VALUES OF $v(k)$, THE SMALLEST DENOMINATOR MISSING FROM A k-TERM UNIT FRACTION DECOMPOSITION OF 1

KOREA SUPERINTELLIGENCE LABS

Abstract. For a positive integer k let D_k be the set of integers that occur as a denominator in some identity $1 = 1/n_1 + \cdots + 1/n_k$ with $1 \le n_1 < \cdots < n_k$, and let $v(k)$ be the smallest integer larger than 1 not in D_k. Erdős and Graham introduced $v(k)$ in 1980 and asked how fast it grows; van Doorn and Tang have recently proved the first lower bound, $v(k) \ge e^{ck^2}$, and their paper prints no value of $v(k)$ beyond $D_2 = \emptyset$ and $D_3 = \{2, 3, 6\}$. We determine $v(k)$ for every $k \le 8$, namely $2, 2, 4, 11, 17, 103, 733, 27539$. We have found no source that states $v(6) = 103$, though a 2020 comment on the OEIS sequence A330808 already implies its non-membership half; $v(7) = 733$ and $v(8) = 27539$ are the seventh and eighth terms of the OEIS sequence A097048, which records data on the neighbouring problem of Erdős and Graham on the length $N(a, b)$ of Egyptian fraction expansions. All eight values are proved here rather than quoted. We also compute $|D_k|$ for $k \le 6$, namely $1, 0, 3, 15, 77, 1152$, show that D_k is finite for every k, and prove the upper bound $v(k) \le |D_k| + 2$ in general. Everything rests on a bounded depth-first search that is proved *exactly* correct — its output is the set of decompositions, not merely a subset of it — so that an empty search is a non-existence theorem. The proofs are machine-checked in Lean 4; Section 9 records the scope of the formalisation and which statements are checked by the proof kernel alone.

1. Introduction

For a positive integer k let S_k be the set of k-tuples $(n_1, \ldots, n_k)$ of integers with

$$1 = \frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_k}, \qquad 1 \le n_1 < n_2 < \cdots < n_k, \tag{1}$$

a *k-term unit fraction decomposition of* 1; throughout, denominators are pairwise distinct. Following [8] we write

$$F(k) := |S_k|, \qquad D_k := \{m : \ m = n_i \text{ for some } (n_1, \ldots, n_k) \in S_k \text{ and some } i\},$$

so that D_k is the set of integers occurring as a denominator in at least one k-term decomposition of 1. In their 1980 monograph [10, p. 35] Erdős and Graham defined

$$v(k) := \min\{m \in \mathbb{N} : \ m > 1, \ m \notin D_k\}$$

and asked how fast it grows (they index by n rather than k). They remarked that $v(k) > ck!$ for some $c > 0$ is easy to see from earlier work of Bleicher and Erdős, and speculated that $v(k)$ may actually grow more like $2^{2^{\sqrt{k}}}$, or even like $2^{2^{k(1-\varepsilon)}}$. Graham repeated the question in [11, §4], and it is Problem #293 on Bloom's Erdős Problems site [6].

The elementary upper bound

$$v(k) \ \le \ |D_k| + 2 \ \le \ kF(k) + 2 \tag{2}$$

is [8, (1.2)]; combined with the upper bound for $F(k)$ of Elsholtz and Planitzer [9] it gives $v(k) \le c_0^{(1/5+o(1))2^k}$, where $c_0 = 1.264085\ldots$ is the Vardi constant. In the other direction van Doorn and Tang [8] prove the first lower bound in the literature: there is an absolute constant $c > 0$ with $v(k) \ge e^{ck^2}$ for every positive integer k. Their proof runs through Vose's theorem [15] on short Egyptian fraction expansions, together with a monotonicity lemma $D_k \subseteq D_{k+1}$ for $k \ge 2$ ([8, Lemma 2.1]).

What is conspicuously absent from that account is the sequence itself. The only values printed in [8] are $D_2 = \emptyset$ and $D_3 = \{2, 3, 6\}$, inside the proof of its Lemma 2.1. No numerical value of $v(k)$ for $k \geq 4$ appears, and neither $F(k)$ nor $|D_k|$ is tabulated; the paper contains no table at all.[1]

Results. This paper supplies the missing values and the proofs.

Theorem 1.1 (Section 4)**.** *$v(1) = 2$, $v(2) = 2$, $v(3) = 4$, $v(4) = 11$, $v(5) = 17$, $v(6) = 103$, $v(7) = 733$ and $v(8) = 27539$.*

Theorem 1.2 (Section 5)**.** *D_k is finite for every k, and $|D_4| = 15$, $|D_5| = 77$, $|D_6| = 1152$.*

Together with $D_1 = \{1\}$ and the two sets printed in [8] this yields Table 1.

k	1	2	3	4	5	6	7	8
$v(k)$	2	2	4	11	17	103	733	27539
$F(k)$	1	0	1	6	72	2320	$*$	–
$\lvert D_k\rvert$	1	0	3	15	77	1152	$*$	–

TABLE 1. The three sequences. Entries marked $*$ ($F(7) = 245765$, $|D_7| = 83231$) come from an auxiliary C program and are not part of the formal development.

Three comments on what is and is not new here.

$v(6) = 103$. We have found no source that states this value as $v(6)$, but it is not unrecorded either, and it is worth being exact about what was already known. The OEIS sequence A330808 [4], "Minimum number of unit fractions that must be added to $1/n$ to reach 1" — in which, its first comment says, "[t]he unit fraction $1/n$ and the unit fractions to be added to it need not be distinct" — carries a comment, contributed in 2020, listing the record positions

$$a(1) = 0,\ a(2) = 1,\ a(3) = 2,\ a(5) = 3,\ a(11) = 4,\ a(17) = 5,\ a(103) = 6,\ a(733) = 7,\ a(27539) = 8.$$

Since $a(103) = 6$ says that $102/103$ is not a sum of five unit fractions even when repetitions are allowed, it gives $103 \notin D_6$ outright, and that is the half of $v(6) = 103$ which needs an exhaustive search; Knott's expository page [5] gives the same, listing $102/103$ among the fractions of denominator at most 130 that need six unit fractions. The other half does not follow. From $a(m) \leq 5$ for $2 \leq m \leq 102$ one gets decompositions of $(m-1)/m$ that may repeat a denominator or use m itself, which do not place m in D_6; at $m = 2$ that is not a hypothetical, since $1 - 1/2 = 1/2$. What we add at $k = 6$ is therefore the identification with $v(6)$ in the distinct-denominator sense of Erdős and Graham, and a proof of both halves. Beyond A330808 we found nothing: OEIS searches for $2, 2, 4, 11, 17, 103, 733, 27539$ and for its segments return nothing against a working positive control, the problem page [6] records no value of $v(k)$, [8] had no citing work, and a zbMATH search through the public API found none. MathSciNet was not searched. All of this was checked on 2026-08-28.

$v(7) = 733$ and $v(8) = 27539$ are not new. Searching the OEIS for the single value 27539 returns A097048 [2], "$a(n)$ = least denominator Y of the proper fractions X/Y which need n or more terms as an Egyptian fraction", whose terms are $2, 3, 5, 11, 17, 79, 733, 27539$ and whose example line states that "27538/27539 is the simplest rational that cannot be expressed as the sum of 7 or fewer distinct unit fractions". Since a 7-term decomposition of $27538/27539$ avoiding the denominator 27539 is in particular a 7-term decomposition, that statement gives $27539 \notin D_8$ outright; and the definition of the sequence gives the same conclusion at the seventh term, whose recorded extremal fraction is $732/733$ (the seventh term of the companion numerator sequence A097049 [3] is 732), so that $733 \notin D_7$ as well. Both values are therefore already recorded: 733 in the original submission of

[1] We checked this in both arXiv versions of [8], the second of which (arXiv:2512.22083v2, May 2026) is the accepted manuscript; we have not read the publisher's version of record. The two versions differ in how small m are handled: the first appealed to a data file [16] for decompositions of 1 with at most 41 terms containing each $m \leq 432$, and the published version replaces that appeal by an explicit construction with at most $m + 1$ terms.

A097048 by Pegg and Reble in 2004, and 27539 in van der Sanden's extension of 2010 [14]. We reached them independently rather than first. What we add at $k = 7, 8$ is a proof: a complete set of witnesses and a completed exhaustive search, both machine-checked.

The two sequences are different. The values of $v(k)$ and of A097048 agree at $k = 1, 4, 5, 7, 8$ and differ at $k = 2, 3, 6$ (Section 7). Two distinct effects cause this. Membership $m \in D_k$ requires a $(k-1)$-term decomposition of $(m-1)/m$ *none of whose denominators is* m, whereas $N(m-1, m) \le k-1$ imposes no such condition; and A097048 minimises over all numerators X, not only over $X = Y - 1$. The first effect is already visible at $k = 3$, the second bites first at $k = 6$.

Section 3 describes the search and the sense in which it is proved correct; Sections 4 and 5 prove Theorems 1.1 and 1.2; Section 6 reproduces, as machine checks, every numerical statement made in [8]; Sections 7 and 8 treat the relation to Erdős Problem #304 and the controls that certify the formal statement of the problem. Section 9 describes the verification.

2. Notation

We identify a tuple $(n_1, \dots, n_\ell)$ with the finite strictly increasing sequence it names, and write $\sigma(n_1, \dots, n_\ell) = \sum_{i=1}^{\ell} 1/n_i \in \mathbb{Q}$ for its reciprocal sum, with $\sigma() = 0$. Thus $S_k = \{(n_1, \dots, n_k) : 1 \le n_1 < \cdots < n_k,\ \sigma = 1\}$ and D_k is the set of entries of tuples in S_k. We shall constantly use the following localised version of S_k, in which the target value, the number of terms, a lower bound for the denominators and one forbidden denominator are all parameters.

Definition 2.1. For integers $a \ge 0$, $b \ge 1$, $\ell \ge 0$, $\lambda \ge 1$ and $f \ge 0$ put

$$\mathcal{E}_f\big(\tfrac{a}{b};\ \ell,\ \lambda\big) := \Big\{(n_1, \dots, n_\ell) :\ \lambda \le n_1 < n_2 < \cdots < n_\ell,\ n_i \ne f \text{ for all } i,\ \sigma(n_1, \dots, n_\ell) = \tfrac{a}{b}\Big\}.$$

The value $f = 0$ forbids nothing, since all entries are at least $\lambda \ge 1$; with that choice $S_k = \mathcal{E}_0(1/1; k, 1)$. Note that a/b is not assumed to be in lowest terms: the recursion below produces unreduced fractions and never reduces them.

Finally, $v(k)$ is by definition the least element of $M_k := \{m \in \mathbb{N} : m > 1,\ m \notin D_k\}$, so that the assertion "$v(k) = V$" unfolds into the two statements

$$V > 1 \ \text{ and } \ V \notin D_k, \qquad \text{and} \qquad 1 < m < V \implies m \in D_k. \tag{3}$$

Every value claimed in this paper is proved in exactly this two-part form. That M_k is non-empty — so that $v(k)$ exists at all — follows from Proposition 5.1 below.

3. A search that is exactly the set of decompositions

All the arithmetic here is elementary; the point of this section is that the search we use is proved *complete* as well as sound, because completeness is what turns an empty search into a theorem of non-existence.

Definition 3.1. For $f \ge 0$, $a \ge 0$, $b \ge 1$, $\lambda \ge 1$ define a finite list $\mathrm{enum}(f, a, b, \lambda, \ell)$ of tuples by recursion on ℓ:

$$\mathrm{enum}(f, a, b, \lambda, 0) = \begin{cases} [\,()\,] & a = 0, \\ [\,] & a > 0, \end{cases}$$

and, for $\ell \ge 0$, $\mathrm{enum}(f, a, b, \lambda, \ell + 1) = [\,]$ if $a = 0$, while for $a > 0$ it is the concatenation, over the integers

$$\max\big(\lambda, \lceil \tfrac{b}{a} \rceil\big) \ \le\ n \ \le\ \Big\lfloor \tfrac{(\ell+1)b}{a} \Big\rfloor, \qquad n \ne f, \tag{4}$$

of the lists obtained by prepending n to every member of $\mathrm{enum}(f,\ an - b,\ bn,\ n + 1,\ \ell)$.

Prepending n replaces the target a/b by $a/b - 1/n = (an - b)/(bn)$; the left inequality in (4) is exactly $b \le an$, so the subtraction never leaves $\mathbb{N}$.

Theorem 3.2. *For all $f \geq 0$, $a \geq 0$, $b \geq 1$, $\lambda \geq 1$ and $\ell \geq 0$, the members of the list* $\mathrm{enum}(f, a, b, \lambda, \ell)$ *are exactly the elements of $\mathcal{E}_f(a/b; \ell, \lambda)$.*

Proof sketch. Induction on ℓ. For $\ell = 0$ the empty tuple has $\sigma = 0$, which equals a/b if and only if $a = 0$. For the inductive step, soundness is immediate from the induction hypothesis: the recursive call has lower bound $n + 1$, which forces strict increase, and $\sigma(n, t) = 1/n + (an - b)/(bn) = a/b$. Completeness requires the two bounds in (4) to be valid, i.e. that they exclude no admissible first denominator. If $(n_1, \ldots, n_{\ell+1}) \in \mathcal{E}_f(a/b; \ell + 1, \lambda)$ then $1/n_1 \leq a/b$, because the remaining terms are non-negative, which is the left-hand inequality; and $a/b = \sigma \leq (\ell + 1)/n_1$, because all $\ell + 1$ denominators are at least n_1, which is the right-hand one. The second of these is the only inequality in the development that needs an argument beyond rearrangement: for a tuple of length L with all entries $\geq n > 0$ one has $\sigma \leq L/n$, by induction on L. □

Theorem 3.2 has two immediate consequences that carry all the computations below. First, taking $f = 0$, $a = b = \lambda = 1$, the list $\mathrm{enum}(0, 1, 1, 1, k)$ enumerates S_k without repetition, so D_k is precisely the set of entries occurring in it and $F(k)$ is its length. Secondly, an *empty* localised search is a non-existence theorem, which we use in the following shape.

Lemma 3.3 (Deletion)**.** *Let $k \geq 2$ and $m \geq 2$. If $\mathcal{E}_m\big(\frac{m-1}{m}; k - 1, 2\big) = \emptyset$ then $m \notin D_k$.*

Proof sketch. Suppose $m \in D_k$, witnessed by $(n_1, \ldots, n_k) \in S_k$ containing m. Deleting the entry m leaves $k - 1$ distinct integers, still strictly increasing, with reciprocal sum $1 - 1/m = (m - 1)/m$ and none of them equal to m. They are all at least 2: a decomposition of 1 with $k \geq 2$ terms cannot contain the denominator 1, since the remaining terms would contribute a strictly positive amount. So the deleted tuple lies in $\mathcal{E}_m((m - 1)/m; k - 1, 2)$, contradicting the hypothesis. □

The gain is large. Membership of m in D_k is a statement about the whole of S_k, whereas Lemma 3.3 localises it to decompositions of a fraction very close to 1; the bounds (4) then force the first denominator of such a decomposition to be at most $\lfloor (k - 1)m/(m - 1) \rfloor = k - 1$ for $m > k$, which prunes the search by orders of magnitude. At $k = 7$ the localised search visits 35 631 nodes where enumerating S_7 visits 69 494 952.

For the positive direction — exhibiting $m \in D_k$ — one checks a witness, and it is worth recording that no rational arithmetic is involved.

Lemma 3.4. *Let $n_1, \ldots, n_\ell$ be positive integers with product P. Then $\sum_i 1/n_i = 1$ if and only if $\sum_i P/n_i = P$, an identity between integers. Consequently, if a strictly increasing list of positive integers of length k satisfies that integer identity and contains m, then $m \in D_k$.*

Lemma 3.5 (Covering)**.** *Let $T = (t_0, t_1, \ldots, t_{n-1})$ be a list of k-term decompositions of 1 such that t_i contains the entry $\lambda + i$ for each i. Then every integer m with $\lambda \leq m < \lambda + n$ lies in D_k.*

Lemmas 3.3, 3.4 and 3.5 reduce each value of $v(k)$ to one finite table plus one finite search, and (3) assembles them.

4. The values of $v(k)$

Theorem 4.1. *$v(1) = 2$, $v(2) = 2$, $v(3) = 4$, $v(4) = 11$, $v(5) = 17$ and $v(6) = 103$.*

Proof sketch. For $k \leq 5$ both halves of (3) are read off the *complete* enumeration $\mathrm{enum}(0, 1, 1, 1, k)$ of S_k, which by Theorem 3.2 is exactly S_k; the lists have $1, 0, 1, 6, 72$ members respectively. For $k = 1$ the single decomposition is $1 = 1/1$, so $D_1 = \{1\}$ and $v(1) = 2$. For $k = 2$ the enumeration is empty, so $D_2 = \emptyset$ and again $v(2) = 2$. For $k = 3, 4, 5$ the entry sets are $\{2, 3, 6\}$, a 15-element set and a 77-element set; the smallest integer > 1 absent from them is 4, 11, 17 respectively.

At $k = 6$ the same route is available in principle but is expensive to certify (see Remark 4.4), so $v(6) = 103$ is proved in the two-part shape instead. The upper half is a table of 101 six-term decompositions of 1, the i-th containing $i + 2$, which by Lemma 3.5 gives $\{2, 3, \ldots, 102\} \subseteq D_6$; each row is re-checked from scratch by the integer identity of Lemma 3.4. The lower half is the

exhaustive search $\mathcal{E}_{103}(102/103; 5, 2) = \emptyset$, which visits 357 nodes; by Lemma 3.3 it gives $103 \notin D_6$. Both halves are finite computations, and the second is where the claim "103 is missing" rests entirely on exhaustiveness. □

Theorem 4.2. *$v(7) = 733$ and $v(8) = 27539$.*

As explained in Section 1 these two values are not new: they are the seventh and eighth terms of A097048 [2], whose extremal fractions there are 732/733 and 27538/27539 [3], and the OEIS entry's own statement about 27538/27539 already implies $27539 \notin D_8$. We record the theorem because the proof below is a complete verification of both halves of (3), obtained independently of that source.

Proof sketch. Identical in shape to the case $k = 6$. For $k = 7$: a table of 731 seven-term decompositions of 1, the i-th containing $i + 2$, gives $\{2, \dots, 732\} \subseteq D_7$; and the exhaustive search $\mathcal{E}_{733}(732/733; 6, 2) = \emptyset$, which visits 35 631 nodes, gives $733 \notin D_7$. For $k = 8$: a table of 27 537 eight-term decompositions, whose largest denominator is 113 423 713 055 421 844 361 000 442, gives $\{2, \dots, 27538\} \subseteq D_8$; and the exhaustive search $\mathcal{E}_{27539}(27538/27539; 7, 2) = \emptyset$, which visits 95 837 829 nodes, gives $27539 \notin D_8$. Every row of both tables is re-checked from scratch inside the proof. □

Remark 4.3 (Provenance of the tables). The witness tables and the node counts come from a separate C implementation of the same search, which at eight terms replaces the linear scan at the two-terms-left level by the divisor solution of $1/n + 1/n' = a/b$, i.e. $(an - b)(an' - b) = b^2$. Nothing in the proofs depends on that program being correct: a table is only a list of candidate tuples, each of which is re-verified by Lemma 3.4, and the non-membership half is run again by the plain search. This mattered in practice. All 27 537 rows of the $k = 8$ table were independently re-checked with exact rational arithmetic outside the proof assistant, which exposed two 64-bit overflow bugs in the C search: one in the divisor construction, and one a truncation of denominators past 2^{64} which corrupted the recorded witnesses and sat also in the comparison against the forbidden denominator, where it could in principle have produced a spurious non-membership. After the fix the answer was unchanged and was confirmed a third time by the naive enumerator.

Remark 4.4 (Measured cost). For orientation, and because it explains the shape of the proof at $k = 6$: reducing the 27 679-node enumeration of S_6 inside the proof kernel costs about 167 seconds and 10.6 gigabytes, whereas the two-part proof of $v(6) = 103$ costs about 10 seconds. The cost per search node inside the kernel is some three orders of magnitude above the cost of the same node in compiled code. The whole development checks in about 286 seconds, of which 251 are the single search at $k = 8$.

5. The cardinality of D_k, and the bound $v(k) \le |D_k| + 2$

Proposition 5.1. *For every k, D_k is finite; indeed D_k is the set of entries of the finite list* $\mathrm{enum}(0, 1, 1, 1, k)$. *In particular $\{m > 1 : m \notin D_k\}$ is non-empty and $v(k)$ is well defined.*

Theorem 5.2. *$|D_4| = 15$, $|D_5| = 77$ and $|D_6| = 1152$.*

Proof sketch. By Proposition 5.1 each cardinality is the number of distinct entries of the complete enumeration of S_k, which has 6, 72 and 2320 members respectively; the three counts are the result of that finite computation. □

For $k \le 3$ the cardinalities are $|D_1| = 1$, $|D_2| = 0$ and $|D_3| = 3$, immediately from $D_1 = \{1\}$ and Theorem 6.1 below. The resulting sequence $1, 0, 3, 15, 77, 1152$ — extended by $|D_7| = 83231$, which is an unformalised computation of the auxiliary C program — we have not found in the OEIS or in the literature. The OEIS queries for $1, 0, 3, 15, 77, 1152, 83231$ and for its segments return nothing against a working positive control, and neither a zbMATH search through the public API nor a web search for 83231 in this context returns anything; MathSciNet was not searched. Checked on 2026-08-28.

Proposition 5.3. *For every k, $v(k) \le |D_k| + 2$.*

Proof. All $v(k) - 2$ integers m with $1 < m < v(k)$ lie in D_k by minimality, so the interval $\{2, 3, \dots, v(k) - 1\}$ injects into D_k and $v(k) - 2 \le |D_k|$. □

This is the first inequality of (2); we have not formalised the second, $|D_k| \le kF(k)$, which is immediate from the definitions. It is worth recording how weak (2) is in the range where we can measure it: at $k = 6$ Proposition 5.3 reads $103 \le 1154$, and $kF(k) + 2 = 13922$. The gap of an order of magnitude at the first interesting value is a concrete illustration of why the lower bound of [8] must be proved by a different route — through Vose's theorem — rather than by counting.

6. The computed statements of [8], re-derived

Any formalisation of a problem may formalise the wrong problem. The check we can offer against that is that the definitions used here reproduce, from the search alone, every numerical assertion made in [8], and reproduce the published enumeration of $F(k)$.

Theorem 6.1. *$D_2 = \emptyset$ and $D_3 = \{2, 3, 6\}$. Moreover $(2, 3, 7, 42)$ and $(2, 4, 6, 12)$ are 4-term decompositions of 1, and $D_3 \subseteq D_4$.*

This is the chain $D_2 = \emptyset \subseteq D_3 = \{2, 3, 6\} \subseteq \{2, 3, 7, 42\} \cup \{2, 4, 6, 12\} \subseteq D_4$ asserted with "as one can verify" in the proof of [8, Lemma 2.1], which is the base case $k \le 3$ of that lemma; the general case, an induction on k, we have not formalised.

Theorem 6.2. *$F(k) = 1, 0, 1, 6, 72, 2320$ for $k = 1, \dots, 6$.*

These are the first six terms of A006585 [1]; the auxiliary C program also reproduces $F(7) = 245765$, which is the seventh. The agreement is the evidence that the sets enumerated here are the S_k of (1) and not a near miss — in particular that the distinctness condition and the treatment of the denominator 1 match.

Theorem 6.3. *Let $D = \{3, 6, 9, 12, 15, 18, 24, 27, 30, 45, 54, 60, 72\}$. Then $\sum_{d \in D} 1/d = 1$; consequently every $d \in D$ lies in D_{13}. Moreover any two distinct $d > d'$ in D satisfy $d/d' \ge 10/9$.*

The set D and both of its properties are used in the proof of the main theorem of [8], where each is asserted with "one can verify"; the ratio inequality is displayed equation (2.4) there — the number is the same in both arXiv versions — and is what rules out collisions among the auxiliary denominators.

7. Relation to Erdős Problem #304

For integers $1 \le a < b$ let $N(a, b)$ be the least t for which there are integers $1 < n_1 < \cdots < n_t$ with $a/b = \sum_i 1/n_i$, and $N(b) = \max_{1 \le a < b} N(a, b)$; estimating $N(b)$ is Erdős Problem #304 [7]. The concluding section of [8] observes the one-way implication

$$1 < b < v(k) \implies b \in D_k \implies N(b - 1, b) \le k - 1, \tag{5}$$

by deleting the term $1/b$ from a k-term decomposition containing it, and remarks that this is why the two problems are closely connected.

The connection is not an equivalence, and the sequence A097048 shows where it fails. Its terms $2, 3, 5, 11, 17, 79, 733, 27539$ agree with $v(1), \dots, v(8)$ at $k = 1, 4, 5, 7, 8$ and differ at $k = 2, 3, 6$. Two separate effects are responsible.

The forbidden denominator. The implication (5) cannot be reversed, because a short decomposition of $(m - 1)/m$ may be forced to use the denominator m itself, and then it does not certify $m \in D_k$.

Proposition 7.1. *$\mathcal{E}_0(3/4; 2, 2) = \{(2, 4)\}$ and $\mathcal{E}_4(3/4; 2, 2) = \emptyset$. Hence $N(3, 4) = 2$, while $4 \notin D_3$ by Lemma 3.3.*

That is, $3/4 = 1/2 + 1/4$ is the only decomposition of $3/4$ into two distinct unit fractions with denominators > 1, and it re-uses the denominator 4; that $N(3, 4) = 2$ rather than 1 is clear, as $3/4$ is not a unit fraction. If one deletes the forbidden denominator from the search it reports $v(3) = 5$, which is the third term of A097048. The same phenomenon — a forced repetition of m itself — accounts for $v(2) = 2$ against the second term 3.

The numerator. A097048 minimises over all numerators, so its k-th term can be smaller than the least b for which $(b-1)/b$ is hard. At $k = 6$ this is what happens: the sixth term of A097048 is 79, its extremal fraction being $77/79$, whereas $78/79$ needs only five terms and consequently $79 \in D_6$ — as the covering half of Theorem 4.1 shows directly, $\{2, \dots, 102\} \subseteq D_6$. So $v(6) = 103 > 79$. At $k = 7$, by contrast, neither effect operates: the extremal fraction recorded there is again of the shape $(b-1)/b$, and the following sharpening shows that the forbidden denominator is not doing the work either.

Proposition 7.2. *$\mathcal{E}_0(732/733; j, 2) = \emptyset$ for every $j \le 6$; that is, $732/733$ is not the sum of six or fewer distinct unit fractions with denominators > 1, so $N(732, 733) \ge 7$. Combined with $v(7) = 733$ and (5), the integer 733 is the least b with $N(b-1, b) \ge 7$.*

Here the non-membership $733 \notin D_7$ holds for the stronger reason: not merely is there no 6-term decomposition of $732/733$ avoiding 733, there is no 6-term decomposition at all. Proposition 7.2 is consistent with, and no stronger than, the tabulated data of A097048; we record it because it isolates which of the two effects is operating at each k.

A third sequence. The separation of the two effects can be read off a third recorded sequence. The record positions in the comment on A330808 [4] quoted in Section 1 are $2, 3, 5, 11, 17, 103, 733, 27539$: they concern the fractions $(b-1)/b$ only, which removes the numerator effect, but they allow repeated denominators, which keeps the forbidden one. That sequence is A097048 with its sixth term 79 replaced by 103, and it differs from $v(k)$ exactly at $k = 2$ and $k = 3$ — the two indices at which, by the analysis above, the forbidden denominator bites. We have not verified A330808 itself; we record the agreement because it is an independent check on which of the two effects is responsible where.

8. Controls

Every claim above has the form "V is the least integer > 1 missing from D_k", a shape that a misstated definition can satisfy vacuously. The following statements are all proved in the same development, and they refute the specific ways the statement could be too weak.

Proposition 8.1.
(i) Not too small: $102 \in D_6$ *and* $732 \in D_7$, *so neither is* $v(6)$ *or* $v(7)$.
(ii) Not too large: 104 *is not the least element of* M_6, *and* 734 *is not the least element of* M_7.
(iii) Distinctness is load-bearing: $1/3 + 1/3 + 1/3 = 1$ *and* $1/2 + 1/4 + 1/4 = 1$, *but neither* $(3, 3, 3)$ *nor* $(2, 4, 4)$ *is a decomposition in the sense of* (1); *dropping strict increase would raise* $F(3)$ *from* 1 *to at least* 3.
(iv) The sum condition is load-bearing: $(2, 3, 7, 43, 1807, 3263443, 10650056950806)$ *is a* 7*-term decomposition of* 1, *while replacing its last entry by* 10650056950807 *leaves a strictly increasing tuple of positive integers that is not one.*
(v) The clause $m > 1$ is load-bearing: $1 \in D_1$, *and* $1 \notin M_k$ *for every* k *by that clause rather than by the arithmetic.*
(vi) D_k really grows: $103 \notin D_6$ *and* $103 \in D_7$.

Item (iv) is the Sylvester decomposition of 1 perturbed by one in the last denominator; the perturbed tuple is the natural near miss, since 10650056950807 is the corresponding Sylvester number. Item (vi) is the smallest instance of $D_k \subsetneq D_{k+1}$ visible in our range and is a special case of [8, Lemma 2.1].

Remark 8.2. Erdős and Graham record that $v(k) > ck!$ for some $c > 0$. Finitely many values cannot refute an asymptotic statement, but they do constrain the implied constant: $v(5) = 17 < 120 = 5!$, so any $c > 0$ with $c \cdot k! \le v(k)$ for all k satisfies $c \le 17/120$. The ratios $v(k+1)/v(k)$ over our range are $1, 2, 2.75, 1.55, 6.06, 7.12, 37.6$, which is the behaviour one expects from e^{ck^2} and not from $k!$.

9. Verification

Every statement labelled a theorem, proposition or lemma above is a theorem of a formal development in Lean 4 [13] built on Mathlib [12], and each proof sketch above corresponds to a complete machine-checked proof; the finite computations are performed by the proof assistant rather than quoted from an external program. One clause is outside that development and we flag it: the quantity $N(a, b)$ of Section 7 is not formalised, so the assertions about N in Propositions 7.1 and 7.2 are consequences of the formalised statements standing next to them, obtained in one line by unfolding the definition of N. Within the development, two levels of trust occur, and we distinguish them. A statement is *kernel-checked* when its proof, including all of its finite computations, is carried out by reduction in the Lean kernel, so that it depends only on the three standard axioms of the library (propositional extensionality, choice and quotient soundness). Everything above is kernel-checked with the following exceptions, which are discharged instead by Lean's compiled evaluation (`native_decide`) and so depend on one further axiom, asserting that compiled Boolean evaluation agrees with the kernel: the values $F(5)$ and $F(6)$ of Theorem 6.2; the values $|D_5|$ and $|D_6|$ of Theorem 5.2; the whole of Theorem 4.2; Proposition 7.2; and, in Proposition 8.1, the assertion of (ii) about 734 and the assertion $103 \in D_7$ of (vi), both of which quote the $k = 7$ computation. The corresponding parts of Theorems 1.1 and 1.2, which restate the above, inherit the same status. In particular $v(k)$ for every $k \le 6$ — including $v(6) = 103$ — and $|D_4| = 15$ are kernel-checked, with no appeal to compiled evaluation anywhere in their proofs. A reader who does not accept the extra axiom should read the listed statements as verified computations rather than as kernel-checked theorems; nothing about $k \le 6$ depends on them. The $k = 8$ witness table is stored as a single string and parsed inside the development, which is harmless: the parsed rows are re-checked from scratch by Lemma 3.4, so a misparse can only make the check fail. Repository reference to be added at submission.

10. Limits of the method

The value $v(9)$ is out of reach of everything above, twice over.

On the verification side the obstruction is time, not memory: the search at $k = 8$ visits 95 837 829 nodes at roughly 2.6 microseconds each, and its peak memory is about 450 megabytes over the cost of loading the library. The change that would matter is replacing the linear scan at the two-terms-left level by the divisor solution used in Remark 4.3, which requires a verified integer factorisation rather than a verified loop.

On the search side, van der Sanden records in a comment on A097048 [2] that "Checking just $(p-1)/p$ for prime p finds no example requiring 9 parts for $p \le 800399$", the underlying data being the file `results-single` of [14]. So $v(9)$ is plausibly at least six figures, and a witness table for it would have on the order of 10^6 rows. Two things temper that guess. The quoted search is over the fractions $(p-1)/p$ with p prime only, whereas the exhaustive search over all proper fractions behind A097048 reaches only denominator 27539; and $N(m-1, m) \le 8$ does not by itself give $m \in D_9$, since the short decomposition may be forced to use m itself — which is the whole content of Section 7.

The route that would scale is not more computation but the missing lemma: $D_k \subseteq D_{k+1}$, proved in [8, Lemma 2.1] by an elementary but case-heavy argument, would let every m below $v(k)$ be inherited from level k rather than re-witnessed, and would replace the entire lower half of each computation. That lemma is the single piece of [8] whose formalisation would most change what is reachable here.

References

[1] OEIS Foundation Inc., *Sequence A006585: Egyptian fractions: number of solutions to* $1 = 1/x_1 + \cdots + 1/x_n$ *in positive integers* $x_1 < \cdots < x_n$, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A006585`.

[2] OEIS Foundation Inc., *Sequence A097048:* $a(n)$ *= least denominator* Y *of the proper fractions* X/Y *which need* n *or more terms as an Egyptian fraction*, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A097048`. Contributed by E. Pegg Jr. and D. Reble, 21 July 2004; the eighth term is due to H. van der Sanden, 14 September 2010, as is the comment on $(p-1)/p$ quoted in Section 10, dated 28 February 2015. The offset is 1, so $a(1) = 2$ and $a(8) = 27539$.

[3] OEIS Foundation Inc., *Sequence A097049:* $a(n)$ *= least numerator* X *of the proper fractions* $X/A097048(n)$ *which need* n *or more terms as an Egyptian fraction*, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A097049`. Same contributors and dates as A097048; its terms are $1, 2, 4, 8, 16, 77, 732, 27538$.

[4] OEIS Foundation Inc., *Sequence A330808: minimum number of unit fractions that must be added to* $1/n$ *to reach* 1, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A330808`. Contributed by J. E. Schoenfield, 11 January 2020, including the comment on record positions quoted in Section 1.

[5] R. Knott, *Egyptian fractions*, University of Surrey, `https://r-knott.surrey.ac.uk/Fractions/egyptian.html` (consulted 28 August 2026).

[6] T. F. Bloom, *Erdős Problem #293*, `https://www.erdosproblems.com/293`.

[7] T. F. Bloom, *Erdős Problem #304*, `https://www.erdosproblems.com/304`.

[8] W. van Doorn and Q. Tang, *The smallest denominator not contained in a unit fraction decomposition of* 1 *with fixed length*, Math. Proc. Cambridge Philos. Soc., published online 8 July 2026, pp. 1–9, `doi:10.1017/S0305004126102102`; arXiv:2512.22083.

[9] C. Elsholtz and S. Planitzer, *Sums of four and more unit fractions and approximate parametrizations*, Bull. Lond. Math. Soc. **53** (2021), no. 3, 695–709, `doi:10.1112/blms.12452`.

[10] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monogr. Enseign. Math. **28**, L'Enseignement Mathématique, Université de Genève, Geneva, 1980, 128 pp.; Zbl 0434.10001.

[11] R. L. Graham, *Paul Erdős and Egyptian fractions*, in: L. Lovász, I. Z. Ruzsa and V. T. Sós (eds.), Erdős Centennial, Bolyai Soc. Math. Stud. **25**, Springer, Berlin, and János Bolyai Math. Soc., Budapest, 2013, 289–309, `doi:10.1007/978-3-642-39286-3_9`. The question about $v(k)$ is in its Section 4.

[12] 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, `https://doi.org/10.1145/3372885.3373824`.

[13] 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, `https://doi.org/10.1007/978-3-030-79876-5_37`.

[14] H. van der Sanden, *least_eg*: code and results, linked from A097048, `https://github.com/hvds/seq/tree/master/least_eg`.

[15] M. D. Vose, *Egyptian fractions*, Bull. London Math. Soc. **17** (1985), no. 1, 21–24, `doi:10.1112/blms/17.1.21`.

[16] W. van Doorn, *Smallest missing denominator data*. `https://github.com/Woett/Smallest-missing-denominator-data`