Short representations and counting with distinct odd unit fractions

Pengcheng Chen

hexagon:2611.00949v1

Abstract

Every rational number a/b in $(0,1)$ with b odd is a sum of $O(1+\log\log b)$ distinct unit fractions with odd denominators. This gives the order of both the maximum and the mean minimum length at each odd denominator. For the representations of 1 with exactly k terms, we show that the double logarithm of their number has order k along the odd integers. In fact, for each sufficiently large odd k, one can fix the first $k-2$ denominators and still obtain at least $\exp(\exp(ck))$ different endings. Each largest denominator in this family is divisible by all of the first $\lfloor\exp(ck)\rfloor$ odd primes. We also obtain doubly exponentially many possible largest denominators, each attained by doubly exponentially many representations of that length.

1 Introduction

Every rational number in $(0,1)$ with odd denominator has a representation by distinct odd unit fractions. Two quantitative questions concern the shortest such representation of a/b and the number of representations of 1 at a prescribed length. The qualitative existence theorem goes back to Breusch and Stewart, as discussed in the introduction to [6]. When even denominators are allowed, the uniform $O(\log\log b)$ length bound and the double-exponential count have been obtained in [3] and further developed in [2]. The odd restriction changes both the elimination of repeated terms and the lengths at which a representation of 1 can occur.

An *odd representation* of a rational x is a tuple of odd integers $1<d_1<\cdots<d_k$ satisfying $x=\sum_{i=1}^{k}1/d_i$. Write $L_o(a,b)$ for the least length of an odd representation of a/b, and $S_o(k)$ for the number of odd representations of 1 with k terms. The order of the denominators is fixed throughout. Clearing the odd denominators in a representation of 1 and reducing modulo 2 shows that its length is odd. All logarithms are natural. All constants in the main theorems are absolute.

Theorem 1.1 (length). For every odd integer $b\geq 3$ and every integer $1\leq a<b$,

$$L_o(a,b)\leq C(1+\log\log b).$$

As b tends to infinity through the odd integers, both $\max_{1\leq a<b}L_o(a,b)$ and $(b-1)^{-1}\sum_{a=1}^{b-1}L_o(a,b)$ have order $\log\log b$. The mean includes all numerators, whether or not they are coprime to b.

Let Q_r be the product of the first r odd primes. Define $K(r)$ to be the least length of an odd representation of 1 whose largest denominator is a multiple of Q_r, with value $+\infty$ if none exists.

Theorem 1.2 (largest denominator). As $r\to\infty$,

$$K(r)\asymp\log r.$$

Theorem 1.3 (a fixed initial segment). There is a $c > 0$ such that, for each sufficiently large odd k, there exist odd integers $1 < a_1 < \cdots < a_{k-2}$ for which

$$\sum_{i=1}^{k-2} \frac{1}{a_i} + \frac{1}{v} + \frac{1}{w} = 1, \qquad a_{k-2} < v < w,$$

has at least $\exp(\exp(ck))$ odd pairs (v, w) with pairwise distinct w, all satisfying $Q_{\lfloor \exp(ck) \rfloor} \mid w$. In particular, $\log\log S_o(k) \asymp k$ along the odd integers.

The initial segment has the greatest possible length for a family with more than one member. Once $k-1$ denominators are fixed, the remaining reciprocal is determined. Elsholtz obtained the lower bound $\exp(\exp(ck/\log k))$ for distinct odd denominators [5, Corollary 1.2]. The lower bound in Theorem 1.3 attains the order of the double logarithm in the elementary upper bound even after the first $k-2$ denominators have been fixed.

Corollary 1.4 (multiplicity at a given maximum). For some $c > 0$ and every sufficiently large odd k, there are at least $\exp(\exp(ck))$ different integers M with the following two properties. Each M is divisible by $Q_{\lfloor \exp(ck) \rfloor}$, and each is the largest denominator of at least $\exp(\exp(ck))$ odd representations of 1 of length k.

Sections 2 and 3 prove the upper bound in Theorem 1.1. Following the smooth-fraction method of Bloom and van Doorn [2], we expand two integers in a common base. The base is an odd prime chosen so that every digit can be represented with boundedly many terms and none of their denominators absorbs a factor of the base. A finite family of substitutions then removes repeats within each digit position while keeping positions disjoint. A short initial sequence of odd unit fractions reduces a general numerator to the range of this construction.

Section 4 applies the length bound to $(Q_r - 1)/Q_r$, a choice also made in [2, §3.2]. We arrange that a multiple of Q_r is the largest denominator in the resulting representation of 1. Chen, Elsholtz and Jiang fix the term $1/(N+2)$ when splitting $1/N$, use divisors to vary the other two terms, and retain a multiple of N as the new maximum [6, Lemma 3.1]. They also count representations subject to a fixed divisor $3^{20}5^{20}7^{20}$ of the largest denominator [6, Lemma 6.2]. Here a length of order $\log r$ suffices for the growing divisor Q_r. Section 5 uses its 2^r divisors after adjusting the length to obtain the family in Theorem 1.3.

2 Smooth digits with odd denominators

Write $P(y) = \prod_{3 \le p \le y} p$, where the product is over primes, and let $\mathcal{D}(y)$ be its set of positive divisors. An empty product is 1. The digit construction in [2, Lemmas 5–7] and [3] represents integers by sums of fractions whose numerators have only small prime factors. We need all these fractions to have odd numerators and odd denominators. We also need their denominators to avoid one auxiliary prime, chosen before the digits are given.

Theorem 2.1 (uniform digit expansions and distinct layers). There are absolute positive integers J, F, R_* and a constant $c > 0$ with the following properties. Given $R \ge 1$, for every sufficiently large odd integer b, put

$$y = c\log b, \qquad P = P(y), \qquad N = P^2.$$

There is a prime $B > \max\{y, R, 3\}$ with $B \nmid b$ such that

$$N \le b^{1/8}, \qquad RN^2 \le b/2, \qquad L = \left\lceil \frac{\log b}{\log B} \right\rceil \ll \log\log b.$$

The threshold for b may depend on R. The implied constant is absolute.

(i) Every integer $0 \le m < B$ has a representation

$$m = \sum_{i=1}^{t} \frac{u_i v_i}{w_i}, \qquad t \le J, \quad u_i, v_i \mid P, \quad 1 \le w_i \le N, \quad w_i \text{ odd}, \quad B \nmid w_i.$$

The factors u_i, v_i are positive. Repetitions are allowed, and $m = 0$ uses the empty sum.

(ii) Suppose $R \ge R_*$. For any digits $0 \le m_{\varepsilon j} < B$, indexed by $\varepsilon \in \{0, 1\}$ and $1 \le j \le L$, each fraction $m_{\varepsilon j}/(b^\varepsilon N B^j)$ can be replaced by a sum of at most F unit fractions so that all denominators in their union are distinct. Every denominator in block (ε, j) has the form $b^\varepsilon B^j h$, where h is odd, $1 \le h \le R_* N^2 < b$, and $B \nmid h$.

2.1 Two estimates for sets of divisors

We use the Chebyshev bounds for $\vartheta(t) = \sum_{p \le t} \log p$ and $\pi(t) = \#\{p \le t\}$. Fix constants $0 < c_\vartheta \le C_\vartheta$, $C_\vartheta \ge 1$, and $a_0 > 0$ such that for all sufficiently large t,

$$c_\vartheta t \le \vartheta(t) \le C_\vartheta t, \qquad \log |\mathcal{D}(t)| = (\pi(t) - 1) \log 2 \ge a_0 t / \log t.$$

Take C_ϑ large enough that $\log P(t) \le C_\vartheta t$ for all $t > 0$.

Finite-field covering. If $A \subseteq \mathbb{F}_p$ and $|A| > p^{3/4}$, every $a \in \mathbb{F}_p^\times$ has the form $u_1 v_1 + u_2 v_2$ with $u_i, v_i \in A$. This is the two-dimensional case of the Hart–Iosevich covering theorem [4, Theorem 1.1]. For a short proof, set $E = A \times A$, $e = |E|$, and take the ℓ lines $v \cdot z = a$ indexed by $v \in E \setminus \{0\}$. These are distinct since $a \ne 0$. Each contains p points, and two distinct lines meet in at most one point. If all lines miss E, let n_z count the lines through each $z \notin E$. Then

$$\sum_{z \notin E} n_z = \ell p, \qquad \sum_{z \notin E} n_z^2 \le \ell p + \ell(\ell - 1).$$

Cauchy–Schwarz gives $\ell p^2 \le (p^2 - e)(p + \ell)$, whence $e(p + \ell) \le p^3$. But $\ell \ge e - 1$, so $e^2 \le p^3$, contrary to $e^2 = |A|^4 > p^3$. An incidence with E gives the required representation. Only $a = -1$ will be needed.

A uniform bound for bad primes. For a finite set $\mathcal{A}$ of positive integers, call a prime p bad for $\mathcal{A}$ if the reduction of $\mathcal{A}$ modulo p has at most $p^{3/4}$ elements. If $\mathcal{A} \subseteq [1, P(y)]$ has size $M \ge 2x^{3/4}$ and $x \ge 2$, then

$$\#\{p \le x : p \text{ bad for } \mathcal{A}\} \le C_{\mathrm{bad}} y x^{3/4}, \qquad C_{\mathrm{bad}} = 2C_\vartheta / \log 2.$$

Indeed, let a_i be the sizes of the nonempty residue classes at a bad prime $p \le x$. It divides the difference of at least $\sum_i a_i^2 - M \ge M^2/x^{3/4} - M \ge M^2/(2x^{3/4})$ ordered pairs of distinct elements. For each such pair the positive integer $|u - v| \le P(y)$ has at most $\log P(y)/\log 2$ distinct prime factors. Counting pairs and their prime divisors in the two orders proves the bound. Its constant applies to every subset and every admissible cutoff x.

We use Vinogradov's quantitative three-prime theorem [1] in the following form. There are constants $v > 0$ and $m_0 \ge 7$ such that every odd $m \ge m_0$ has at least $vm^2/(\log m)^3$ ordered representations as a sum of three odd primes, with repetition allowed. Starting with the count allowing 2 gives the same assertion after reducing v and increasing m_0, since a triple containing 2 must be a permutation of $(2, 2, m-4)$. If E_m primes up to m are bad for a given set, at most $3mE_m$ triples contain one of them. Choose its position, its value, and one other entry. The sum determines the last entry.

2.2 Smooth numerators at the scale of a single integer

Put $H(m) = K \log(2m) \log\log(3m)$. We claim that one can fix $K \geq 1$ and $J_0 \geq 10$ so that every positive integer m is a sum of at most J_0 fractions uv/w satisfying

$$u, v \in \mathcal{D}(H(m)), \qquad 1 \leq w \leq P(H(m))^2, \qquad w \text{ odd}.$$

To prove this, choose $K \geq 4/a_0$. For $Y = H(m)$ we have $Y/\log Y \sim K \log m$. Thus for all large m, $\log|\mathcal{D}(Y)| \geq 2\log m$ and in particular $|\mathcal{D}(Y)| \geq 2m^{3/4}$. Apply the bad-prime estimate at $x = m$. For odd $m \geq m_0$ sufficiently large the proportion of triples meeting a bad prime is at most

$$\frac{3C_{\text{bad}}}{v} \frac{Y(\log m)^3}{m^{1/4}} = o(1).$$

Fix $m_1 \geq m_0$ above which this bound is less than $1/2$ and the size condition holds. Each odd $m \geq m_1$ is a sum of three good odd primes.

For each such prime p, cover -1 using $\mathcal{D}(Y)$ mod p and lift the four factors. Then

$$u_1v_1 + u_2v_2 + 1 = pw, \qquad p = \frac{u_1v_1}{w} + \frac{u_2v_2}{w} + \frac{1}{w}.$$

The left side of the first identity is positive and odd. As $p \geq 3$, the quotient w is positive and odd and $w \leq (2P(Y)^2+1)/3 \leq P(Y)^2$. Summing over the three primes gives nine terms. For even $m \geq m_1 + 1$, use the expansion of $m - 1$ and append $1/1$. The factors still lie in $\mathcal{D}(H(m))$ since H is increasing for $m \geq 1$. Each remaining m is the sum of m copies of $1/1$. This proves the assertion with $J_0 = \max\{10, m_1\}$ and fixes $J = \max\{J_0 + 1, 10\}$ for Theorem 2.1.

2.3 Choosing a single base and residue class

Take $c = 1/(16C_\vartheta)$ and fix

$$0 < \eta \leq \min\{1/16, a_0/16\}, \qquad 0 < \delta \leq \min\{1/8, 1/(16C_\vartheta K)\}.$$

Choose a constant $D \geq 2$ with $c_\vartheta D - C_\vartheta \geq 2$. For $y = c\log b$ let $X = \exp(\eta y/\log y)$. For every sufficiently large y,

$$\vartheta(DX) - \vartheta(X) \geq 2X > \log b.$$

Consequently some prime $B \in (X, DX]$ does not divide b, since otherwise the product of the primes in this interval would divide b. Put $\lambda = \log B$. Uniformly for this choice,

$$\eta y/\log y < \lambda \leq 2\eta y/\log y.$$

Here $\log D \leq \eta y/\log y$ for all sufficiently large y. We can also ensure $X > y > \max\{R, 3\}$. We have $\log N \leq 2C_\vartheta y = (\log b)/8$, and $RN^2 \leq Rb^{1/4} \leq b/2$ once $b \geq (2R)^{4/3}$. Finally,

$$b \leq B^L, \qquad L \leq 1 + \frac{\log y}{c\eta} \ll \log\log b.$$

The full divisor set $\mathcal{D}(y)$ has size at least B^4, since $\log|\mathcal{D}(y)| \geq a_0 y/\log y \geq 4\lambda$. It splits into pairs $(d, 3d)$ with $d \mid P/3$. For $f(t) = 2t^2 + 1$ we have $f(3d) - 9f(d) = -8$. Thus at least one element of each pair satisfies $B \nmid f(t)$. Every divisor is nonzero modulo B since $B > y$. Some residue r therefore satisfies

$$r \neq 0, \quad 2r^2+1 \neq 0 \quad \text{in } \mathbb{F}_B, \qquad \mathcal{F} = \{u \in \mathcal{D}(y) : u \equiv r \pmod{B}\}, \qquad |\mathcal{F}| \geq \frac{|\mathcal{D}(y)|}{2B} \geq \frac{B^3}{2}.$$

Fix this $\mathcal{F}$ for all digits. Its size is at least $2B^{3/4}$, so at every cutoff $2 \leq x \leq B$ the number of primes bad for $\mathcal{F}$ is at most $C_{\text{bad}} y x^{3/4}$.

2.4 Uniform control of all digits

Let $s = \delta\lambda/\log\lambda$ and $T = \lfloor e^s \rfloor$. The bounds for λ imply, uniformly as $y \to \infty$, that $\lambda \to \infty$, $\lambda \leq y$, $s \to \infty$, and $s/\lambda \to 0$. Take y large enough that $m_0 \leq T < B$, $\log 2 \leq s$, and $s + \log 3 \leq \lambda$.

For $1 \leq m \leq T$ we have $\log(2m) \leq 2s$ and $\log\log(3m) \leq \log\lambda$. Hence $H(m) \leq 2K\delta\lambda < y$ for all sufficiently large y, the last inequality following from $\lambda/y \leq 2\eta/\log y$. The single-integer expansion of Section 2.2 now has $u, v \mid P$ and

$$\log\left(P(H(m))^2\right) \leq 4C_\vartheta K\delta\lambda \leq \lambda/4.$$

Every denominator w is at most N and at most $B^{1/4} < B$. In particular $B \nmid w$.

For an odd digit $T < m < B$, sieve the three-prime representations using $\mathcal{F}$ at cutoff $x = m$. The proportion of triples meeting a bad prime is at most

$$\frac{3C_{\text{bad}}}{v}\frac{y(\log m)^3}{m^{1/4}} \leq \frac{3C_{\text{bad}}}{v}\, y^4 \exp\!\left(-\frac{\delta\eta\, y}{4(\log y)^2}\right).$$

Indeed $m > \lfloor e^s \rfloor$ gives $m > e^s$, $\log m \leq \lambda \leq y$, and $s \geq \delta\eta y/(\log y)^2$. The last bound tends to zero independently of m and the chosen base. Above a fixed threshold it is less than $1/2$, giving three good odd primes for every such digit. For each of these primes repeat the covering step of Section 2.2 with factors in $\mathcal{F}$. We obtain $pw = u_1v_1 + u_2v_2 + 1$ with w positive, odd and at most N. Now $pw \equiv 2r^2 + 1 \not\equiv 0 \pmod{B}$, so $B \nmid w$. The three primes give nine terms in all.

If $T < m < B$ is even, represent the odd integer $m-1$ and append $1/1$. This uses at most J_0+1 terms when $m-1 \leq T$, and ten when $m-1 > T$. Zero uses no terms. All estimates above require only finitely many eventual inequalities in y, uniform for $\eta y/\log y < \lambda \leq 2\eta y/\log y$. Their thresholds can be fixed before individual digits are chosen. The two further requirements $y > R$ and $b \geq (2R)^{4/3}$ account for the dependence on R. This proves part (i).

2.5 Making the layers distinct

The odd reciprocal-basis theorem of van Albada and van Lint, in the tail form [5, Lemma 2.3], gives a finite set of odd integers exceeding any prescribed bound whose reciprocals sum to 1. Consequently we can choose arbitrarily many such sets, all of whose elements exceed 1, with each set disjoint from the preceding ones.

We first treat a list of at most J odd positive denominators, allowing repetitions. Before seeing the list, set $F_0 = 0$. At position $i = 1, \ldots, J$ choose $F_{i-1} + 1$ pairwise disjoint sets of odd multipliers, each with reciprocal sum 1. Let F_i be F_{i-1} plus the largest size of any set chosen at this position. Let $F = F_J$, and take R_* to be the largest multiplier in the entire finite collection. To replace an input $1/x$, choose one set S at its position and use the terms $1/(tx)$ for $t \in S$. Each previously used denominator d can forbid at most one choice of S, since $tx = d$ determines t uniquely. There are at most F_{i-1} previously used denominators. A choice remains, with no repetition within its own set either. Processing the list once gives at most F distinct denominators, each equal to its original denominator times an odd t with $1 < t \leq R_*$. These constants depend only on the already fixed J.

Now take $R \geq R_*$. In block (ε, j) a digit term uv/w contributes the reciprocal with denominator

$$b^{\varepsilon} B^j h_0, \qquad h_0 = \frac{N}{uv} w.$$

Since $uv \mid P^2 = N$, the integer h_0 is positive, odd and at most N^2. It is prime to B since $B \nmid Nw$. Apply the preceding replacement independently in each block. Its new coefficients $h = th_0$ still satisfy $B \nmid h$ since $t \leq R_* < B$, and $h \leq R_* N^2 < b$. Thus the exact B-adic valuation of each output denominator is j. At fixed j every denominator with $\varepsilon = 0$ is less than bB^j, whereas every denominator with $\varepsilon = 1$ is at least bB^j. Within each block the replacement already ensured distinctness. All denominators exceed 1 because $j \geq 1$. This proves part (ii).

In particular, two strings of L base-B digits with the normalizations in Theorem 2.1 give at most $2FL$ distinct odd unit fractions. The next section obtains those strings by Euclidean division.

3 Short representations with distinct odd denominators

We prove the uniform length bound for every fraction a/b with $b \geq 3$ odd and $1 \leq a < b$. Throughout this section a/b denotes the indicated quotient, with no coprimality assumption. This convention allows us to apply the digit theorem to the denominator produced by the recurrence below. We write log for the natural logarithm and $\log_2 x = \log x / \log 2$.

3.1 Expanding the quotient and remainder

Lemma 3.1 (small numerators). There are absolute constants b_S and $C_S > 0$ such that, whenever $b \geq b_S$ is odd and $1 \leq a \leq \sqrt{b}$, the rational number a/b is a sum of at most $C_S \log\log b$ distinct unit fractions with odd denominators greater than one.

Proof. Let J, F, R_* be the absolute constants of Theorem 2.1. Apply that theorem with $R = R_*$. After increasing its threshold to $b_S \geq e^e$, it supplies, for each eligible b, an odd prime B, an odd integer $N = P^2$ and an integer L satisfying

$$B \nmid bN, \qquad B > R_*, \qquad N \leq b^{1/8}, \qquad R_* N^2 < b, \qquad b \leq B^L, \qquad L \leq C_0 \log\log b$$

for an absolute C_0. Here P is the product of the odd primes up to y in Theorem 2.1, so $B > y$ gives $B \nmid N$. The choice of B works simultaneously for every digit below B.

Since $aN \leq b^{5/8} < b$, Euclidean division gives integers d, e with

$$aNB^L = bd + e, \qquad 0 \leq e < b \leq B^L, \qquad 0 \leq d < B^L.$$

Write $d = \sum_{i=0}^{L-1} d_i B^i$ and $e = \sum_{i=0}^{L-1} e_i B^i$, with $0 \leq d_i, e_i < B$. Dividing the last identity for aNB^L by bNB^L gives

$$\frac{a}{b} = \sum_{i=0}^{L-1} \frac{d_i}{NB^{L-i}} + \sum_{i=0}^{L-1} \frac{e_i}{bNB^{L-i}}. \tag{3.1}$$

The quotient and remainder construction in (3.1) follows Bloom and van Doorn [2, Lemma 7]. Apply Theorem 2.1(ii) to its two rows, with $m_{0,j} = d_{L-j}$ and $m_{1,j} = e_{L-j}$ for $1 \leq j \leq L$. The theorem supplies at most F odd denominators greater than one for each digit and gives distinctness across the entire union. The length is at most $2FL \leq 2FC_0 \log\log b$. This proves the lemma with $C_S = 2FC_0$. □

3.2 A finite increasing odd-denominator prefix

Lemma 3.2. There is an absolute integer b_* with the following property. For every odd $b \geq b_*$ and every $1 \leq a < b$ there are increasing odd integers $3 \leq q_0 < \cdots < q_{t-1}$ and integers $\alpha \geq 0$, $\beta > 0$ such that

$$\frac{a}{b} = \sum_{i=0}^{t-1} \frac{1}{q_i} + \frac{\alpha}{\beta}, \qquad \beta = b \prod_{i=0}^{t-1} q_i, \qquad t \leq 10 + \frac{\log\log b}{\log 2}.$$

If $\alpha > 0$, then $b^3 \leq \beta < b^6$, $\alpha \leq \sqrt{\beta}$, and $\alpha/\beta < 1/q_{t-1}$.

Proof. Start from $(\alpha_0, \beta_0) = (a, b)$ and $q_{-1} = 1$. At a state with $\alpha_i > 0$ and $\beta_i < b^3$, put $\rho_i = \alpha_i/\beta_i$. Let g_i be the least odd integer at least 3 and at least $1/\rho_i$, and define

$$q_i = \max\{g_i, q_{i-1} + 2\}, \qquad \alpha_{i+1} = \alpha_i q_i - \beta_i, \qquad \beta_{i+1} = \beta_i q_i. \tag{3.2}$$

Stop at the first t for which $\alpha_t = 0$ or $\beta_t \geq b^3$. With $\rho_i = \alpha_i/\beta_i$ also at the terminal state, (3.2) gives

$$0 \leq \rho_{i+1} = \rho_i - 1/q_i < \rho_i < 1, \qquad \beta_i = b \prod_{j<i} q_j.$$

All β_i are odd. At a nonterminal state, $\beta_i \geq 1/\rho_i$ and $\beta_i > q_{i-1}$, the latter from the product formula (and $b > 1$ initially). As β_i and q_{i-1} are odd, $\beta_i \geq q_{i-1} + 2$. Hence $q_i \leq \beta_i$ and $\beta_{i+1} \leq \beta_i^2$. The process stops in finite time since each step multiplies β_i by at least three.

Call a step ordinary if $q_i = g_i$. At an ordinary step, $1/\rho_i \leq q_i < 1/\rho_i + 2$, where the strict upper bound also holds for $q_i = 3$ because $\rho_i < 1$. Consequently

$$0 \leq \alpha_{i+1} < 2\alpha_i, \qquad \rho_{i+1} < \frac{2\rho_i^2}{1 + 2\rho_i} < 2\rho_i^2. \tag{3.3}$$

If the ordinary step has $q_i \geq 5$, then $q_i - 2 \geq 3$ is odd, and minimality gives $q_i - 2 < 1/\rho_i$. Thus $\rho_i < 1/(q_i - 2)$. It follows that

$$\rho_{i+1} < \frac{2}{q_i(q_i - 2)} < \frac{1}{q_i}. \tag{3.4}$$

If another step is needed, its greedy odd denominator is already strictly larger than q_i. It follows by induction that every subsequent step is ordinary.

Before the first ordinary step of denominator at least five, the algorithm takes successive terms of $3, 5, 7, \ldots$. Indeed, the first step is ordinary, and any subsequent step before that first ordinary denominator at least five is forced to take $q_{i-1} + 2$. At most six such initial terms can occur, since

$$\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} = \frac{46027}{45045} > 1,$$

whereas every partial sum is at most $a/b < 1$. Put $D_0 = 3 \cdot 5 \cdot 7 \cdot 9 \cdot 11 \cdot 13$. During these initial steps $\beta_i \leq D_0 b$ and $\alpha_i \leq D_0 a$ by (3.2). All later steps satisfy (3.3), giving the bound $\alpha_i \leq D_0 2^i a$ at every state.

If the process continues past seven steps, an ordinary step of denominator at least five has already occurred. By (3.4) and monotonicity $\rho_7 < 2/15$. Put $\kappa = \log(15/4)$. Iterating (3.3), for $7 \leq i < t$ we obtain

$$2\rho_i \leq \exp(-\kappa 2^{i-7}).$$

At such a state $\alpha_i \geq 1$ and $\beta_i < b^3$, whence $\rho_i > b^{-3}$. Taking

$$I = 7 + \left\lceil \log_2\!\left(\frac{3\log b + 1}{\kappa}\right)\right\rceil$$

shows that $t \le I$, since the two bounds for ρ_I would otherwise contradict each other. For $b \ge 3$ we have $\log b > 1$ and $\kappa > 1$, and therefore

$$2^t \le 2^I \le \frac{256(3\log b + 1)}{\kappa} \le 1024 \log b.$$

This gives the asserted bound on t, including the case $t \le 7$.

Choose $b_* \ge \max\{3, b_S\}$ so large that $b_*^2 > D_0$ and $1024 D_0 \log x \le \sqrt{x}$ for all $x \ge b_*$. If $\alpha_t > 0$, the preceding state has $\beta_{t-1} < b^3$, so

$$b^3 \le \beta_t \le \beta_{t-1}^2 < b^6, \qquad \alpha_t \le D_0 2^t a \le 1024 D_0 b \log b \le b^{3/2} \le \sqrt{\beta_t}.$$

During the initial segment the denominator is at most $D_0 b < b^3$. Thus the final step that crossed the threshold was ordinary with $q_{t-1} \ge 5$, and (3.4) gives $\alpha_t/\beta_t < 1/q_{t-1}$. Telescoping (3.2) proves the claimed identity. □

3.3 Completion of the short representation theorem

Theorem 3.3 (uniform upper bound). There is an absolute $C_H > 0$ such that, for every odd integer $b \ge 3$ and every integer $1 \le a < b$, there is a finite set S of distinct odd integers greater than one satisfying

$$\frac{a}{b} = \sum_{x \in S} \frac{1}{x}, \qquad |S| \le C_H(1 + \log\log b).$$

Proof. Suppose first that $b \ge b_*$. Apply Lemma 3.2. If the final numerator vanishes, the prefix gives the representation. Otherwise the unreduced denominator β is odd, $\beta \ge b^3 \ge b_S$, and $1 \le \alpha \le \sqrt{\beta}$. Apply Lemma 3.1 with input (α, β), choosing its auxiliary prime and its digits for this denominator. Every term of the resulting tail satisfies $1/x \le \alpha/\beta < 1/q_{t-1}$, hence $x > q_{t-1}$. The prefix and tail thus have disjoint denominators. Using $\beta < b^6$, the length is at most

$$10 + \frac{\log\log b}{\log 2} + C_S \log\log \beta \le 10 + C_S \log 6 + \left(\frac{1}{\log 2} + C_S\right)\log\log b.$$

The same upper bound covers a vanishing tail.

There are only finitely many pairs with b odd, $3 \le b < b_*$ and $1 \le a < b$. Fix an odd integer $D \ge b_*$. For each such pair the large-denominator result just proved applies to $(Da)/(Db) = a/b$ since Db is odd and $1 \le Da < Db$. Let $M_0 \ge 1$ bound the lengths obtained for these finitely many pairs. As $\log\log b > 0$ for $b \ge 3$, we may take

$$C_H = \max\left\{M_0,\ 10 + C_S \log 6,\ (\log 2)^{-1} + C_S\right\}.$$

All constants are fixed independently of a and b. □

4 Divisibility of the largest denominator

Let p_j be the jth odd prime and put $Q_r = \prod_{j=1}^{r} p_j$. Write $K(r)$ for the shortest length of a representation of 1 by distinct odd denominators greater than 1 whose largest denominator is divisible by Q_r. Write $K_\omega(r)$ for the corresponding minimum when the largest denominator may have any r or more distinct prime factors. We prove Theorem 1.2 in the form

$$c \log r \;\le\; K_\omega(r) \;\le\; K(r) \;\le\; C \log r \qquad (r \text{ sufficiently large}),$$

with absolute positive constants c, C. The upper bound uses Theorem 1.1 to represent $1 - 1/Q_r$. The point is to retain divisibility by Q_r while making the largest denominator exceed every denominator in that representation.

4.1 Splitting one denominator and avoiding a finite set

For odd $u \ge 3$, define

$$U(u) = \begin{cases} (u+1)(u+2)/2, & u \equiv 1 \pmod 4, \\ u(u+3)/2, & u \equiv 3 \pmod 4, \end{cases} \quad V(u) = \begin{cases} uU(u), & u \equiv 1 \pmod 4, \\ (u+2)U(u), & u \equiv 3 \pmod 4. \end{cases}$$

The half factors $(u+1)/2$ and $(u+3)/2$ are odd in their respective rows. Both rows give $u < u+2 < U(u) < V(u)$ and $u \mid V(u)$. In the first row the sum of the last two reciprocals is $(u+1)/(uU(u))$, and in the second it is $(u+3)/((u+2)U(u))$. Both expressions equal $2/(u(u+2))$. Thus

$$\frac{1}{u} = \frac{1}{u+2} + \frac{1}{U(u)} + \frac{1}{V(u)}.$$

For a proper positive divisor d of an odd integer $n \ge 3$, set $u = n/d$ and form the triple

$$T_n(d) = \{d(u+2),\ dU(u),\ dV(u)\}.$$

Its three denominators are odd, distinct and greater than n. Their reciprocals sum to $1/n$. The smallest is $n + 2d$ and the largest is a multiple of n.

Lemma 4.1. Let F be a finite set of positive integers. At least $\tau(n) - 1 - 2|F|$ proper divisors d of n satisfy $T_n(d) \cap F = \varnothing$, where $\tau(n)$ denotes the number of positive divisors of n. For these divisors the sets $F \cup T_n(d)$ are pairwise distinct.

Proof. For fixed n write $X_1(u), X_2(u), X_3(u)$ for the displayed coordinates of $T_n(d)$ in increasing order. The first is $n(1 + 2/u)$, which is strictly decreasing in u and at most $5n/3$. Direct substitution gives

$$\frac{n}{2}(u+3) \le X_2(u) < \frac{n}{2}(u+5), \qquad \frac{n}{2}(u+1)(u+2) \le X_3(u) \le \frac{n}{2}(u+2)(u+3).$$

These intervals show that X_2 and X_3 are each strictly increasing on the odd arguments in question. Indeed the next possible argument v is at least $u+2$, and $(u+2)(u+3) < (v+1)(v+2)$. Also $X_2(u) \ge 3n$. An element of F below $3n$ can therefore occur only as X_1 and excludes at most one divisor. An element at least $3n$ excludes at most one divisor through each of X_2 and X_3. This proves the bound. For an admissible divisor, removing F from the union leaves a triple with smallest element $n + 2d$, which recovers d. □

4.2 A maximum divisible by the odd primorial

Let $Q \ge 3$ be odd, and suppose that a set E of m distinct odd integers greater than 1 satisfies

$$\sum_{e \in E} \frac{1}{e} = 1 - \frac{1}{Q}, \qquad \tau(Q) > 8m + 7.$$

Keep E fixed. Make a list consisting of Q^2 and one entry Qe for each $e \in E$, allowing repeated entries in this list. Its reciprocal sum is

$$\frac{1}{Q^2}+\frac{1}{Q}\sum_{e\in E}\frac{1}{e}=\frac{1}{Q}.$$

Replace the entries of the list one at a time. To replace an entry n, take F to consist of E together with all triples already chosen. At this stage $|F|\le m+3m=4m$. Since $Q\mid n$, every divisor of Q is a divisor of n, so $\tau(n)-1-2|F|>0$. Lemma 4.1 gives a triple $T_n(d)$ disjoint from F. This procedure also applies when the list contains two copies of Q^2.

After all $m+1$ replacements the union of E and the chosen triples represents 1 and has $4m+3$ distinct denominators. For each $e\in E$ the triple replacing Qe has largest element greater than $Qe>e$. Consequently the largest denominator of the union is the largest element of one of the triples. Each such largest element is divisible by its parent n, hence by Q.

Now take $Q=Q_r$. Its squarefreeness gives $\tau(Q_r)=2^r$. Bertrand's postulate, iterated from the prime 2, gives $p_j\le 2^{j+1}$. In particular

$$\log Q_r\le r(r+1)\log 2\le r^2\quad(r\ge 3),\qquad \log\log Q_r\le 2\log r.$$

Apply Theorem 1.1 with numerator Q_r-1 and denominator Q_r. It gives a set E as above with $m\le C_H(1+\log\log Q_r)=O(\log r)$. Thus $2^r>8m+7$ for all sufficiently large r. The construction proves $K(r)\le 4m+3\ll\log r$.

The same E also gives a short representation containing Q_r itself. If $Q_r\notin E$, add Q_r to E. If $Q_r\in E$, keep that entry and insert a triple representing $1/Q_r$ which avoids E, using Lemma 4.1 with $n=Q_r$. The resulting length is at most $m+3$.

4.3 The lower bound and adjustment of the length

Lemma 4.2 (denominator height). If $1=\sum_{i=1}^{k}1/x_i$ with positive integers $x_1\le\cdots\le x_k$, then $x_i\le k^{2^{i-1}}$ for $1\le i\le k$. Repeated denominators are allowed.

Proof. Put $P_0=1$ and $P_j=x_1\cdots x_j$. For $0\le j<k$ the residual sum is positive and its product with P_j is an integer. Therefore

$$\frac{1}{P_j}\le 1-\sum_{i=1}^{j}\frac{1}{x_i}\le\frac{k-j}{x_{j+1}},\qquad P_{j+1}\le kP_j^2.$$

Induction gives $P_j\le k^{2^j-1}$ for $0\le j\le k$. The same residual inequality now gives $x_i\le kP_{i-1}\le k^{2^{i-1}}$ for every i. □

If x_k is odd and has at least r distinct prime factors, then $3^r\le x_k$. Hence Lemma 4.2 gives $r\log 3\le 2^{k-1}\log k$. Since $\log 3>1$, $2^{k-1}\le e^k$ and $\log k\le e^k$, it follows that $k\ge(\log r)/2$. This proves the lower bound for $K_\omega(r)$. Divisibility by Q_r supplies r distinct prime factors, so $K_\omega(r)\le K(r)$, completing the proof of Theorem 1.2.

For later use, take a representation of 1 by distinct odd denominators and replace its largest denominator n by $T_n(1)$. The length increases by two and the new largest denominator $V(n)$ is strictly larger than n and divisible by it. The map $n\mapsto V(n)$ also preserves strict order on odd integers $n\ge 3$. Indeed $V(n)/n$ lies between $(n+1)(n+2)/2$ and $(n+2)(n+3)/2$, and the latter is strictly smaller than the lower endpoint for $n+2$. The last three denominators of the new representation are exactly $n+2, U(n), V(n)$. Deleting these three entries and subtracting two from the smallest of them recovers the entry to restore, so this operation is injective. Multiplication of the original identity by the odd product of its denominators shows modulo 2 that its length is odd. Iteration therefore attains every larger odd length while preserving every specified divisor of the original maximum.

5 Counting at a prescribed length

5.1 Two variable denominators

Lemma 5.1. Let $N \geq 3$ be odd and let $Q \mid N$. For every divisor $d > 1$ of Q there is a decomposition

$$\frac{1}{N} = \frac{1}{N+2} + \frac{1}{v_d} + \frac{1}{w_d}, \qquad N+2 < v_d < w_d,$$

with v_d, w_d odd and $N \mid w_d$. As d varies, the largest denominators w_d are distinct.

Proof. Write $A = N(N+2)$. The factors in each pair $\{d, A/d\}$ are odd and have product 3 modulo 4. Exactly one, say q, is 1 modulo 4. Both exceed one because $1 < d \leq N$ and $A/d \geq N+2$. Put $t = A/q$, $h = (q+1)/2$, and set

$$v_d = th, \qquad w_d = Ah.$$

The integer h is odd. Also $2v_d = A + t > A > 2(N+2)$ and $w_d = qv_d > v_d$. The identity follows from

$$\frac{1}{v_d} + \frac{1}{w_d} = \frac{q+1}{Ah} = \frac{2}{A} = \frac{1}{N} - \frac{1}{N+2}.$$

The smaller member of $\{d, A/d\}$ is d. Two such factor pairs sharing an element coincide, so different d give different q and thus different w_d. Each w_d is a multiple of N. □

This is a complementary-divisor version of [6, Lemma 3.1], where the fixed denominator is also $N+2$. For example, at $N = Q = 15$, the choices $d = 3, 5, 15$ give $(v_d, w_d) = (129, 10965), (153, 765), (135, 2295)$, respectively. In all three decompositions the first term is $1/17$.

Proof of Theorem 1.3. Fix $B_0 > 0$ such that $K(r) \leq B_0 \log r$ for all sufficiently large r. For a given large odd k, choose $r = \lfloor \exp(k/(2B_0)) \rfloor$. Theorem 1.2 gives an odd representation of length at most $k/2$ whose largest denominator is divisible by Q_r. Its length is odd. Use the canonical split of Section 4 repeatedly on its largest denominator to reach exactly $k-2$ terms. Each new maximum is a multiple of the old one. Write N for the resulting maximum, so $Q_r \mid N$.

Replace $1/N$ by the three terms of Lemma 5.1 with $Q = Q_r$. The $k-3$ old denominators below N, followed by $N+2$, form the first $k-2$ denominators in every resulting representation. The $\tau(Q_r) - 1 = 2^r - 1$ choices have distinct maxima, each divisible by Q_r. With $c = 1/(4B_0)$, for all sufficiently large k we have $r \geq \lfloor e^{ck} \rfloor$ and $\log\log(2^r - 1) = k/(2B_0) + O(1) > ck$. This gives the required lower bound for the family. Lemma 4.2 bounds its ambient count by

$$S_o(k) \leq \prod_{i=1}^{k} k^{2^{i-1}} = k^{2^k - 1}.$$

Taking two logarithms proves $\log\log S_o(k) \asymp k$ along odd k. □

5.2 Many representations for each of many largest denominators

Proof of Corollary 1.4. Apply Theorem 1.3 at length m, where m is a sufficiently large odd integer. Let P be the set of $m-2$ common denominators, let $u = \max P$, and number the T endings (v_i, w_i) in increasing order of w_i. The v_i are pairwise distinct too, since $1/v_i + 1/w_i$ is fixed. Split the indices into the lowest $\lfloor T/2 \rfloor$ and the remaining highest indices.

For a low index i and a high index j, replace $1/u$ in representation i by the reciprocals of the m numbers

$$D_j = \{uz : z \in P \cup \{v_j, w_j\}\}.$$

Their reciprocals sum to $1/u$. The scaling operation appears for a terminal term in [6, equation (4.1)]. Here u is the last term of the common prefix. Every element of D_j exceeds u, so a collision can occur only with v_i or w_i. For fixed j, each element of D_j rules out at most one i in each coordinate. At least $\lfloor T/2 \rfloor - 2m$ low indices therefore succeed. From the resulting set one recovers $\{v_i, w_i\}$ by removing the fixed background $(P \setminus \{u\}) \cup D_j$. These representations are distinct and all have maximum uw_j, since $w_i < w_j$. The $\lceil T/2 \rceil$ high indices give different maxima. Each retains the prescribed prime factors of w_j.

The new length is $2m - 1$. For a given large odd k, take the largest odd $m \le (k + 1)/2$. If $k = 2m + 1$, apply the canonical split of Section 4 once to the largest denominator in each representation. At a fixed maximum every member receives the same three larger denominators, so all representations in that family remain distinct. The largest-denominator update is strictly increasing on odd integers and preserves multiples of Q_r, so the families for distinct maxima remain distinct as well. The case $k = 2m - 1$ needs no additional step.

For some $c_0 > 0$, the input family has $T \ge \exp(\exp(c_0 m))$ members and its maxima are divisible by $Q_{\lfloor \exp(c_0 m) \rfloor}$. For large m both $\lceil T/2 \rceil$ and $\lfloor T/2 \rfloor - 2m$ are at least $T/4$. Since $m = k/2 + O(1)$, any fixed $0 < c < c_0/2$ gives the bounds of the corollary for sufficiently large k. □

5.3 Sharpness of the short representation bound

We conclude the proof of Theorem 1.1 using the pairing argument of Erdős [7, Theorem 2 and p. 209], also given in [2, Appendix A.2]. Set $L_o(0, b) = 0$. For $0 \le a < b$, combine shortest representations for a/b and $(b - 1 - a)/b$ and append $1/b$. The resulting list sums to one and may have repetitions. If $n = L_o(a, b) + L_o(b - 1 - a, b)$, then the list has length $n + 1$ and contains the denominator b. Lemma 4.2 gives $b \le (n + 1)^{2^n}$. Since $n \ge 1$ and $b \ge 3$,

$$\log\log b \le n \log 2 + \log\log(n + 1) \le 2n.$$

In particular $L_o(b - 1, b) \ge \frac{1}{2} \log\log b$. Call a numerator a long if $L_o(a, b) \ge \frac{1}{4} \log\log b$. Under the involution $a \mapsto b - 1 - a$, every numerator which is not long is paired with a long one. Thus there are at least $\lceil b/2 \rceil = (b + 1)/2$ long numerators among $0, \ldots, b - 1$. All are positive, since $\log\log b > 0$. This gives the lower bounds for the maximum and the mean at each odd b. The upper bounds follow from Theorem 3.3. □

A Scope of the Lean verification

The formal project uses Lean 4.36.0. Its definition of a representation is a finite set of odd natural numbers greater than one, with the sum of their reciprocals taken in the rationals. The cardinality of the set is the length. The counting function is the number of such finite sets representing one at the prescribed length. Sorting gives the tuples used in this paper.

The analytic input is explicit. The proposition `ThreePrimeInput` says that, for every sufficiently large odd n, there are $\gg n^2/(\log n)^3$ ordered representations by three odd primes, with repetitions allowed. Vinogradov's Theorem 2 gives an asymptotic for the unweighted ordered count with all primes, with singular series bounded below uniformly on odd n. At most three ordered triples contain 2, since they must be permutations of $(2, 2, n - 4)$. Thus the stated input follows after increasing the threshold. The Lean project verifies the deduction from this input and includes reductions from weighted formulations. The underlying three-prime estimate is cited in the mathematical proof [1, Theorem 2, p. 194], [2, proof of Lemma 6].

The declarations are in the `OddEgyptian` namespace. The theorem `englishManuscript_of_threePrimeInput` proves `EnglishManuscriptResults` from the prime-counting input. Its six fields give the pointwise, maximum and mean length bounds of Theorem 1.1, the bounds on the least length with a prescribed divisor of the maximum in Theorem 1.2, the common-prefix family and total count in Theorem 1.3, and the multiplicity statement in Corollary 1.4. The construction in Section 4 also gives the minimum-length bound when the largest denominator may contain any r distinct primes.

In `IsPrefixFamily`, a fixed set P of $k-2$ denominators lies strictly below the two varying terms, and the members of a second finite set index distinct maxima. In `HasMaximumFamilies`, every maximum has its own finite set of distinct k-term representations of one whose largest denominator is exactly that maximum. The same integer ceiling of $\exp(\exp(ck))$ bounds both cardinalities from below. These definitions retain divisibility by the product of the first $\lfloor \exp(ck) \rfloor$ odd primes.

Acknowledgement. AI assistance was used in developing the proofs, drafting the manuscript and writing the Lean formalization.

References

[1] I. M. Vinogradov. Some theorems concerning the theory of primes. *Matematicheskii Sbornik* (N.S.) **2(44)** (1937), 179–195. Original article

[2] T. F. Bloom and W. van Doorn. Optimal bounds for the length and number of unit fraction representations. Preprint, 2026. Version 1

[3] OpenAI. Short Egyptian fractions. Preprint, September 2026. Authors' manuscript

[4] D. Hart and A. Iosevich. Sums and products in finite fields: an integral geometric viewpoint. Preprint, 2007. arXiv:0705.4256

[5] C. Elsholtz. Egyptian fractions with odd denominators. *Quarterly Journal of Mathematics* **67** (2016), 425–430. Preprint

[6] Y.-G. Chen, C. Elsholtz and L.-L. Jiang. Egyptian fractions with restrictions. *Acta Arithmetica* **154** (2012), 109–123. Publisher's version

[7] P. Erdős. Az $1/x_1 + \cdots + 1/x_n = a/b$ egyenlet egész számú megoldásairól. *Matematikai Lapok* **1** (1950), 192–210. Original article