

Conjectures.io

Formal proof verified through Conjectures

ERDŐS PROBLEM 416(I): THE DOUBLING LAW FOR DISTINCT TOTIENT VALUES

Liam Kruer and Jensen Kohlmeyer

ABSTRACT. Let $V(x)$ count the distinct positive values of Euler's totient function that do not exceed x . The accepted submission proves $V(2x)/V(x) \rightarrow 2$ as real $x \rightarrow \infty$, answering part (i) of Erdős Problem 416. Its route is to count finite families of prime–core pairs, control missing values and repeated representations separately, and use uniform prime counting to balance the pair counts at the two endpoints. We give a complete proof of the finite counting and final limit reductions, explain the concrete record construction and analytic interfaces, and map those interfaces to the full accepted Lean source. The extensive sieve, prime-distribution, and geometric arguments are supplied by that source, rather than rederived in this exposition.

1. COUNTING DISTINCT TOTIENT VALUES

For a positive integer m , Euler's totient function is

$$\phi(m) = |\{1 \leq a \leq m : \gcd(a, m) = 1\}|.$$

Define

$$T(x) = \{n \in \mathbb{N} : 1 \leq n \leq x, \exists m \geq 1, \phi(m) = n\}, \quad V(x) = |T(x)|.$$

The preimage m is unrestricted: the cutoff applies to the value $\phi(m)$, not to m . Also, equal values are counted only once. For example, $\phi(3) = \phi(4) = \phi(6) = 2$, contributing a single value to $T(x)$ when $x \geq 2$.

Theorem 1.1 (Accepted target). *As x tends to infinity through the real numbers,*

$$\frac{V(2x)}{V(x)} \rightarrow 2.$$

Equivalently, for every $\eta > 0$ there exists X such that for all real $x \geq X$, $|V(2x)/V(x) - 2| < \eta$.

The denominator is eventually positive: $\phi(1) = 1$ gives $V(x) \geq 1$ for $x \geq 1$. This is an eventual statement over all large real cutoffs, not just a selected subsequence.

Date: 18 September 2026.

Key words and phrases. Euler totient function, distinct values, prime counting, Lean.

Submitted by JenW1N. Working proof exposition and source guide prepared with Codex assistance from the accepted Lean submission. The complete proof is the linked formal source; the longer construction and analytic lemmas are summarised here. This prose is not itself formally verified or a journal publication.

1.1. History and the exact part resolved. In his 1974 paper, printed page 201, Erdős defines this counting function with the letter f and asks whether $f(2x)/f(x)$ has a limit, observing that the limit, if it exists, must be two [1]. Thus the announcement has a directly documented 52-year history as of 2026. This is a documented publication date, not a claim about the first time the question was considered.

Ford's work gives a much more refined order-of-magnitude estimate for $V(x)$ but explicitly distinguishes it from the scaling-limit question [2]. A bound up to multiplicative constants cannot, by itself, give a limiting ratio: bounded multiplicative fluctuations can survive division. The accepted result establishes the doubling limit. It does not supply a full asymptotic equivalent for $V(x)$ or resolve all of Problem 416.

2. AN EXACT FINITE COUNTING INEQUALITY

The central elementary step separates two obstructions: values that have no retained representation, and values with several retained representations. It is useful independently of the specific totient construction.

Let P, T, T_0 be finite sets, with $T_0 \subseteq T$, and let $f : P \rightarrow T$. Put

$$P_0 = f^{-1}(T_0), \quad I = f(P), \quad I_0 = f(P_0) = I \cap T_0.$$

Define the missing-value and excess-representation counts

$$M = |T| - |I|, \quad M_0 = |T_0| - |I_0|, \quad E = |P| - |I|, \quad E_0 = |P_0| - |I_0|.$$

We have $0 \leq M_0 \leq M$. Also $0 \leq E_0 \leq E$: each nonempty fibre contributes its cardinality minus one, and P_0 retains exactly the fibres over I_0 .

Lemma 2.1 (Finite counting error). *With this notation,*

$$||T| - 2|T_0|| \leq ||P| - 2|P_0|| + M + E. \quad (1)$$

Proof. The definitions give the exact identity

$$|T| - 2|T_0| = (|P| - 2|P_0|) + (M - 2M_0) - (E - 2E_0).$$

Since $M_0 \in [0, M]$, we have $|M - 2M_0| \leq M$. Similarly $|E - 2E_0| \leq E$. The triangle inequality proves (1). $\square$

For the application, take $T = T(y)$ and $T_0 = T(y/2)$. The family $P = P_y$ consists of retained prime-core pairs, and $f = f_y$ maps a pair to its numerical totient value. Write

$$A_y = |P_y|, \quad B_y = |\{a \in P_y : f_y(a) \leq y/2\}|.$$

Then the three error terms to control are

$$\begin{aligned} M_y &= V(y) - |f_y(P_y)| && \text{(missing values),} \\ E_y &= A_y - |f_y(P_y)| && \text{(excess representations),} \\ D_y &= A_y - 2B_y && \text{(pair imbalance).} \end{aligned}$$

Lemma 2.1 yields

$$|V(y) - 2V(y/2)| \leq |D_y| + M_y + E_y. \quad (2)$$

Small error for representations alone would not suffice; both M_y and E_y must be controlled.

3. CONCRETE CORE RECORDS

The construction uses the totient identity

$$\phi(mp) = (p-1)\phi(m) \quad \text{when } p \text{ is prime and } p \nmid m.$$

It retains a finite set of numerical cores $b = \phi(m)$ and varies the top prime p . The map to values is

$$f(b, p) = b(p-1).$$

A record supplying a core consists of a bounded positive tail a and a decreasing list of distinct primes $p_1, \dots, p_L$, with all these primes larger than every prime factor of a . Thus

$$m = a \prod_{j=1}^L p_j, \quad b = \phi(a) \prod_{j=1}^L (p_j - 1) = \phi(m).$$

The source additionally imposes prime-gap, size, and geometric conditions on the list. In `LowerCoreRecord`, the tail is bounded by a fixed R , the last prime by a fixed cutoff, and the prime coordinates satisfy an explicit facet inequality. These are finite arithmetic data; a desired counting law is not a field of the record.

There may be several records for the same numerical core. The source's `CoreSelection`.`fullSelection` chooses a representative record for each core, and the pair family uses those numerical cores. For sufficiently large endpoints, the top prime is larger than the primes of its selected record. This ensures coprimality and makes every retained pair an actual totient value. The theorem `powerRawPairs_actual_eventually` supplies that assertion.

The retained value window has a lower cutoff $y^{99/100}$. It deliberately excludes very small values, whose total contribution is negligible. In fact, the elementary upper bound $V(z) \leq z$ and an eventual lower bound $V(y) \geq cy/\log y$ for some $c > 0$ give

$$0 \leq \frac{V(y^{99/100})}{V(y)} \leq \frac{\log y}{c y^{1/100}} \rightarrow 0.$$

The lower bound follows from the injection $p \mapsto p-1$ on primes, together with prime counting. No doubling law is used in this power-cutoff estimate.

4. THE ANALYTIC INTERFACES SUPPLIED BY THE PROOF

The substantial work in the accepted source proves that these finite families have the required coverage, collision, and prime-counting properties. We state their combined interface explicitly so the role of every input is visible.

Proposition 4.1 (Retained-family estimates). *For every $0 < \varepsilon < 1/2$, there is a family of finite sets P_y and maps $f_y : P_y \rightarrow T(y)$ such that, eventually as real $y \rightarrow \infty$,*

$$M_y \leq \varepsilon V(y) + V(y^{99/100}), \tag{3}$$

$$E_y = o(V(y)), \tag{4}$$

$$D_y = o(A_y). \tag{5}$$

The family may depend on ε . Its auxiliary tail and terminal cutoffs are fixed before the endpoint tends to infinity.

This proposition summarizes proved declarations in the submitted file. The detailed analytic derivation is in that file; the following paragraphs identify the ingredients rather than claim a separate short proof of the proposition.

4.1. Coverage. The source first restricts to values above the power cutoff, then controls exceptional families associated with prime normality, large square factors, geometric facets, concentration, terminal primes, and residual factors. Outside the excluded families, an actual totient preimage supplies an admissible lower-core record. It is therefore represented in the selected pair family. The resulting estimate is `exists_powerRawPairs_fullSelection_coverage`, giving exactly (3). The existential preimage remains unrestricted; it is not silently bounded by the value cutoff.

4.2. Collisions. Even after choosing one record per core, different pairs may produce the same integer $b(p-1)$. The formal argument maps these collisions into structured tail and prime data, bounds the resulting families, and proves `powerRawPairs_collisions_negligible`. The excess count E_y is bounded by the number of pair records belonging to nonsingleton fibres. Hence the collision estimate implies (4). This step is what permits the passage from prime-pair counts to *distinct* totient values.

4.3. Uniform prime counting. The selected cores are subpower in the endpoint: their logarithms are bounded by a function $G(y) = o(\log y)$, uniformly across the finite family. Uniform prime estimates allow the same selected cores at endpoint y to be counted both up to y and up to $y/2$. The source's `power_corePairs_count_asymptotic` gives a common positive normalizing mass $H(y)$ with

$$A_y/H(y) \rightarrow 1, \quad B_y/H(y) \rightarrow 1/2.$$

Consequently $D_y/A_y \rightarrow 0$, yielding (5). The normalization is part of the pair-counting argument; it is not asserted to be a full asymptotic formula for $V(y)$.

The source contains the supporting prime number theorem, Mertens, and sieve developments, including attributed ports from PrimeNumberTheoremAnd [3]. Intermediate reductions with analytic hypotheses are followed by proofs supplying those hypotheses. The final target has no residual analytic assumption.

5. FROM THE ESTIMATES TO THE DOUBLING LAW

We now give the full final deduction from Proposition 4.1. This also explains why the normalization $o(A_y)$ in (5) is sufficient. Since $f_y(P_y) \subseteq T(y)$,

$$A_y = |f_y(P_y)| + E_y \leq V(y) + E_y.$$

By (4), eventually $A_y \leq 2V(y)$. Therefore $D_y = o(A_y)$ implies $D_y = o(V(y))$. Combining (2)–(5) with the power-cutoff estimate gives, for each fixed ε ,

$$|V(y) - 2V(y/2)| \leq \varepsilon V(y) + o(V(y)).$$

Given any $\delta > 0$, first choose $0 < \varepsilon < \min(\delta/2, 1/2)$ and its corresponding family, then take y sufficiently large for the remaining error to be at most $(\delta - \varepsilon)V(y)$. Hence

$$\forall \delta > 0, \quad |V(2x) - 2V(x)| \leq \delta V(2x) \quad \text{eventually.} \quad (6)$$

The auxiliary family changes when the requested accuracy changes; the counting function V does not.

Lemma 5.1 (Relative error controls the quotient). *If $v > 0$, $w \geq 0$, $0 \leq \delta \leq 1/2$, and $|w - 2v| \leq \delta w$, then $|w/v - 2| \leq 4\delta$.*

Proof. We have $(1 - \delta)w \leq 2v$, so $w \leq 4v$. Dividing the given error bound by v now gives

$$|w/v - 2| \leq \delta w/v \leq 4\delta.$$

□

For any requested $\eta > 0$, use $\delta = \min(1/2, \eta/8)$ in (6), and apply Lemma 5.1 with $v = V(x)$ and $w = V(2x)$. The resulting error is at most $4\delta < \eta$. This proves Theorem 1.1. The argument derives a bound on $V(2x)/V(x)$ from the relative error; it does not assume such a bound as an unproved regularity property.

6. FORMAL SOURCE GUIDE AND VERIFICATION

The full development is available as the [accepted Lean proof](#), with its [public verification and review record](#). The following line numbers refer to the exact accepted file.

Source declaration or component	Line
Exact count V, and membership equivalence	33
<code>finite_counting_error</code>	45376
<code>quotient_error_of_relative_doubling_error</code>	52168
<code>doubling_of_finite_counting_contracts</code>	52214
<code>LowerCoreRecord</code>	62597
<code>powerRawPairs_actual_eventually</code>	63447
<code>power_corePairs_count_asymptotic</code>	63482
<code>powerRawPairs_collisions_negligible</code>	63919
<code>exists_powerRawPairs_fullSelection_coverage</code>	64294
<code>doubling_limit</code>	64526
<code>Final Bounty.target</code>	64567

The [machine-readable verification report](#) is also public.

The public source retrieved on 18 September 2026 is byte-identical to the accepted local file, with SHA-256

`af5e0facbfca599610dc80e9c1a484c5fa50c8ce23a7aef409ff6ff32e751a1`.

The matching production record reports successful statement comparison and Lean kernel verification, and the manual review is approved. No fresh Lean replay or independent-kernel run was performed in preparing this paper. The permitted axioms in the record are propositional extensionality, quotient soundness, and classical choice. The public formal result, not this prose, is the object that the kernel accepted.

The manuscript credit follows the announcement series; the public submission credit is JenW1N. Supporting libraries and their contributors retain their separate attribution and licence notices in the complete file. This source guide makes no claim of independent verification of every analytic lemma, journal acceptance, or certified worldwide priority.

REFERENCES

- [1] P. Erdős, *Remarks on some problems in number theory*, Math. Balkanica **4** (1974), 197–202, especially p. 201. [Original paper](#).
- [2] K. Ford, *The distribution of totients*, Ramanujan J. **2** (1998), 67–151; [revised arXiv version](#), especially its discussion on p. 3.
- [3] PrimeNumberTheoremAnd contributors, [PrimeNumberTheoremAnd](#), supporting formal number theory; attribution retained in the submission.
- [4] Conjectures.io, [Erdős 416\(i\): accepted submission and public review](#), submitted by JenW1N; accessed 18 September 2026.
- [5] T. F. Bloom, [Erdős Problem 416](#), problem statement and historical references.