THE SHARP CONSTANT IN THE DEGREES-OF-FREEDOM BOUND FOR BINARY ISOTONIC REGRESSION: AN EXACT CERTIFICATE, AND AN ERRATUM TO A NUMERICAL CHECK

KOREA SUPERINTELLIGENCE LABS

Abstract. Rossellini, Barber, Ren and Soloff [1] bound the number of distinct fitted values of isotonic regression on a binary sample of size n by an explicit quantity $f(n)$ built from the Euler totient, and write $f(n) = c\,n^{2/3} + \delta_n$ with $c = 3/(4\pi^2)^{1/3}$. A proposition of that paper asserts $\delta_n \le 0.63\,n^{1/6}$ for all $n \le 2 \cdot 10^{23}$, and reduces it by two concavity interpolations to the Farey-type inequality $R_K := B_K - c\,n_K^{2/3} \le 0.63\,n_K^{1/6}$ for all $K \le 10^8$, where $n_K = 1+\sum_{j\le K} j\varphi(j)$ and $B_K = 1+\sum_{j\le K}\varphi(j)$ are the sum of the denominators and the number of terms of the Farey series of order K; that inequality is verified there by a printed floating-point script. We determine the optimal constant. It is

$$\alpha^* = 5\cdot 10^{-1/6} - 3\sqrt{10}\,(4\pi^2)^{-1/3} = 0.6203305642542976\ldots,$$

a transcendental number. We prove $R_K \le \alpha^* n_K^{1/6}$ for every $K \le 10^6 + 1$, with equality at $K = 3$; hence $\delta_n \le \alpha^* n^{1/6}$ for every $n \le n_{10^6} = 202\,642\,368\,741\,515\,820$, which is 17 of the 23 decades the source asserts, and no smaller constant works on any range containing $n = 10$, where $f(10) = 5$ and $\delta_{10} = \alpha^* \cdot 10^{1/6}$ exactly. The finite verification is an exact integer certificate: the inequality is cubed to remove the cube root in c, sixth roots enter as checked rationals $m/2^{20}$, and π enters only through a twenty-digit rational bound, so no floating-point arithmetic occurs anywhere in it. We also record an erratum to the source's printed script, whose cumulative sums are accumulated in binary64 through an unintended integer-type promotion: at $K = 10^6$ the exact n_K is $202\,642\,368\,741\,515\,820$ against the script's $202\,642\,368\,741\,518\,368$, and near $K = 9.3\cdot 10^7$ the script's value of the quantity it maximises has the wrong sign. The proposition's conclusion is unaffected, the maximum sitting at $K = 3$ in small exact integers; this corrects the certificate, not the theorem. Everything above is machine-checked in Lean 4 except three things, each flagged in the text: the transcendence, a one-line consequence of Lindemann's theorem proved by hand; the decimal digits of α^* beyond the ten of the certified enclosure; and the measurements of the printed script over its full range, of which only the exact value at $K = 10^6$ is formal.

1. Introduction

1.1. **Distinct fitted values of binary isotonic regression.** Isotonic regression fits a monotone step function to data by least squares. When the responses are binary, the number of distinct values the fit can take is limited by arithmetic rather than by the data. Each block of the fitted step function is an average of 0's and 1's over a run of observations, so in lowest terms its value is a fraction a/b whose denominator b is at most the length of that run; distinct blocks occupy disjoint runs, so the denominators of the distinct fitted values of a fit on n observations sum to at most n. Maximising the number of distinct reduced fractions in $[0,1]$ subject to a bound on the sum of their denominators is a question about Farey series, which is where the Euler totient φ enters: for $j \ge 2$ there are exactly $\varphi(j)$ reduced fractions of denominator j in $[0,1]$, and buying all of them costs $j\,\varphi(j)$ against the budget; denominator 1 contributes the two fractions $0/1$ and $1/1$, which is the source of the additive constants below. The reasoning just given is the source's own, from the proof of its Theorem 1, where the level-set averaging property of isotonic regression is what reduces the count to a Farey series; it is included for orientation, and nothing below depends on it.

Rossellini, Barber, Ren and Soloff [1] make this exact. Writing φ for the Euler totient, they consider

$$f(n) = \sup\Big\{\, 1 + \sum_{j=1}^{K-1} \varphi(j) + \ell \;:\; K \ge 1,\ 0 \le \ell \le \varphi(K),\ 1 + \sum_{j=1}^{K-1} j\,\varphi(j) + \ell K \le n \Big\}, \tag{1}$$

which is the greedy optimum of the counting problem just described, and which their Theorem 1 identifies as the exact worst case: writing $\hat{k}(y)$ for the number of distinct fitted values of the isotonic least-squares fit to y, that theorem states

$$\sup_{y\in\{0,1\}^n} \hat{k}(y) = \frac{3}{(4\pi^2)^{1/3}}\, n^{2/3} + \delta_n, \qquad |\delta_n| = O(n^{1/3}\log n),$$

with δ_n "specified exactly in the proof" by (1); they then write

$$f(n) = c\, n^{2/3} + \delta_n, \qquad c = \frac{3}{(4\pi^2)^{1/3}} = 0.88105159648984077\ldots \tag{2}$$

The leading term is the classical Farey count: the average order of the totient gives $\sum_{j\le K}\varphi(j) = 3K^2/\pi^2 + O(K\log K)$ (see [3]), whence $\sum_{j\le K} j\varphi(j) = 2K^3/\pi^2 + O(K^2\log K)$ by Abel summation, so eliminating K between the two gives the exponent $2/3$ and the constant c. The interest here is in the error term δ_n: it is the only part of (2) not given in closed form, and the source devotes an appendix to bounding it.

1.2. **The source's numerical proposition.** On the size of δ_n the source proves one thing, in an appendix proposition — Proposition 9 of the v1 e-print, internal label `prop:error_term_certificate`; all numbering we cite is that of v1:

$\delta_n \le \alpha\, n^{1/6}$ for all $n \le 2\cdot 10^{23}$, where $\alpha = 0.63$.

The proof interpolates twice by concavity, which reduces the claim to a statement about the two integer sequences

$$n_K = 1 + \sum_{j\le K} j\,\varphi(j), \qquad B_K = 1 + \sum_{j\le K}\varphi(j), \qquad R_K = B_K - c\, n_K^{2/3}, \tag{3}$$

and then says, verbatim: *"Numerically, we check that $R_K \le \alpha\, n_K^{1/6}$ for all $K \le 10^8$ (see the Python code that follows the proof)."* The code that follows is reproduced here as (6) in Section 6, because it is the object of one of our results.

Two questions are left open by that sentence, and this paper answers them.

Is 0.63 *the right constant?* It is a rounding of something, and the source does not say of what, nor whether its proposition is tight. We show that the constant is attained, identify it in closed form, and prove that it is transcendental.

What does the printed script actually certify? It is the only evidence offered for the inequality over the asserted range. We show that its arithmetic silently leaves the exact integers — n_K passes 2^{53}, the largest integer a binary64 number represents exactly, at $K \approx 3.5\cdot 10^5$ — and that at large K the quantity it maximises is computed with errors far exceeding the values being maximised, while confirming, in exact arithmetic, that the conclusion it was used to reach is nevertheless true.

1.3. **Results.** Throughout, c is as in (2) and n_K, B_K, R_K as in (3), and

$$\alpha^* := \frac{R_3}{n_3^{1/6}} = \frac{5 - c\cdot 10^{2/3}}{10^{1/6}}. \tag{4}$$

The first result identifies this number.

Proposition 1.1 (the constant, in closed form)**.** $\alpha^* = 5\cdot 10^{-1/6} - 3\sqrt{10}\,(4\pi^2)^{-1/3}$*, and* $0.6203305642 < \alpha^* < 0.6203305643$*. Moreover* α^* *is transcendental.*

The next three are the substance.

Theorem 1.2 (the sequence inequality, with the optimal constant)**.** *For every integer* K *with* $0 \le K \le 1\,000\,001$*,*

$$R_K \;\le\; \alpha^*\, n_K^{1/6}.$$

Theorem 1.3 (attainment and optimality)**.** *$R_3 = \alpha^* n_3^{1/6}$, where $n_3 = 10$ and $B_3 = 5$. Consequently, if $a \in \mathbb{R}$ satisfies $R_K \le a\, n_K^{1/6}$ for every $K \le 1\,000\,001$, then $\alpha^* \le a$. Moreover $f(10) = 5$ and $\delta_{10} = R_3 = \alpha^* \cdot 10^{1/6}$ exactly, so no constant smaller than α^* satisfies the source's proposition on any range of n containing 10.*

Theorem 1.4 (the bound on δ_n)**.** *Let $N = n_{10^6} = 202\,642\,368\,741\,515\,820$. For every integer $n \le N$, every value admissible in (1) — and hence $f(n)$ itself — is at most $c\,n^{2/3} + \alpha^* n^{1/6}$; that is,*

$$\delta_n \;\le\; \alpha^*\, n^{1/6} \qquad \text{for every } n \le 2.02 \cdot 10^{17}.$$

Since $\alpha^* < 0.63$, Theorem 1.4 implies the source's proposition on the band it covers, which is 17 of the 23 decades asserted there; and by Theorem 1.3 the constant cannot be improved on any band containing $n = 10$, so on this range the question of the sharp constant is closed. The verification behind Theorems 1.2 and 1.4 uses exact integer arithmetic only (Section 4).

Finally, the erratum. We separate its two halves: a single exact value, which we prove, and a measurement of the printed script over its whole range, which is a computation about floating-point arithmetic and is reported as such.

Proposition 1.5 (an exact value the printed script misses)**.** *$n_{10^6} = 202\,642\,368\,741\,515\,820$ and $B_{10^6} = 303\,963\,552\,393$. In particular $n_{10^6} \ne 202\,642\,368\,741\,518\,368$, which is the value the script of Section 6 returns.*

1.4. **What is not claimed.** We do not reprove the source's Theorem 1: that $f(n)$ of (1) bounds the number of distinct fitted values is the source's combinatorial result, it is taken here as the definition of the object of study, and nothing below depends on it or establishes it. All our statements are about f, δ_n, n_K, B_K and R_K as defined in (1), (2) and (3).

We do not close the source's own range. Theorem 1.2 covers $K \le 10^6 + 1$ and Theorem 1.4 covers $n \le 2.02 \cdot 10^{17}$, against $K \le 10^8$ and $n \le 2 \cdot 10^{23}$ in [1]. The obstruction is the cost of computing 10^8 totients inside a verified evaluator, not the certificate, whose margin is ample throughout (Remark 8.1).

We do not prove that α^* is optimal over the source's whole range, only over the band of Theorem 1.2. The measurement of Section 6 puts the maximum of $R_K / n_K^{1/6}$ over $1000 \le K \le 10^8$ at 0.0934, nearly seven times below α^*, so optimality over the full range is extremely likely; but that is a numerical scan, not a theorem here.

We do not claim the erratum affects the source's conclusion. It does not, and Section 6 says why in detail.

2. Preliminaries

2.1. **Farey series, and the two sequences.** For $K \ge 1$ let F_K be the Farey series of order K: the reduced fractions a/b in $[0,1]$ with $1 \le b \le K$, listed in increasing order. Then

$$|F_K| = 1 + \sum_{j \le K} \varphi(j) = B_K, \qquad \sum_{a/b \in F_K} b = 1 + \sum_{j \le K} j\,\varphi(j) = n_K,$$

the extra 1 in each case accounting for $0/1$. These are entries A005728 and A240877 of the OEIS [6, 7]; the sum $\sum_{j \le K} j\varphi(j)$ alone is A011755 [8]. We use the same formulas for $K = 0$, where both empty sums vanish and $n_0 = B_0 = 1$; the sequences are then defined for all $K \ge 0$, n_K is strictly increasing, and the first values are

$$(n_K)_{K \ge 0} = 1, 2, 4, 10, 18, 38, 50, \ldots, \qquad (B_K)_{K \ge 0} = 1, 2, 3, 5, 7, 11, 13, \ldots$$

The pair (K, ℓ) in (1) is exactly a pair for which the constraint reads $n_{K-1} + \ell K \le n$ and the value reads $B_{K-1} + \ell$; we call such a pair *admissible for n*, so that $f(n) = \sup\{B_{K-1} + \ell : (K, \ell) \text{ admissible for } n\}$. Taking $\ell = \varphi(K)$ turns an admissible pair for n_K into the value B_K, which is why the sequence R_K controls δ_n along a subsequence and why the source interpolates between consecutive K.

2.2. **The normalised error and its first values.** Put $\rho_K = R_K/n_K^{1/6}$, the quantity the source's numerical check maximises. Its first values are, to ten places,

K	n_K	B_K	ρ_K
1	2	2	0.5358023194
2	4	3	0.6189983850
3	10	5	**0.6203305643**
4	18	7	0.5860173430
5	38	11	0.5680619125
6	50	13	0.5430339179

and ρ_K decays thereafter; the numerical scan of Section 6 puts $\max_{1000\le K\le 10^8}\rho_K$ at 0.0934306933, attained at $K = 1352$ — a measurement, not a statement proved here. The maximum is at $K = 3$ — proved below for $K \le 10^6 + 1$ (Theorems 4.2 and 5.1), and measured beyond that — and this is what makes the sharp constant explicit: $n_3 = 10$ and $B_3 = 5$ are small integers, so ρ_3 is the closed-form number α^* of (4), and the whole difficulty of the problem is the *proof* that no later index beats it.

3. The optimal constant

Proposition 3.1. *Let $\alpha^* = R_3/n_3^{1/6}$ as in (4). Then*

$$\alpha^* = 5\cdot 10^{-1/6} - \frac{3\sqrt{10}}{(4\pi^2)^{1/3}}, \qquad 0.6203305642 < \alpha^* < 0.6203305643.$$

Proof. Write $w = 10^{1/6}$, so $w^6 = 10$, $\sqrt{10} = w^3$ and $10^{2/3} = w^4$. Since $n_3 = 10$ and $B_3 = 5$,

$$\alpha^* = \frac{5 - c\,w^4}{w} = \frac{5}{w} - c\,w^3 = 5\cdot 10^{-1/6} - \frac{3\sqrt{10}}{(4\pi^2)^{1/3}},$$

the last step by $c = 3/(4\pi^2)^{1/3}$. For the enclosure, note that $x \mapsto x^6$ and $x \mapsto x^3$ are increasing on $[0,\infty)$, so the rational bounds

$$1.4677992676220 < w < 1.4677992676221, \qquad 0.8810515964898 < c < 0.8810515964899$$

follow from $1.4677992676220^6 < 10 < 1.4677992676221^6$ and, using $c^3 = 27/(4\pi^2)$, from

$$0.8810515964898^3\cdot 4\pi^2 < 27 < 0.8810515964899^3\cdot 4\pi^2;$$

the latter two comparisons need only the twenty-digit enclosure

$$3.14159265358979323846 < \pi < 3.14159265358979323847.$$

Substituting the four rational bounds into $\alpha^* = 5/w - c\,w^3$, in the directions that make each side extremal, gives the stated ten-digit enclosure. □

Proposition 3.2 (transcendence; outside the formal development)**.** *α^* is transcendental.*

Proof. Suppose α^* were algebraic. The algebraic numbers form a field containing $10^{-1/6}$, $\sqrt{10}$ and $4^{1/3}$, so

$$\pi^{-2/3} = \left(5\cdot 10^{-1/6} - \alpha^*\right)\cdot\frac{4^{1/3}}{3\sqrt{10}}$$

would be algebraic, hence so would its inverse cube π^2, hence so would π — and π is transcendental by Lindemann's theorem [4]. □

Proposition 3.2 is the one statement in this paper that is not machine-checked; Section 8 says why. It settles the shape of the answer: the sharp constant in the source's proposition is not an algebraic number that a rounding such as 0.63 approximates badly, but a transcendental one, and (4) is a closed form for it.

4. The integer certificate, and the sweep

4.1. **The reduction to natural-number arithmetic.** Fix K and write $n = n_K$, $B = B_K$. The inequality to be certified is $B \le c\,n^{2/3} + q\,n^{1/6}$ for a rational q, which is $R_K \le q\,n_K^{1/6}$. We take

$$q = \frac{31}{50} = 0.62,$$

a rational strictly between $\rho_2 = 0.61899\ldots$ and $\alpha^* = 0.62033\ldots$; with this choice the test *fails* at $K = 3$ and at no other index of the sweep, which is exactly right, since $K = 3$ is the maximiser and is handled separately and exactly.

The certificate for a single index consists of a natural number m and two comparisons of natural numbers:

$$m^6 \le n \cdot 2^{120} \qquad \text{and} \qquad 4P^2\big(B \cdot 52428800 - 31m\big)^3 \le 27\,n^2\,(10^{20})^2\,52428800^3, \tag{5}$$

where $P = 31415926535897932384$7 and $52428800 = 50 \cdot 2^{20}$ is the common denominator of $q = 31/50$ and of $b := m/2^{20}$. (The subtraction in (5) is in $\mathbb{N}$; when $B \cdot 52428800 \le 31m$ the conclusion below is immediate anyway.)

Lemma 4.1. *Let $n \ge 1$, B, m be natural numbers satisfying* (5)*. Then $B \le c\,n^{2/3} + \frac{31}{50}\,n^{1/6}$.*

Proof. The first comparison says $b^6 = (m/2^{20})^6 \le n = (n^{1/6})^6$, hence $b \le n^{1/6}$ by monotonicity of the sixth power on $[0,\infty)$; therefore $q\,b \le q\,n^{1/6}$ and it is enough to prove $B - q\,b \le c\,n^{2/3}$. Both sides are nonnegative in the nontrivial branch $B \cdot 52428800 > 31m$, so, again by monotonicity, it is enough to prove the cubed inequality $(B - qb)^3 \le c^3 n^2$. Now $c^3 = 27/(4\pi^2)$, so the cubed inequality reads $4\pi^2(B - qb)^3 \le 27n^2$. Multiplying by 52428800^3 clears the denominators of q and b, turning the left side into $4\pi^2(B \cdot 52428800 - 31m)^3$; and $\pi < P/10^{20}$, so the second comparison of (5) implies it. □

Three features of (5) deserve mention, because they are what make the check exact. First, nothing about the provenance of m is trusted: m is produced by an unverified Newton iteration for $\lfloor (n2^{120})^{1/6} \rfloor$, and the inequality $m^6 \le n2^{120}$ is part of what is checked, so a bad root makes the certificate fail rather than weakening it. Second, π enters only through the rational upper bound $P/10^{20}$, with twenty correct digits; the tightest of the 10^6 comparisons has its two sides within a relative $5 \cdot 10^{-9}$ of each other (Remark 8.1), whereas replacing π by $P/10^{20}$ perturbs the left side by a relative 10^{-20}, so π is nowhere near the binding constraint. Third, every quantity in (5) is a natural number — over the sweep the two sides of the second comparison never exceed 330 bits — so the check involves no floating-point arithmetic at all.

4.2. **The sweep.** The finite search is a single pass over K: starting from $(n_0, B_0) = (1, 1)$, for $K = 1, 2, \ldots, 1\,000\,001$ the totient $\varphi(K)$ is computed, the pair (n_K, B_K) is updated by (3), and the test

$$K = 3 \quad \text{or} \quad (5) \text{ holds at } (n_K, B_K, m_K)$$

is evaluated. All $1\,000\,001$ indices pass. This is the only exhaustive computation behind Theorem 1.2, and it is performed once, in exact integer arithmetic.

Theorem 4.2. *For every integer K with $0 \le K \le 1\,000\,001$,*

$$R_K \;\le\; \alpha^*\,n_K^{1/6}, \qquad \textit{equivalently} \qquad B_K \;\le\; c\,n_K^{2/3} + \alpha^*\,n_K^{1/6}.$$

Proof sketch. Three cases. For $K = 0$ we have $n_0 = B_0 = 1$ and the claim is $1 \le c + \alpha^*$, which follows from $c > 0.881$ and $\alpha^* > 0.620$ (Proposition 3.1). For $K = 3$ the claim is an equality by the definition (4) of α^*; see Theorem 5.1. For every other K in the range, the sweep certifies (5) at (n_K, B_K), so Lemma 4.1 gives $R_K \le \frac{31}{50} n_K^{1/6}$, and $\frac{31}{50} < \alpha^*$ by Proposition 3.1, so $R_K \le \alpha^* n_K^{1/6}$. The exhaustive part is the sweep over the $1\,000\,001$ indices; the passage from the two integer comparisons to the real inequality is Lemma 4.1, proved once for a general index. □

5. Attainment, optimality, and the bound on the error term

Theorem 5.1. *$R_3 = \alpha^* n_3^{1/6}$; that is, the bound of Theorem 4.2 is an equality at $K = 3$. Consequently, if $a \in \mathbb{R}$ satisfies $R_K \le a\, n_K^{1/6}$ for every $K \le 1\,000\,001$, then $\alpha^* \le a$: no constant smaller than α^* works on the band.*

Proof. The equality is the definition (4) of α^*, together with $n_3 = 10 > 0$, so that $n_3^{1/6} \neq 0$. For the second part, apply the hypothesis at $K = 3$ and use the equality: $\alpha^* n_3^{1/6} = R_3 \le a\, n_3^{1/6}$, and divide by $n_3^{1/6} > 0$. □

Optimality on a band of sequence indices is one thing; optimality in the source's proposition, which is a statement about all n in a range, is another, and it is the second that Theorem 1.3 asserts. It follows from the equality above together with the fact that $n = 10$ realises it.

Theorem 5.2. *Let $N = n_{10^6} = 202\,642\,368\,741\,515\,820$ and let $n \le N$ be a positive integer. Then for every pair (K, ℓ) admissible for n in the sense of (1),*

$$B_{K-1} + \ell \;\le\; c\, n^{2/3} + \alpha^*\, n^{1/6}.$$

Hence $f(n) \le c\, n^{2/3} + \alpha^* n^{1/6}$, i.e. $\delta_n \le \alpha^* n^{1/6}$, for every $n \le 2.02 \cdot 10^{17}$.

Proof sketch. Put $G(x) = c\, x^{2/3} + \alpha^* x^{1/6}$ for $x \ge 0$. Then G is concave (a nonnegative combination of the concave functions $x^{2/3}$ and $x^{1/6}$, the coefficients being positive by Proposition 3.1) and increasing. Let (K, ℓ) be admissible for n and write $K = J + 1$, so that $\ell \le \varphi(J + 1)$ and $n_J + \ell(J + 1) \le n$. From $n_J \le n \le N = n_{10^6}$ and the strict monotonicity of (n_K) we get $J \le 10^6$, so both J and $J + 1$ lie in the band of Theorem 4.2, and $B_J \le G(n_J)$, $B_{J+1} \le G(n_{J+1})$. Set $\lambda = \ell/\varphi(J + 1) \in [0, 1]$. The two interpolation identities

$$(1 - \lambda)n_J + \lambda n_{J+1} = n_J + \ell(J + 1), \qquad (1 - \lambda)B_J + \lambda B_{J+1} = B_J + \ell$$

hold by (3), since $n_{J+1} - n_J = (J + 1)\varphi(J + 1)$ and $B_{J+1} - B_J = \varphi(J + 1)$. Therefore

$$B_J + \ell = (1 - \lambda)B_J + \lambda B_{J+1} \le (1 - \lambda)G(n_J) + \lambda G(n_{J+1}) \le G\big(n_J + \ell(J + 1)\big) \le G(n),$$

the middle step by concavity of G and the last by monotonicity together with $n_J + \ell(J + 1) \le n$. Since $B_{K-1} + \ell = B_J + \ell$, this is the claim. The statement about $f(n)$ follows because $f(n)$ is the supremum of the left-hand sides over admissible pairs. □

The two interpolations in that proof are the source's argument for its proposition; they are reproved here rather than cited, in the slightly stronger form that bounds every admissible pair directly, so that no preliminary reduction to a single distinguished index is needed. The new content of Theorem 5.2 is that the constant is α^* rather than 0.63, and that the finite input is the exact certificate of Section 4 rather than a floating-point computation.

Corollary 5.3 (the constant is attained)**.** *The pair $(K, \ell) = (3, 2)$ is admissible for $n = 10$, its value is $B_2 + 2 = 5$, and $5 = c \cdot 10^{2/3} + \alpha^* \cdot 10^{1/6}$. Hence $f(10) = 5$, $\delta_{10} = \alpha^* \cdot 10^{1/6}$, and no constant $a < \alpha^*$ satisfies $\delta_n \le a\, n^{1/6}$ on any range of n containing 10.*

Proof. Admissibility: $K = 3 \ge 1$, $\ell = 2 = \varphi(3)$, and $n_2 + \ell K = 4 + 6 = 10 \le 10$. Its value is $B_{3-1} + 2 = B_2 + 2 = 3 + 2 = 5$. The identity $5 = c \cdot 10^{2/3} + \alpha^* \cdot 10^{1/6}$ is (4) rearranged, using $n_3 = 10$, $B_3 = 5$. So $f(10) \ge 5$; and $10 \le N$, so Theorem 5.2 gives $f(10) \le c \cdot 10^{2/3} + \alpha^* \cdot 10^{1/6} = 5$. Hence $f(10) = 5$ and $\delta_{10} = 5 - c \cdot 10^{2/3} = R_3 = \alpha^* \cdot 10^{1/6}$. If $a < \alpha^*$ then $\delta_{10} = \alpha^* \cdot 10^{1/6} > a \cdot 10^{1/6}$, so the bound with constant a fails at $n = 10$. □

Corollary 5.3 is what turns "α^* works" into "α^* is the constant": the source's proposition, read as an assertion about a range of n containing 10 — and its asserted range, all $n \le 2 \cdot 10^{23}$, contains 10 — cannot hold with any constant below α^*, and holds with α^* on the band of Theorem 5.2.

6. The printed verification script

The source's evidence for its proposition over $K \le 10^8$ is the following program, printed in full after the proof and reproduced here verbatim from the v1 e-print:

```
import numpy as np

def totients(K):
    """Return totient sequence phi(0)=0, phi(1), ..., phi(K)."""
    phi = np.arange(K + 1, dtype=np.uint64)

    for p in range(2, K + 1):
        if phi[p] == p:
            phi[p::p] -= phi[p::p] // p

    return phi
```

(6)

```
K = 100_000_000
Ks = np.arange(K + 1)

phi = totients(K)

nK = 1+np.cumsum(Ks*phi)
BK = 1+np.cumsum(phi)
c = 3/(4*np.pi**2)**(1/3)
RK = BK - c*nK**(2/3)

print(np.log10(nK[-1]), np.max(RK/nK**(1/6)))
```

In the line `nK = 1+np.cumsum(Ks*phi)` the array `Ks` has type `int64` and `phi` has type `uint64`. There is no common *integer* type for that pair in the numerical library used, so the product — and therefore the cumulative sum that reaches $2 \cdot 10^{23}$, and therefore $R_K = B_K - c\, n_K^{2/3}$, a difference of two quantities of size $3 \cdot 10^{15}$ — is computed in binary64. (Running (6) under NumPy 2.2.6 prints `float64` for the type of `Ks*phi`.) The companion line `BK = 1+np.cumsum(phi)` stays in `uint64` and is exact.

The first consequence is a statement about integers, and we prove it.

Proposition 6.1. $n_{10^6} = 202\,642\,368\,741\,515\,820$ *and* $B_{10^6} = 303\,963\,552\,393$. *In particular* $n_{10^6} \neq 202\,642\,368\,741\,518\,368$.

Proof sketch. The two values are computed by an exact evaluation of the recursion (3) for $K \le 10^6$, with φ evaluated by repeatedly stripping the least prime factor; that routine is proved equal to the Euler totient, so the evaluation is an exact integer computation and no approximation enters. The final inequality is a comparison of two integer literals. The second literal is the value that (6) prints at $K = 10^6$; that it does so is a run of (6), not part of the proof. □

The second consequence is a measurement of (6) over its own range. It is a computation about floating-point arithmetic, it is not part of the formal development, and we report it as such.

Remark 6.2 (the defect of (6); outside the formal development). The script (6) was run verbatim at its own $K = 10^8$ and each of its floating-point values was compared with the exact integer computed by block-exact accumulation. Conversion from binary64 to an integer is exact, so the differences in n_K below are exact integer differences; the exact R_K and $\rho_K = R_K / n_K^{1/6}$ at the probe indices were then

evaluated in 50-digit arithmetic from the exact integers.

K	$n_K^{\text{script}} - n_K$	R_K	R_K^{script}	ρ_K vs. ρ_K^{script}
10^3	0	1.139224	1.139224	0.046997, 0.046997
10^6	2 548	9.146410	9.144348	0.0119343, 0.0119316
10^7	8 108 664	61.62756	60.86719	0.0254285, 0.0251147
92 926 626	46 612 619 350	+16.2505	−480.500	+0.0021996, −0.0650382
10^8	9 552 344 498	−204.4249	−294.000	−0.026673, −0.038361

The worst absolute error of the script's n_K over $K \le 10^8$ is $46\,678\,656\,004 \approx 4.7 \cdot 10^{10}$, at $K = 92\,926\,707$; the worst relative error is $3.1 \cdot 10^{-13}$. In R_K the error reaches 89.6 at $K = 10^8$, which is 44% of the value there, and 496.8 near $K = 9.29 \cdot 10^7$. In the quantity the script actually maximises, the largest distortion the scan located is $|\rho_K^{\text{script}} - \rho_K| \approx 0.0673$, near $K = 9.29 \cdot 10^7$; at $K = 92\,926\,708$ its exact value is 0.067310, and at the probe index $K = 92\,926\,626$ of the table it is 0.067238, where the exact ρ_K is +0.00220 and the script returns −0.06504: the wrong sign, and thirty times the true magnitude. (The scan itself is only accurate to about 10^{-5} at that scale, so it locates the maximum but does not pin its index; the two values just quoted are exact.) For orientation, the largest tail value the script reports is 0.0934 at $K = 1352$, so the worst distortion is 72% of it.

Remark 6.3 (what this does and does not affect). The conclusion of the source's proposition is true, and we have proved a sharper form of it on 17 of its 23 decades (Theorem 5.2). The reason the defect is harmless there is structural: the maximum of ρ_K is attained at $K = 3$, where $n_3 = 10$ and $B_3 = 5$ are small integers and binary64 is exact — indeed (6) prints the maximum as 0.620330564254298, which is α^* to fifteen digits, and its argmax as $K = 3$ — while the whole tail is bounded by 0.0934, nearly seven times below the asserted 0.63. It is that factor of seven, and nothing in the script, that makes the script's answer safe. What the measurement shows is that the printed certificate does not establish the inequality at large K: at $K \approx 9.3 \cdot 10^7$ the number it computes for ρ_K has the wrong sign, and its worst error in ρ_K exceeds two thirds of the largest tail value it reports. This is an erratum about the certificate, not about the theorem.

7. Controls

The statements of Sections 4–5 are inequalities established in part by an exhaustive computation, so it is worth recording explicitly that the computation is not vacuous, that the constant cannot be lowered, that the exemption of $K = 3$ from the sweep is needed rather than convenient, and that the sequences being swept are the intended ones.

Proposition 7.1 (controls)**.** (i) (the sequences are the intended ones) *At $K = 3, 10, 1000$ the pairs (n_K, B_K) are*

$$(10, 5), \qquad (218, 33), \qquad (202\,870\,720,\ 304\,193),$$

which are the values of the OEIS entries A240877 and A005728 at those indices.

(ii) (the exemption of $K = 3$ is load-bearing) *The integer test* (5) *with $q = 31/50$ fails at (n_3, B_3) and passes at (n_2, B_2) and at (n_4, B_4).*

(iii) (the constant 31/50 used inside the sweep is not a valid global constant) $R_3 \le \frac{31}{50} n_3^{1/6}$ *is false.*

(iv) (no smaller constant works, e.g. the ten-digit truncation) $R_3 \le 0.6203305642\, n_3^{1/6}$ *is false.*

(v) (the $n^{1/6}$ term is needed) $B_3 \le c\, n_3^{2/3}$ *is false; that is, $c\, n^{2/3}$ alone already fails at $n = 10$.*

Proof. (i) is an exact evaluation of (3). (ii) is three evaluations of the two comparisons (5); the failure at $K = 3$ is expected, since $\rho_3 = \alpha^* > 31/50$, and the two neighbouring successes show that the exemption is not masking a defect of the test at the indices adjacent to it. (iii) and (iv) follow from Theorem 5.1 together with the enclosure of Proposition 3.1: $R_3 = \alpha^* n_3^{1/6}$ and $31/50 < 0.6203305642 < \alpha^*$. (v) follows from $c < 0.8810515964899$ and $10^{2/3} < w^4$ with $w < 1.4677992676221$, whence $c \cdot 10^{2/3} < 5$. □

Item (ii) is the one that matters for reading Theorem 4.2: the sweep is not a uniform check with a single constant. It verifies $R_K \le \frac{31}{50} n_K^{1/6}$ at every index except $K = 3$, where that inequality is genuinely

false, and the maximiser is then treated exactly, by the identity that defines α^*. Choosing q strictly between ρ_2 and α^* is what makes this split possible: the exempted index is the only one the test rejects.

8. Verification

Every statement above except Proposition 3.2 and the two remarks of Section 6 is machine-checked in Lean 4 [9] over the Mathlib library [10]; the development is six modules and contains no `sorry`. The only external inputs to the integer certificate are Mathlib's Euler totient and its twenty-digit enclosure of π. The analytic content — Lemma 4.1, the concavity and monotonicity of G, the two interpolation identities, and the passage from a verified sweep to Theorem 4.2 — is checked by the kernel alone, an axiom print returning `propext`, `Classical.choice` and `Quot.sound` and nothing else, the second entering only through library facts about π and about the least prime factor. Compiled evaluation is confined to four finite computations in exact integer arithmetic: the sweep over $1 \le K \le 1\,000\,001$ behind Theorems 4.2 and 5.2, the single evaluation of (n_{10^6}, B_{10^6}) behind Proposition 6.1, the single evaluation of (n_{1000}, B_{1000}) behind Proposition 7.1(i), and the three-index evaluation behind Proposition 7.1(ii); the axiom print of each certified theorem names the three standard axioms together with exactly those of these four that it uses. Three formulations differ from the printed ones, and we say which, since none of them is mathematics. First, the development does not define f as a supremum: what is proved is the pair of statements "every admissible value is at most $G(n)$" (Theorem 5.2) and "the value 5 is admissible at $n = 10$" (Corollary 5.3), from which $f(10) = 5$ is one line of order theory over a set of natural numbers. Second, Proposition 3.2 is proved by hand: Mathlib has the transcendence of π, but not a convenient interface to the field of algebraic numbers inside $\mathbb{R}$, which is all the argument needs; the closed form and the enclosure of Proposition 3.1 *are* machine-checked. Third, Remarks 6.2 and 6.3 are floating-point measurements and are asserted only as such; their integer half is Proposition 6.1, which is checked. Independently of the kernel, both sequences were recomputed to $K = 10^8$ by two implementations written separately from the development — a big-integer divisor sieve to $K = 10^7$ and a sieve with exact block accumulation to $K = 10^8$ — which agree with each other and with the development at every index compared. The source of the development is currently held in a private repository: repository reference to be added at submission.

Remark 8.1 (cost, and what was not run; outside the formal development). The following are timings, not theorems, and nothing above depends on them. The sweep of Theorem 4.2 is one compiled evaluation taking 225 s at 6.7 GB peak resident memory, the library import accounting for the memory; the whole development is about 50 CPU-minutes. The cost is dominated by the least-prime-factor totient, and was measured to grow by a factor of about 11 per decade of K: 20 s at $K = 10^5$ and 225 s at $K = 10^6$. Extrapolating at that rate, extending the band to $K \le 10^7$ (that is, to $n \le 2.03 \cdot 10^{20}$) is about 40 minutes in a single evaluation, and closing the source's own range $K \le 10^8$ is 7 to 8 CPU-hours; neither was run here. Nothing mathematical stands in the way: a separate unverified implementation of the same certificate clears $K \le 10^6 + 1$ with a worst ratio of 0.999999995 between the two sides of the second comparison in (5), and that relative margin decays only like $n^{-1/2}$, so the twenty digits of π never bind. The sweep is also splittable across evaluations, since the state at a split point is a pair of exact integers that can be supplied as literals. What would change the price is a verified linear totient sieve, which is an array-mutation correctness argument and is the real work.

On the reference list

Reference [1] was read here in its arXiv e-print, version 1 — both the compiled document and its LaTeX sources — and every quotation above is from that file; the internal label `prop:error_term_certificate` is the source's own, the proposition it labels is Proposition 9 there, and the numbering of any later version may differ. As checked against the abstract page on 3 September 2026, the e-print had a single version, dated 29 July 2026, with primary category stat.ML and cross-lists cs.LG and stat.ME, no journal reference, and was not withdrawn; the title and the order of the four authors are as printed in the entry below.

Reference [2] is cited for one comparison only and is used in no proof. It is, as far as a search of an 86 GB full-text corpus of arXiv could determine, the only other place where the constant $(4\pi^2)^{1/3}$ arises from the Farey-type asymptotics $\sum_{j\le K}\varphi(j)\sim 3K^2/\pi^2$: its lemma on Euler's totient records that asymptotic and its main theorem produces the constant $3\,(3/4\pi^2)^{1/3}\approx 1.27$ in a point-line incidence construction, both verified here in that paper's own e-print. That is a different constant for a different problem, and we record it only so that a reader who meets $(4\pi^2)^{1/3}$ in both places is not misled into thinking they are the same computation.

Reference [4] is cited for the transcendence of π in Proposition 3.2; its bibliographic data are those of the publisher's record for the article.

Reference [5] is cited once, for context: the classical degrees-of-freedom literature for shape-restricted regression measures effective dimension by an *expected* quantity under additive Gaussian errors; writing D for the divergence of the maximum-likelihood estimator, that paper shows for monotone regression that $\mathrm{E}_{\sigma,\mu}[D]\le Cn^{1/3}$, with C depending only on $(\mu(x_n)-\mu(x_1))/\sigma$. That is a different quantity from the deterministic worst-case count (1). Its bibliographic data were read off the article's own first page.

The three OEIS entries [6, 7, 8] were read in full — comments, formulas, references and cross-references — on 3 September 2026. None of them mentions $B_K - c\,n_K^{2/3}$, any normalisation by $n_K^{1/6}$, isotonic regression, or degrees of freedom, and a search of the OEIS index of constants for the decimal digits of α^* returns nothing.

References

[1] R. Rossellini, R. F. Barber, Z. Ren and J. A. Soloff, *An analysis of binary isotonic regression: degrees of freedom and implications for calibration*, arXiv:2607.27301v1 [stat.ML], 29 July 2026.

[2] M. Balko, A. Sheffer and R. Tang, *The constant of point-line incidence constructions*, arXiv:2210.04821v1 [math.CO], 10 October 2022.

[3] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press; the Farey series and the average order of φ.

[4] F. Lindemann, *Ueber die Zahl π*, Mathematische Annalen **20** (1882), 213–225, doi:10.1007/BF01446522.

[5] M. Meyer and M. Woodroofe, *On the degrees of freedom in shape-restricted regression*, The Annals of Statistics **28** (2000), no. 4, 1083–1104, doi:10.1214/aos/1015956708.

[6] OEIS Foundation Inc., *Entry A005728: Number of fractions in Farey series of order n*, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A005728`.

[7] OEIS Foundation Inc., *Entry A240877: Sum of the denominators of the Farey series of order n*, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A240877`.

[8] OEIS Foundation Inc., *Entry A011755:* $a(n)=\sum_{k=1}^{n} k\,\varphi(k)$, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A011755`.

[9] 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, doi:10.1007/978-3-030-79876-5_37.

[10] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, 367–381, doi:10.1145/3372885.3373824.