Forced prime squares and a hundred-billion-digit lower bound in Carmichael's totient conjecture

Tyler Hughes [co-author]

Third draft

AI-written draft for author review. The authors must check the mathematics, sources and attribution before submission.

Abstract

Euler's totient function can take the same value at several integers. Carmichael conjectured that every value it takes has at least two preimages. Suppose instead that only x maps to m. Any way of replacing a divisor of x while preserving its totient must then fail. A switching rule due to Carmichael and Klee, in Ford's formulation, explains why each failure forces a prime square into x. The exact exponent of three divides this process into two cases. An independently rechecked finite certificate gives $x > 10^{100,000,160,994}$, compared with Ford's $x > 10^{10,003,700,000}$. A reciprocal-sum estimate for the forced primes proves that the value m exceeds the same power of ten. The same divisors give an upper bound on the lower limiting fraction of singleton values among totients. Lean checks the forcing rules and a smaller explicit certificate giving $x > 10^{10,117}$. We also treat Ford's separate question about values whose every preimage is divisible by ten. Here a switch can force one of several primes to divide an input. An externally checked tree of these alternatives gives $m \geq 10^{60}$.

2020 Mathematics Subject Classification. 11A25, 11N64, 11Y11. **Keywords.** Euler totient, inverse totient, primality certificate, Carmichael conjecture.

1 What a single preimage would have to contain

Euler's function $\varphi(n)$ counts the integers from 1 through n coprime to n. For example, $\varphi(36) = \varphi(42) = 12$. In passing from $36 = 6 \cdot 6$ to $42 = 7 \cdot 6$, the first factor 6 adds no new prime to the second. It therefore multiplies the totient by 6. The new prime 7 contributes the same factor, $7 - 1$.

Write $A(m) = \#\{n \geq 1 : \varphi(n) = m\}$. Carmichael's conjecture says $A(m) \neq 1$ for every positive m [2, 3]. In the event $A(m) = 1$, call the unique *input* x and the *value* $m = \varphi(x)$. For a real X, let $V(X)$ count the distinct totients at most X, and let $V_1(X)$ count those with one preimage.

Theorem 1. *Suppose $A(m) = 1$ and $\varphi(x) = m$. The finite prime certificates give*

$$x > 10^D, \qquad m > 10^D, \qquad D = 100{,}000{,}160{,}994.$$

They also give $\liminf_{X\to\infty} V_1(X)/V(X) < 10^{-50,000,080,497}$. A smaller certificate checked inside Lean gives $x > 10^{10,117}$.

Theorem 2. *Suppose $A(m) > 0$ and $10 \mid x$ for every x with $\varphi(x) = m$. Then $m \geq 10^{60}$, by a finite external descent. The explicit lower bound checked inside Lean is $m \geq 32{,}000$.*

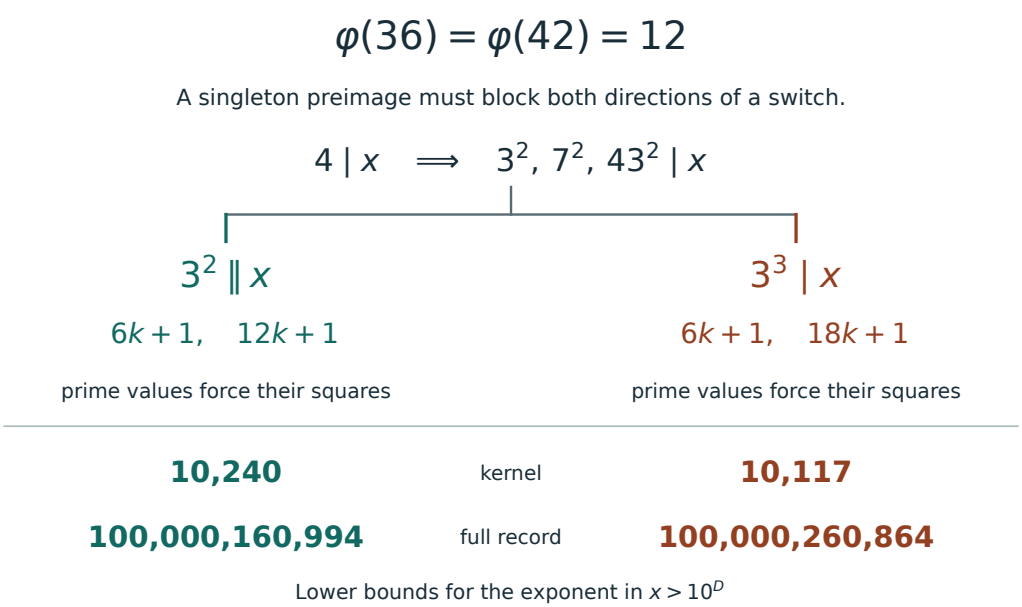


Figure 1. The two cases for the exponent of three give different ways to create the next prime. Each row at the foot reads across the two cases. Its smaller entry works for every singleton preimage. The full record was checked by separate integer-arithmetic software.

The switch in seven lines. The equation used above can be checked directly from the definition.

```
from math import gcd
def phi(n):
    return sum(gcd(a, n) == 1 for a in range(1, n + 1))
p, z = 7, 6
x, y = (p - 1) * z, p * z
assert (x, y) == (36, 42)
assert phi(x) == phi(y) == 12
```

The switch gives a divisibility rule in Section 2. The arithmetic consequences in Section 3 use only the divisors it supplies. Section 4 specifies what the finite checks established for each theorem. Sections 5 and 6 describe the formal results and the remaining infinitude question.

2 Switching a divisor

For $n \geq 1$, put $\mathrm{rad}(n) = \prod_{p \mid n} p$, with $\mathrm{rad}(1) = 1$, and write $\nu_p(n)$ for the exponent of p in n. A divisor d of $x = dw$ is *unitary* if $\gcd(d, w) = 1$. The notation $p^j \parallel x$ means $\nu_p(x) = j$. Two facts from the prime factorization of φ will recur. If $e \mid w/\mathrm{rad}(w)$, the numbers w/e and w have the same prime divisors and $\varphi(w/e) = \varphi(w)/e$. If every prime of u divides z, then $\varphi(uz) = u\varphi(z)$.

Lemma 3 (Ford's switching rule [5, Lemma 7.2])**.** *Let x be a singleton preimage and let $x = dw$ with d unitary. Suppose $\mathrm{rad}(\varphi(d)) \mid x$, and take $e \mid w/\mathrm{rad}(w)$. If $P = 1 + e\varphi(d)$ is prime, then $P^2 \mid x$.*

Proof. If $P \nmid x$, replace dw by $P(w/e)$. The new integer contains P and has totient $(P-1)\varphi(w/e) = \varphi(d)\varphi(w) = \varphi(x)$. If $P \parallel x$, write $x = Pz$. Every prime of $P - 1 = e\varphi(d)$ divides x and is smaller than P, so divides z. Now $(P-1)z$ is a different input with totient $(P-1)\varphi(z) = \varphi(x)$. Both possibilities contradict uniqueness. □

If x were odd, $2x$ would have its totient. If $2 \parallel x$, $x/2$ would have its totient. Hence $4 \mid x$. In Lemma 3, set $d = 1$. Taking $e = 2$ forces $3^2 \mid x$. Then $e = 6$ forces $7^2 \mid x$, and $e = 42$ forces $43^2 \mid x$. The evident next choice gives $1 + 2 \cdot 3 \cdot 7 \cdot 43 = 1807 = 13 \cdot 139$. To continue, use the exponent of 3. Carmichael introduced this split, and Ford gives both recurrences below in his explicit search [3, 7, 5].

Case	Condition on x	Initial divisor b	Candidates P
I	$3^2 \parallel x$	36	$6k+1,\ 12k+1$
II	$3^3 \mid x$	108	$6k+1,\ 18k+1$

Here k is a squarefree product of primes at least 7 whose squares already divide x, including the empty product 1. For $6k+1$ use $d = 1, e = 6k$. In Case I the factor $d = 9$ is unitary and $\varphi(d) = 6$. Taking $e = 2k$ gives the second form $12k + 1$. In Case II the extra power of three makes $e = 18k$ available with $d = 1$. Thus a prime candidate supplies its square for the next step. Starting at $k = 1$ gives $7, 13$ in Case I and $7, 19$ in Case II. For example, Case I then permits $79 = 6 \cdot 13 + 1$, and Case II permits $127 = 18 \cdot 7 + 1$.

3 Two consequences of the forced divisor

Let S_i contain the distinct certified primes in Case i. The forcing rules give $d_i = b_i \prod_{p \in S_i} p^2 \mid x$ in that case. The finite record in Section 4 supplies $d_i > 10^{D_i}$, where

$$D_{\mathrm{I}} = 100{,}000{,}160{,}994, \qquad D_{\mathrm{II}} = 100{,}000{,}260{,}864.$$

Write $D = D_{\mathrm{I}}$. A small margin in the binary conversion will matter for both corollaries. Set $u = 1{,}936{,}274,\ v = 6{,}432{,}163$, so that $10^u < 2^v$. For the Case I bit sum $B = 332{,}193{,}344{,}302$,

$$6Bu - (6D + 5)v = 3{,}541 > 0.$$

Consequently $d_{\mathrm{I}} \geq 2^B > 10^{D+5/6} > 6 \cdot 10^D$, since $6^6 < 10^5$. Case II gives $d_{\mathrm{II}} > 10^{D+2}$ directly.

3.1 Bounding the totient value

Divisibility $d_i \mid x$ implies $\varphi(d_i) \mid \varphi(x)$, by the prime-power formula. We will show that $\varphi(d_i) > d_i/6$. Let T_i consist of the smallest 1,000 primes in the certified base of Case i and set $t_i = \sum_{s \in T_i} 1/s$. Summing $\lceil 10^5/s \rceil$ over these seeds gives 27,595 and 23,756 respectively, so $t_i < 1/3$. The largest seed in each case exceeds 10^8. There are fewer than 10^6 further base entries, whose reciprocal sum is therefore less than $1/100$.

Every additional prime has the form $P = ck + 1$, where k is a product of one to four distinct seeds. For products of j seeds, the sum of $1/k$ is at most t_i^j, by expanding the power. Summing over both coefficients c and allowing every such product gives

$$\sum_{P \text{ additional}} \frac{1}{P} < \left(\frac{1}{6} + \frac{1}{12}\right) \sum_{j=1}^{4} t_i^j < \frac{1}{4} \frac{t_i}{1 - t_i} < \frac{1}{8}.$$

The same upper bound applies when the second coefficient is 18. Thus $\sum_{p \in S_i} 1/p < 1/3 + 1/100 + 1/8 < 1/2$. Using $\prod(1 - u_j) \geq 1 - \sum u_j$ for $0 \leq u_j \leq 1$, we obtain

$$\frac{\varphi(d_i)}{d_i} = \frac{1}{3} \prod_{p \in S_i} (1 - 1/p) > \frac{1}{6}.$$

The margin $d_i > 6 \cdot 10^D$ now gives $m \geq \varphi(d_i) > 10^D$.

3.2 Counting singleton values among all totients

Write $a_i = 6\prod_{p\in S_i} p$, so $d_{\mathrm{I}} = a_{\mathrm{I}}^2$ and $d_{\mathrm{II}} = 3a_{\mathrm{II}}^2$. Both a_i exceed $2 \cdot 10^{50,000,080,497}$. Indeed D is even, $d_{\mathrm{I}} > 6 \cdot 10^D$ and $d_{\mathrm{II}}/3 > 10^{D+2}/3 > 4 \cdot 10^D$.

Let $V(X; a^2)$ count totients $n \le X$ whose every preimage is a multiple of a^2, where $a \ge 2$. Ford proves $V(X; a^2) \le V(X/a)$ [5, Lemmas 7.4–7.5]. To see the injection, take $X/a < n \le X$ and any preimage y. If $a^2 \nmid y$, leave n unchanged. Otherwise write $y = a^t r$ with $a \nmid r$. Then $\varphi(ar) = n/a^{t-1}$, and $a^2 \nmid ar$. The resulting totients are distinct because two distinct members of $(X/a, X]$ cannot have ratio an integral power of a. Removing their number from $V(X)$ gives the inequality.

Put $A = \min(a_{\mathrm{I}}, a_{\mathrm{II}})$. Every singleton preimage belongs to one of the two cases, so $V_1(X) \le V(X/a_{\mathrm{I}}) + V(X/a_{\mathrm{II}}) \le 2V(X/A)$. If $V_1(X)/V(X)$ eventually exceeded a constant $c > 2/A$, iteration of $V(X) \le (2/c)V(X/A)$ would give $V(X) = O(X^{\log(2/c)/\log A})$, with exponent less than one. Every prime $p \le X + 1$ supplies the distinct totient $p - 1$, contradicting the prime number theorem. Thus the lower limit is at most $2/A < 10^{-50,000,080,497}$.

4 The finite records

4.1 From a prime witness to a decimal exponent

For each candidate $P = ck + 1$ the construction supplies a complete factorization of $P - 1$. For each distinct prime $q \mid P - 1$, the certificate supplies a residue a_q satisfying

$$a_q^{P-1} \equiv 1 \pmod P, \qquad a_q^{(P-1)/q} \not\equiv 1 \pmod P.$$

These Lucas tests certify primality [1]. Indeed a_q is a unit. Its order divides $P - 1$ and contains the full power of q in $P - 1$. Every such order divides $\varphi(P)$. Consequently $P - 1 \mid \varphi(P)$, which forces P prime. Each derived entry also needs a valid expression $ck + 1$ with its distinct factors of k already forced in the same case. This second check connects primality to divisibility of x.

Ford's search already used a base of 1,000 primes per case, an initial batch with one to three seed factors, and subsequent batches using four seeds, with one fixed smallest index and three larger indices. He also used the Lucas test above, following Schlafly and Wagon [9, 5]. The larger computation reported here takes the smallest 1,000 primes of its own certified base and runs the same batch design to the exponents in Table 1. Its final batch stops partway through. Each prime is counted once within a case. The valuations at two and three identify the coefficient c, unique factorization identifies the seed indices, and the checker rejects indices out of order and primes already in the base. An entry may occur in both cases, where it has different forcing hypotheses.

For a chunk with prime product G, the recorded integer $\lfloor \log_2(G^2) \rfloor$ is the bit length of G^2 minus one. Partitioning the primes into chunks and summing these integers with $\lfloor \log_2 b_i \rfloor$ gives $d_i \ge 2^B$. The strict integer inequality $10^{1,936,274} < 2^{6,432,163}$ gives $2^B > 10^{D_i}$. A missed prime would leave the divisor valid. Counting the same square twice would threaten the bound, which is why membership and distinctness matter.

Across both cases the record contains 1,965,316,556 entries, including the bases. A separately written checker replayed every prime derivation, Lucas witness and chunk sum across the bases and 699 batch files. It rejected all five deliberately damaged inputs. A separate regeneration reproduced the initial batch in each case byte for byte. The retained records include hashes, batch summaries and regeneration code. The disk holding the complete set of per-prime batch files was later retired. The accompanying quick check therefore reproduces the table from the recorded integers. Repeating the full primality replay requires regenerating those files.

Table 1. Exact accounting for the forced divisor. The first three bit rows are added to form B. The last row is $\lfloor B(1{,}936{,}274)/(6{,}432{,}163)\rfloor$. The candidate count covers the batches following the construction of the base.

	Case I	Case II
Batches, including the initial one	171	179
Candidates tested	43,866,971,434	45,276,159,886
Primes in the base	155,792	117,842
Additional primes	1,120,013,445	845,029,477
Distinct within the case	1,120,169,237	845,147,319
$\lfloor \log_2 b_i \rfloor$	5	6
Bits from products of base primes squared	37,637,307	37,045,809
Bits from products of additional primes squared	332,155,706,990	332,156,630,246
B, sum of integer lower bounds	332,193,344,302	332,193,676,061
D_i, certified decimal exponent	100,000,160,994	100,000,260,864

4.2 How much a candidate contributed

A prime P adds $2\log_{10} P$ to the decimal exponent of the divisor. For a simple yield model, assume the factors of k give uniformly distributed units mod each prime $q \geq 5$, with independent local survival corrections. For $c = 6, 12, 18$, exactly one of the $q-1$ units makes $ck+1 = 0 \bmod q$. The ratio of the surviving fraction to $1-1/q$ is $q(q-2)/(q-1)^2$. Parity and divisibility by three together supply the factor 3. Thus the predicted multiplier of the ordinary prime probability $1/\log P$ is $4C_2$, where $C_2 = \prod_{q>2,\ q \text{ prime}}(1-(q-1)^{-2})$. The expected decimal exponent per candidate is $8C_2/\log 10$, about 2.293637.

The seeds are far from uniform at some primes. In Case II, for example, 52.3% of them equal 1 mod 7. For the form $6k+1$, the residue $k = 1 \bmod 7$ is exactly the one forbidden by 7. For each $5 \leq q \leq 541$ prime, the finite model counts the residues of products at the actual seed indices. Scanning backwards, adding a seed of residue s sends a count at size j, residue v, to size $j+1$, residue $sv \bmod q$, retaining the choices that omit the seed. The model multiplies the resulting local survival factors, and uses the uniform product for the tail. Its count includes the whole final index batch. The measured yield uses only candidates processed up to the stopping point.

Decimal exponent per candidate	Case I	Case II
Predicted from seed residues	2.279324	2.209519
Measured from the chunk bit sums	2.279365	2.208427

Here the measured numerator is the additional bit sum times 1,936,274/6,432,163. These predictions describe the search yield. Both the divisor and its lower bound follow from integer certificates alone.

4.3 When every preimage is divisible by ten

For Theorem 2, multiplicity is unrestricted. Fix a value m in the image of φ whose every preimage is divisible by 10, and choose any preimage x. As before, $2 \parallel x$ would give the odd preimage $x/2$. If $5 \parallel x$, write $x = 5w$, where $4 \mid w$. Then $\varphi(4w) = 4\varphi(w) = m$ although $5 \nmid 4w$. Thus $100 \mid x$ and $40 = \varphi(100) \mid m$.

Here a switch need only produce a preimage missing a 2 or a 5. For a unitary d divisible by 2, write $x = dw$, take $e \mid w/\operatorname{rad}(w)$, and choose $t \geq 0$ with $5^t \mid e\varphi(d)$ and either $t = 0$ or $5 \nmid d$. Suppose that distinct odd primes P_i satisfy $\prod_i (P_i - 1) = e\varphi(d)/5^t$. If all P_i were absent from x,

the odd integer $5^t(w/e)\prod_i P_i$ would have totient m. Indeed w/e keeps the primes of w, so the factor 5^t multiplies its totient by 5^t whenever $t > 0$. Each new prime then contributes $P_i - 1$. Thus at least one $P_i \mid x$. With $5 \mid d$, the same argument uses 2^t in place of 5^t, requires $t = 0$ or $2 \nmid d$, and chooses odd $P_i \neq 5$. The constructed preimage then omits 5. Keeping a single P_i gives an ordinary prime-forcing step. Several factors give a disjunction.

For example, on the branch $2^{10} \parallel x$, $5^6 \parallel x$, take $d = 2^{10}$, $e = 1$, $t = 0$. The identity $\varphi(d) = 512 = (3-1)(257-1)$ forces $3 \mid x$ or $257 \mid x$. To make such alternatives finite, assume $m < 10^{60}$. Every root is now an exact valuation pair $(a, b) = (\nu_2(x), \nu_5(x))$ with $a, b \geq 2$ and

$$\varphi(2^a 5^b) = 2^{a+1}5^{b-1} < 10^{60}.$$

At a node, a forced divisor $n \mid x$ implies $\varphi(n) \mid m$. A node closes as soon as $\varphi(n) \geq 10^{60}$. Otherwise its children cover all the possible primes in a disjunction, or split a valuation $p^v \mid x$ into $p^v \parallel x$ and $p^{v+1} \mid x$.

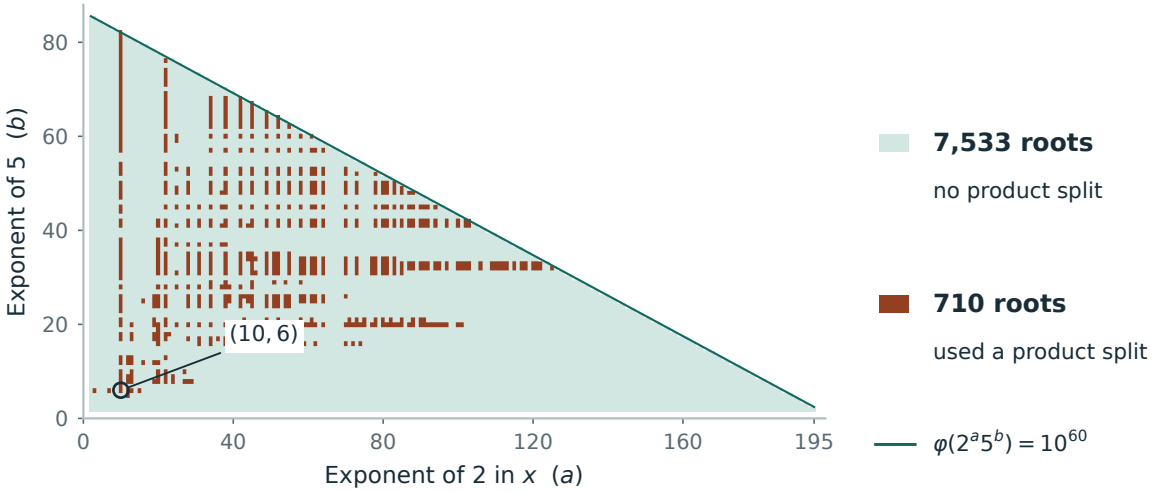


Figure 2. Each cell is one possible pair of initial valuations below the totient cutoff. The darker cells mark roots at which some descendant used a product $\prod(P_i - 1)$ with more than one prime. Light cells used no such split, though 1,366 of them branched on a prime valuation. The ring marks the worked example. White space above the curve already exceeds the cutoff.

All 8,243 roots close. The recorded trees have 21,329 nodes, 59,755 forced-prime entries, 893 product splits and 14,786 closed leaves. Two independent implementations replayed the derivations, branching and leaf inequalities. Their common omission is that both accept a file with all the correct valuation labels but an empty node list at every label. The packaged check requires exactly one initial state at each required pair and a path from it to every node. On the archived forest it verifies these conditions and reruns the original arithmetic and leaf checks. It rejects the empty-tree example. This exhausts the possible inputs below the cutoff and proves Theorem 2. The smaller bound $m > 10^{12}$ also has a separate check using enumeration of leaf multiples without the product-splitting rule.

5 What Lean checks

The library `CarmichaelTotient` in Lean 4 proves the switching rule, both case recursions, the multi-witness primality criterion, and the conversion from bit sums to powers of ten. For a concrete check, the library instantiates the two case certificates with 81 and 80 prime entries. They supply 34,024 and 33,616 bits. The smaller integer comparison $10^{31} < 2^{103}$ gives exponents 10,240 and 10,117. The resulting theorem $10^{10,117} < x$ uses a certificate reduced by the kernel. For the larger data, `carmichael_bound_strict_from_certificate` proves the implication from a valid pair of certificates to the bound. The external checks in Section 4 supply the large instances.

For the common-divisor problem, Lean checks a concrete proof of $m \geq 32{,}000$ and the soundness of `K10RungCertificate`. The 10^{60} forest was checked externally. The pinned library audit covered 159 declarations with the usual axioms of propositional extensionality, choice and quotient soundness. It records zero proof placeholders and no use of `native_decide`.

Three structural results in the library place the computations in a wider setting. Klee shows that a singleton input with no proper singleton divisor has $4 \parallel x$ and no prime divisor of the form 2^h+1, $h \geq 2$ [7]. Pomerance gives a sufficient condition for an individual x to be a singleton input. Require $p^2 \mid x$ whenever p is prime and $p-1 \mid \varphi(x)$. All prime divisors of any competing input then occur in x. Comparing the prime valuations of their totients in descending prime order forces equal exponents [8]. Finally, Erdős's construction preserves a whole fiber. If $k, m \geq 1$, $q > km+1$ and $P = kq+1$ are prime, and $jq+1$ is composite for each other divisor $j \mid km$, then $\varphi^{-1}((P-1)m) = P\,\varphi^{-1}(m)$. The library formalizes this algebraic step [4].

6 Historical scope and open questions

Carmichael withdrew his first argument and gave forced divisors in its place [2, 3]. Klee refined the cases, and Schlafly and Wagon used certificates for large primes [7, 9]. Ford gave the input bound $x > 10^{10,003,700,000}$ and the value bound $m \geq 10^{10^{10}}$ [5]. The exponent of the former is the direct comparison for Table 1. These are primary-source comparisons. The earlier literature search had incomplete access to several historical sources, so it cannot establish an exhaustive priority claim.

The singleton conjecture remains open. An infinite set of forced primes in each of the two cases would suffice, since no positive integer contains infinitely many prime factors. Adjoining 2 and 3 to either generated set gives a set closed under taking prime factors of $p-1$. This set omits 5. Ford shows that such a set has a convergent sum of prime reciprocals [6]. This is consistent with a finite set and with an infinite, sparse one. The observed yield measures a finite range and does not decide which occurs. For the separate preimage-divisor question, the calculation leaves open whether any totient has all of its preimages divisible by ten.

7 Data and reproducibility

The companion package identifies the two earlier manuscripts, the source records, the prime-file manifest and the Lean revision. Its checker executes the displayed Python example verbatim, finds fresh Lucas witnesses for both seed sets, sums the prime batches, and checks the exponent transfers. The complete compressed descent forest and its mathematical checker accompany the paper. Scripts for both figures use the archived data. A complete independent replay of the large prime certificate additionally requires regenerated per-prime files. AI systems assisted with the historical searches, software, independent computational checks and preparation of the drafts.

References

[1] J. Brillhart, D. H. Lehmer, and J. L. Selfridge, *New primality criteria and factorizations of* $2^m \pm 1$, Math. Comp. **29** (1975), no. 130, 620–647.

[2] R. D. Carmichael, *On Euler's ϕ-function*, Bull. Amer. Math. Soc. **13** (1907), no. 5, 241–243.

[3] R. D. Carmichael, *Note on Euler's φ-function*, Bull. Amer. Math. Soc. **28** (1922), no. 3, 109–110.

[4] P. Erdős, *Some remarks on Euler's φ function*, Acta Arith. **4** (1958), 10–19.

[5] K. Ford, *The distribution of totients*, Ramanujan J. **2** (1998), no. 1–2, 67–151; updated version arXiv:1104.3264v2 (2013), 43 pp.

[6] K. Ford, *Sieving by very thin sets of primes, and Pratt trees with missing primes*, Int. Math. Res. Not. IMRN **2014**, no. 11, 2955–2971; arXiv:1212.3498.

[7] V. L. Klee, Jr., *On a conjecture of Carmichael*, Bull. Amer. Math. Soc. **53** (1947), no. 12, 1183–1186.

[8] C. Pomerance, *On Carmichael's conjecture*, Proc. Amer. Math. Soc. **43** (1974), no. 2, 297–298.

[9] A. Schlafly and S. Wagon, *Carmichael's conjecture on the Euler function is valid below* $10^{10,000,000}$, Math. Comp. **63** (1994), no. 207, 415–419.