$. Write $\Delta_N^{(20)}(X)$ for the resulting determinant polynomial. Row reflection gives

$$\deg \Delta_N^{(20)} \leq 18N.$$

The three 20×20 base matrices corresponding to signs $0, 1, -1$ are invertible. Exact inverse certificates modulo 1009 prove their rational nonvanishing. The prime-index reduction then proves uniformly, outside a fixed finite set of primes,

$$\Delta_p^{(20)}(a/b) \neq 0 \quad \text{if } p^2 \nmid b.$$

The finite inverse computations certify the base matrices; the deduction for all large primes is a separate theorem, not an extrapolation from those computations.

For large odd primes with $27N < p^2$, the actual filtered determinant satisfies $v_p(\Delta_N^{(20)}(z)) \geq -Nf(p/N)$ whenever z has p -unit denominator. The physical column pools and their rational inverse maps give

$$f(x) = \begin{cases} 40 & 0 \leq x \leq 10, \\ 60 - 2x & 10 \leq x \leq 12, \\ 36 & 12 \leq x \leq 27/2, \\ 63 - 2x & 27/2 \leq x \leq 14, \\ 91 - 4x & 14 \leq x \leq 16, \\ 75 - 3x & 16 \leq x \leq 24, \\ 27 - x & 24 \leq x \leq 27, \\ 0 & 27 \leq x. \end{cases}$$

Each bound applies to the actual determinant: the pool reconstruction retains the integer filter, and the upper-range central remainders are proved p -integral. The exact integral is $\int_0^{27} f = 2937/4$. The proved sparse row relations add the saving

$$g(x) = 3(x - 22/3)_+ - 6(x - 8)_+ + 3(x - 26/3)_+.$$

It is supported on $[22/3, 26/3]$, where $f = 40$, and has integral $4/3$. Thus $f - g \geq 0$, and

$$\frac{1}{400} \int_0^{27} (f - g)(x) dx = \frac{8795}{4800}.$$

The continuous weighted prime-number theorem proves that this integral is the normalized limiting odd-prime cost. The finitely many excluded primes and primes at most $2\sqrt{65N}$ contribute $o(N^2)$, using the already proved logarithmic entry bounds. Primes above $27N$ are locally integral.

The 2-adic smoothing and exceptional-column expansion also survives the parameter change. Nonzero coefficient minors have distinct coefficient indices within each exceptional family, and their raw column indices are distinct. If m and ℓ are the two exceptional populations, the quadratic part of the selector exponent is

$$\frac{3}{2}m^2 + \frac{1}{2}\ell^2 - 50Nm - 26N\ell + 520N^2 \geq -42N^2, \quad m, \ell \geq 0, \quad m + \ell \leq 20N.$$

Indeed, adding $42N^2$ to the left side gives

$$2(m - 11N)^2 + (20N - m - \ell)(6N + m) + \frac{1}{2}(20N - m - \ell)^2.$$

The full minor expansion, not just this quadratic inequality, yields

$$\|\Delta_N^{(20)}(z)\|_2 \leq (8F(z)(27N)^2)^{20N} 2^{42N^2 + 40N}.$$

At integer points $F(z) \leq F(0)$. Consequently a common denominator $D_N^{(20)}$ clears every integer value and, for every $\epsilon > 0$, eventually

$$\log D_N^{(20)} \leq \left(\frac{8795}{4800} + \frac{21}{200} \log 2 + \epsilon \right) (20N)^2.$$

The formal construction uses a harmless eventual threshold $N \geq 10^{12}$ to absorb all finite prime cutoffs. No effective threshold for the final rational denominator is asserted.

7.2. Free and fixed populations. Recompute both sheet duals for the new parameters. Four rational trial certificates cover the two sheet cases at the two endpoints of the missing population. Their global field bounds, derivatives, norms and finite diagonal corrections are checked by the kernel. The certified dual target is

$$U_{20} = \frac{2256085}{1000000}.$$

The physical determinant energy retains the correction $+\frac{1}{2} \log 2$. Fixed-node fields include the common-factor cancellation in the quadrature weights; their costs are not replaced by the free-node fields. The finite population interpolation gives, uniformly in every missing mass m , an energy loss at most $4.3 m/(20N)$, in addition to an arbitrarily small regularization error. Damping is chosen before taking N large, so this is a uniform bound rather than a numerical cutoff near the endpoints.

Use $18N$ Chebyshev nodes for the residual row polynomials, and a circle of radius $1001/1000$. The exact circle certificate gives $K \leq e^{33}$. The joint common-factor and quadrature-weight bound is $54000NK^N$ per atom. Combining it with the population estimate gives the actual mixed integrand, multiplied by all its atom weights, the bound

$$S_N(54000N)^m \exp\left((20N)^2 \left(-U_{20} + \frac{1}{2} \log 2 + \delta\right) + 119Nm\right),$$

where

$$S_N = (3/2)^{20N} (20N)^{10N} + (1 + 10e^{12}(20N))^{20N}, \quad \log S_N = o(N^2).$$

This estimate covers both sheets. Zero nodes and repeated nodes give zero determinants and are included separately.

The mixed Andréief identity can retain the product of the atom weights on both sides of the integral. The determinant column expansion then uses their joint bound, with atom cost $54000Ne^{119N}$. Since

$$(20N)(18N)(54000N)e^{119N} \leq e^{120N}$$

eventually, for every $\epsilon > 0$ the actual determinant polynomial satisfies

$$\begin{aligned} |\Delta_N^{(20)}(G)| &\leq e^{(20N)^2(-U_{20} + \frac{1}{2} \log 2 + \epsilon)}, \\ |\Delta_N^{(20)}(x) - \Delta_N^{(20)}(G)| &\leq e^{(20N)^2(-U_{20} + \frac{1}{2} \log 2 + \epsilon)} \quad \text{if } |x - G| \leq e^{-120N}. \end{aligned}$$

7.3. The resulting measure. Set

$$A_{20} = \frac{8795}{4800} + \frac{21}{200} \log 2, \quad \gamma_{20} = U_{20} - \frac{1}{2} \log 2 - A_{20}.$$

The certified rational budget gives $400\gamma_{20} > 177/100$. The integer-valued polynomial argument supplies $Q_N^{(20)} = (18N)! D_N^{(20)} \Delta_N^{(20)} \in \mathbb{Z}[X]$ for all

sufficiently large N ; smaller indices can be assigned the zero polynomial. Factorial growth is $o(N^2)$, so the criterion applies with

$$d = 18, \quad \eta = 177/100, \quad \rho = 120.$$

Choose $\varepsilon = 247/5000$ and $c = 1001/1000$ in the square-denominator prime window. Its exponent is

$$(1 + \varepsilon)120 \max \left(\frac{18}{177/100}, \frac{1001/1000}{2(247/5000)} \right) = \frac{377784}{295} < 1281.$$

The Chebyshev criterion gives $216000/59 < 3662$. These prove Theorem 1.1. The Lean statements in `Parametric/TwentyMain` are `catalan_irrationalityMeasure_1281` and its 3662 counterpart; the public series statements appear in `Solution.lean`.

8. FORMALIZATION

All results of this paper are formalized in Lean 4 [6] with Mathlib [5], on top of OpenAI's Lean formalization of the irrationality of G [7]. The development is available at

<https://github.com/slitsey/catalan-irrationality-measure>

and combines analytic and arithmetic proofs with generated rational data and kernel checks for Proposition 5.11 and Section 5.8. The uniform sparse and combined prime-row dependencies and the weighted-PNT denominator saving are also proved in Lean; numerical rank scans alone are not used as proof. The historical sharpened OpenAI certificate is retained separately. It uses OpenAI's definitions directly: Catalan's constant is `catalan`, defined as the series $\sum_j (-1)^j / (2j+1)^2$; the matrix (2.2) and its rational form are OpenAI's `determinant` and `determinantRat`. The current main statements are in `CatalanMeasure.Parametric`; the listed historical lemmas remain in `OAI.InternalCatalan`:

- `catalan_irrationalityMeasure_1281`: for all sufficiently large $b \in \mathbb{N}$ and all $a \in \mathbb{Z}$, $b^{-1281} < |G - a/b|$ (Theorem 1.1);
- `catalan_irrationalityMeasure_3662`: the same with exponent 3662;
- `eventually_realDeterminant_abs_upper_dual`: Proposition 5.11;
- `separate_perturbation`: Theorem 5.1, with $\rho = 366.3$;
- `catalan_irrationalityMeasure` and `catalan_irrationalityMeasure_pnt`: the general exponents of Proposition 4.1.

There is no `sorry`, and the command `#print axioms` reports only the three standard axioms `propext`, `Classical.choice` and `Quot.sound`. The prime number theorem enters through OpenAI's workspace, which already uses it for Proposition 2.1(c).

File	Content
DeltaPoly, Degree	$\Delta_N(X)$; Lemma 3.1
Denominators,	Proposition 3.2; Lemma 3.6; Corollary 3.8
IntegerValued, Integrality	
SparseDenominators,	Propositions 3.3–3.5; the earlier 1981 bound
DenseDenominators,	
DenseGainPNT, DenseMain	
Nonvanishing	Proposition 2.1(d) for $p \nmid \text{den } z$
PrimeWindow, Criterion,	Proposition 4.1 and the reduction to
PNTWindow, Main	Theorem 5.1
Quadrature, ChebyshevNodes,	Lemma 5.2
QuadCircle	
DetExpansion	Lemma 5.4
MixedAndreief	Lemma 5.5
VandermondeProduct	Lemma 5.6
PinMajorant, FixedNodes,	Proposition 5.8, (5.6), and (5.5)
FixedNodeMajorant	
PinDampedDual, PinSupDual	Proposition 5.10
Cert/ListPoly,	the general lemmas behind the checks of
Cert/LogApprox,	Proposition 5.11
Cert/FieldEval,	
Cert/FieldDeriv,	
Cert/Cover, Cert/Norm	
Dual/*, DualTwo	the certificate data, the kernel checks, Proposition 5.11
SharpBarrier	the remark after Proposition 5.11
WeightedEnergy,	Section 5.8
WeightedBounds	
SeparateFixedNodes,	Theorem 5.1
SeparateBounds,	
SeparateLemma6	
SquareNonvanishing,	Proposition 3.7, strengthened
SquarePNTWindow,	Proposition 4.1
SquareMain	
Parametric/TwentyMain,	Section 7; the current 1281/3662 bounds
TwentyIntegerPolynomial	
TwentyDenominators,	Actual local arithmetic and
TwentyTwoAdicNorm,	common-denominator growth
TwentyCombinedOddLoss	
TwentyTrial/*,	Certified population fields, mixed
TwentyUniformEnergy,	integration and $\rho = 120$
TwentyPerturbation	

The current proofs were checked with Lean v4.35.0-rc2 and the matching Mathlib release. OpenAI’s original repository at commit `fd4aeeb` pins Lean v4.34.1 and Mathlib commit `d13f23b`. The parts of that formalization used here were converted to Lean’s module system and ported to the current

toolchain; the repository's README and provenance documents describe the changes.

DISCLOSURE OF AI USE

The original development was produced by the AI system Claude Opus 5.5 (Anthropic): the mathematics, including the perturbation estimate and all proofs; the Lean formalization; the numerical computations and the certificate generator; the literature search; and the text of this paper. Claude worked in a series of conversations at the request of the author, who posed the problem, directed the work, and is responsible for this submission. The separate fixed-node decay bounds and weighted Vandermonde refinement, endpoint certificates, quadratic case refinement, dependency refactor, integer certificate computations and their soundness proofs, square-denominator nonvanishing, the prime-window refinement, uniform sparse and combined prime-row dependencies and denominator saving, the twenty-row parameter family, certified free and fixed population fields, joint atom integration, and corresponding revisions to this paper were developed by OpenAI Codex running GPT-6.1 Sol (OpenAI) at the author's request. The underlying irrationality proof and its formalization are OpenAI's [7]. The formal verification in Lean is intended as an independent check of every step of the argument above.

REFERENCES

- [1] F. Calegari, V. Dimitrov and Y. Tang, *The linear independence of 1, $\zeta(2)$, and $L(2, \chi_{-3})$* , preprint, [arXiv:2408.15403](https://arxiv.org/abs/2408.15403).
- [2] P. Eskandari, *Improved rational approximations to Catalan's constant*, preprint (2026), [arXiv:2609.26354](https://arxiv.org/abs/2609.26354).
- [3] P. J. Forrester, *Meet Andréief, Bordeaux 1886, and Andreev, Kharkov 1882–83*, preprint (2018), [arXiv:1806.10411](https://arxiv.org/abs/1806.10411).
- [4] R. K. Jha, *A finite irrationality measure for Catalan's constant*, preprint v1, October 8, 2026, [doi:10.5281/zenodo.23237269](https://doi.org/10.5281/zenodo.23237269); Lean sources at <https://github.com/rohitkrjha/catalan-measure>.
- [5] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), 367–381.
- [6] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction – CADE 28, Lecture Notes in Computer Science 12699, Springer, 2021, 625–635.
- [7] OpenAI, *Catalan's constant is irrational*, preprint (September 24, 2026), with a Lean formalization, available at <https://github.com/openai/math> (commit `fd4aeeb`, October 7, 2026).
- [8] T. Rivoal and W. Zudilin, *Diophantine properties of numbers related to Catalan's constant*, Math. Ann. **326** (2003), 705–721.
- [9] Z.-W. Sun, *Catalan's constant is irrational*, preprint (2026), [arXiv:2609.04176](https://arxiv.org/abs/2609.04176).
- [10] D. Wachs, *A note on a recent claimed proof of the irrationality of Catalan's constant*, preprint (2026), [arXiv:2609.22339](https://arxiv.org/abs/2609.22339).
- [11] W. Zudilin, *An Apéry-like difference equation for Catalan's constant*, Electron. J. Combin. **10** (2003), R14.

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