hexagon:2611.00781v1

# Asymptotically full minimal signed unit-fraction identities

An alternative proof for Erdős Problem 319

Jensen Kohlmeyer `jenw1n`
Liam Kruer `lkruer`
Will Howe `H0W4RD-dev`
Grant Huffman `Ghuffy1`

16 October 2026

**Abstract**

Let $f(N)$ be the largest size of a subset of $\{1,\dots,N\}$ admitting a signed zero sum of its reciprocals with no proper nonempty signed zero subsum. We prove $f(N) = N - o(N)$, using the subset-sum theorem of Liu and Sawhney. The construction joins finitely many large-prime groups by one CRT-chosen unit fraction. The resulting positive sum can be arbitrarily large, although the density cost of excluding those primes is arbitrarily small. A dense coprime-denominator expansion of that sum provides the opposite side and preserves minimality. This is a complete candidate proof, self-checked but not independently reviewed. Earlier public work already claims this asymptotic and the protected-prime construction used here. This version gives a route through Liu–Sawhney and an accompanying Lean formalization; no first-discovery or formalization-priority claim is made. Substantial AI assistance was used.

## 1 Statement

For a finite set $B \subset \mathbb{N}$, write $\mathcal{R}(B) = \sum_{n\in B} 1/n$. Call $A \subseteq \{1,\dots,N\}$ primitive if there are signs $\delta_n \in \{-1,1\}$ satisfying

$$\sum_{n\in A} \frac{\delta_n}{n} = 0, \qquad \sum_{n\in C} \frac{\delta_n}{n} \neq 0 \quad \text{for every } \varnothing \neq C \subsetneq A.$$

Let $f(N)$ be the maximum of $|A|$ over these sets.

**Theorem 1.** *For every $\varepsilon > 0$ and every sufficiently large $N$, there is a primitive $A \subseteq \{1,\dots,N\}$ with $|A| \geq (1-\varepsilon)N$. Consequently*

$$\boxed{\lim_{N\to\infty} \frac{f(N)}{N} = 1.}$$

The lower bound $(1-1/e+o(1))N$ follows from a dense expansion of 1 and the singleton $\{1\}$ [3]. Earlier public candidate proofs of the density-one asymptotic are recorded in [4, 5]; their existence precludes treating the present result as an uncontested new solution. Our positive side is a fixed, possibly enormous, finite set once $\varepsilon$ is chosen. Only the opposite side grows with $N$. The result is asymptotic and supplies no practical threshold for $N$.

## 2 A finite positive set with only one coprime subsum

For a squarefree integer $D$, call a rational number $D$-integral if its reduced denominator is coprime to $D$. This includes 0 and all integers. The $D$-integral rationals form an additive subgroup of $\mathbb{Q}$.

**Lemma 2.** *Let $K \geq 1$, $L = \operatorname{lcm}(1, \ldots, K)$, and $B = LH_K$, where $H_K = \sum_{m=1}^{K} 1/m$. Let $\mathcal{P}$ be a finite set of at least two primes, each greater than $B$, and put $D = \prod_{p \in \mathcal{P}} p$. There exists a positive integer $t$ coprime to $D$ such that the set*

$$E = \{pm : p \in \mathcal{P},\ 1 \leq m \leq K\} \ \cup\ \{Dt\}$$

*has the following properties:*

1. *All its denominators are distinct, and all have a prime factor in $\mathcal{P}$.*
2. *$r := \mathcal{R}(E)$ is $D$-integral and*
$$r = H_K \sum_{p \in \mathcal{P}} \frac{1}{p} + \frac{1}{Dt}.$$
3. *The only subsets $C \subseteq E$ for which $\mathcal{R}(C)$ is $D$-integral are $\varnothing$ and $E$.*

*Proof.* Since $p > B \geq K$, both $L$ and $B$ are invertible modulo every $p \in \mathcal{P}$. Choose $t$ by the Chinese remainder theorem so that

$$tB(D/p) + L \equiv 0 \pmod{p} \qquad (p \in \mathcal{P}). \tag{1}$$

The prescribed residue $-L\big(B(D/p)\big)^{-1}$ is nonzero modulo $p$, so $\gcd(t, D) = 1$. Take its positive representative, or add a multiple of $D$.

The groups $\{pm : m \leq K\}$ are disjoint: a denominator in the $p$-group has no other prime factor in $\mathcal{P}$. The connector $Dt$ has all these prime factors, so is distinct from every group denominator.

For a subset $C \subseteq E$, let

$$J_p = \{m \leq K : pm \in C\}, \qquad b_p = \sum_{m \in J_p} L/m,$$

and let $h \in \{0, 1\}$ indicate whether $Dt \in C$. Every $b_p$ is an integer in $[0, B]$, with $b_p = 0$ exactly for an empty $J_p$, and $b_p = B$ exactly for a full $J_p$. Multiplying $\mathcal{R}(C)$ by the common denominator $DtL$ gives the integer

$$U_C = t \sum_{p \in \mathcal{P}} (D/p) b_p + hL. \tag{2}$$

Each $p \in \mathcal{P}$ occurs to exponent exactly one in $DtL$. Thus $\mathcal{R}(C)$ is $D$-integral if and only if every such $p$ divides $U_C$.

If $h = 0$, reduction modulo $p$ gives $b_p \equiv 0 \pmod{p}$. As $0 \leq b_p \leq B < p$, this forces $b_p = 0$ for every $p$, hence $C$ empty. If $h = 1$, subtracting (1) from the condition $p \mid U_C$ gives $b_p \equiv B \pmod{p}$. The same bound forces $b_p = B$ for every $p$, hence $C = E$. Finally, (1) directly makes $U_E$ divisible by every $p$, so $r$ is $D$-integral. The formula for $r$ is immediate. □

**Lemma 3.** *For every $0 < \delta < 1/4$ and $R > 0$, there are a squarefree $D$ and a finite set $E$ as in Lemma 2 such that*

$$\rho := \frac{\varphi(D)}{D} \geq 1 - 2\delta, \qquad r = \mathcal{R}(E) \geq R.$$

*Proof.* Choose $K$ so that $H_K \geq R/\delta$, and then choose primes $p > \max\{LH_K, 1/\delta\}$ until

$$\delta \leq \sum_{p \in \mathcal{P}} \frac{1}{p} < 2\delta.$$

Such a finite choice exists because the sum of reciprocal primes diverges. Each summand is less than $\delta$, so at least two primes are chosen. The finite product inequality gives

$$\rho = \prod_{p \in \mathcal{P}} (1 - 1/p) \geq 1 - \sum_{p \in \mathcal{P}} 1/p > 1 - 2\delta.$$

Lemma 2 gives $r \geq H_K \delta \geq R$. □

## 3 Dense expansions avoiding a fixed set of primes

The only deep input is Proposition 5.2 of Liu–Sawhney [1]. For transparency, its hypotheses are recorded in the appendix. The following application also explains why fixed prime factors in the target denominator cause no problem.

**Lemma 4.** *Fix a squarefree $D$ and a positive $D$-integral rational $r$. Put $\rho = \varphi(D)/D$. For every $0 < \kappa < \min(r/2, 1/2)$, set $c = \exp(-(r+\kappa)/\rho)$. For all sufficiently large $N$, there is a set*

$$F_N \subseteq \{n \in [cN, N] : \gcd(n, D) = 1\}$$

*with $\mathcal{R}(F_N) = r$ and*

$$|F_N| \geq \big(\rho(1-c) - \kappa + o(1)\big)N. \tag{3}$$

*In particular,*

$$\liminf_{N \to \infty} \frac{\max |F_N|}{N} \geq \rho(1 - e^{-r/\rho}),$$

*where the maximum allows all coprime denominators at most $N$.*

*Proof.* Let $W_N = \{n \in [cN, N] : \gcd(n, D) = 1\}$. Since $D, c$ are fixed, elementary counting in residue classes gives

$$|W_N| = \rho(1-c)N + O_D(1), \qquad \mathcal{R}(W_N) = \rho \log(1/c) + o(1) = r + \kappa + o(1). \tag{4}$$

Set

$$M_0 = \min(c/2, 10^{-5})N, \quad u = (\log N)^{4/5}, \quad K_0 = M_0 e^{-u}, \quad S = N e^{-6u}, \quad \eta = (\log\log N)^{-2}.$$

The endpoint can be rounded to an integer. Remove from $W_N$ all numbers that are not $S$-smooth and all numbers with $\Omega(n) > 5 \log\log N$. Lemmas 2.2–2.3 of [1] show that this removes $O(N/(\log N)^{1/5}) + O(N/(\log N)^3) = o(N)$ elements. Because these numbers are at least $cN$, their reciprocal mass is $o(1)$.

Next, while a prime power $q$ occurring in the current set satisfies $q\mathcal{R}(W_q) < \eta$, remove all its multiples. Each prime power is used at most once, so the total reciprocal mass removed is at most

$$\eta \sum_{q \leq N} \frac{1}{q} = O(1/\log\log N) = o(1).$$

Call the final set $W_N^*$, and put $Q = \mathrm{lcm}(W_N^*)$. We have

$$\mathcal{R}(W_N^*) = r + \kappa + o(1), \qquad q\mathcal{R}((W_N^*)_q) \geq \eta \tag{5}$$

for every prime power occurring in $W_N^*$.

Write $r = a/b$ in lowest terms. For each prime power $q \mid b$, $\gcd(q, D) = 1$. Again by residue-class counting,

$$\mathcal{R}((W_N)_q) = \frac{\rho}{q}\log(1/c) + o(1) > 0.$$

Here $q$ is fixed. Since the entire cleanup removes only $o(1)$ reciprocal mass, some multiple of $q$ remains in $W_N^*$. Therefore $b \mid Q$. In particular,

$$x = Q\big(\mathcal{R}(W_N^*) - r\big)$$

is an integer, positive for large $N$, and $x/Q = \kappa + o(1)$.

Apply Proposition 5.2 with $d = 1/5$, $e = 1/100$, and the displayed $M_0, K_0, S, \eta$. All its size conditions hold for large $N$. For the last condition, the second term in $G$ gives

$$G \gg_c (\log N)^{3/5}(\log\log N)^{-9},$$

and hence

$$\frac{G^2}{(\log N)^{2e}(\log(N/M_0) + (\log N)^{4/5})} \gg_c \frac{(\log N)^{19/50}}{(\log\log N)^{18}} \longrightarrow \infty.$$

The bounds on $S$ hold because its factor $e^{-6u}$ dominates the $e^{-2u}$ and polylogarithmic factors in those bounds; $S \geq N^{.99}$ eventually. The target condition follows from (5) and $\kappa > 0$:

$$(1 + 1/\log N)x/Q \leq \mathcal{R}(W_N^*) \leq (\log N)x/Q.$$

The proposition supplies $G_N \subseteq W_N^*$ with reciprocal sum $x/Q = \mathcal{R}(W_N^*) - r$. Take $F_N = W_N^* \setminus G_N$, so that $\mathcal{R}(F_N) = r$. This applies the proposition to a small complementary target; its sampling proportion is less than $1/2$ for large $N$. Finally every term in $W_N \setminus F_N$ is at most $N$, so

$$|W_N \setminus F_N| \leq N\mathcal{R}(W_N \setminus F_N) = N(\kappa + o(1)).$$

Together with (4) this proves (3). Taking the lower limit in $N$ and then letting $\kappa \downarrow 0$ gives the final assertion. No uniformity as $D, r, \kappa$ vary is required. □

# 4 Primitivity and completion of the asymptotic

*Proof of Theorem 1.* Fix $0 < \delta < 1/4$ and use Lemma 3 with $R = \log(1/\delta)$. This gives fixed $D, E, r$ and $\rho \geq 1 - 2\delta$, $r \geq \log(1/\delta)$. By Lemma 4, choose $F_N$ with $\mathcal{R}(F_N) = r$ and all denominators coprime to $D$. For large $N$, every element of the fixed set $E$ is at most $N$. Also $E \cap F_N = \varnothing$, since every element of $E$ has a prime factor dividing $D$.

Give $E$ positive signs and $F_N$ negative signs. The total is zero. Suppose $C \subseteq E \cup F_N$ is a signed zero subset, and put $E' = C \cap E$, $F' = C \cap F_N$. Then

$$\mathcal{R}(E') = \mathcal{R}(F').$$

The right side is $D$-integral. Lemma 2 therefore forces $E' = \varnothing$ or $E' = E$. In the first case positivity forces $F' = \varnothing$. In the second, $\mathcal{R}(F') = r = \mathcal{R}(F_N)$ and positivity forces $F' = F_N$. Thus the signed identity is primitive.

Its cardinality and Lemma 4 give

$$\liminf_{N\to\infty} \frac{f(N)}{N} \geq \rho(1-e^{-r/\rho}) \geq (1-2\delta)(1-\delta) \geq 1-3\delta.$$

The middle inequality uses $\rho \leq 1$ and $r/\rho \geq r \geq \log(1/\delta)$. Let $\delta \downarrow 0$, and combine this with $f(N) \leq N$. This proves the asserted limit and the quantified first statement. □

For a small concrete example, $K = 2$ and $\mathcal{P} = \{5, 7, 11\}$ give $D = 385$, $t = 73$, and the primitive identity

$$\frac{1}{5}+\frac{1}{10}+\frac{1}{7}+\frac{1}{14}+\frac{1}{11}+\frac{1}{22}+\frac{1}{28105} = \frac{95}{146} = \frac{1}{2}+\frac{1}{8}+\frac{1}{39}+\frac{1}{22776}.$$

The four denominators on the right are coprime to 385. This illustrates minimality, not the asymptotic density estimate.

## 5 Verification and attribution

The protected-prime construction, the amplification of its reciprocal mass at small prime-density cost, and the final primitivity argument already appear in the public manuscript of `g-01234`, dated 16 July 2026 [4]. A later submission by `peilinliu66-dev`, opened on 3 October 2026, claims a complete Lean proof using a thinned-prime inverse-sum completion route [5]. That submission was pending when checked on 16 October 2026; we have not independently reproduced it. These earlier claims must be considered in any assessment of priority or prize eligibility. The present dense-expansion lemma instead applies Liu–Sawhney's Proposition 5.2. No novelty claim is made for the protected-prime idea. The Fourier analysis is due to Liu–Sawhney and their predecessors. The attached exact audit checks small global CRT examples exhaustively and checks a larger example locally at every selected prime. These finite checks do not replace the proof of the asymptotic or certify the imported analytic theorem.

The accompanying file `Erdos319.lean` proves the limit and explicit fixed-sign witnesses without additional mathematical hypotheses. Its analytic estimates include ports of existing proofs from `plby/lean-proofs`, commit `8822f7ddef30fadbd92e1c6ab4ed897af356af5e`; those background proofs retain their separate attribution. The checked environment is Lean 4.35.0-rc2 with Mathlib commit `065356127b1dc0016f66b7283ce0ce2c4055aa55`. The main theorem is

`Erdos319Formal.erdos_problem_319`.

The original 1980 statement was checked at printed p. 43 [2]: the same signs are used in the whole sum and in every subsum. The conclusion determines the leading asymptotic, not the lower-order deficit or the exact maximum for each finite $N$.

The listed team prepared this AI-assisted proof presentation and formalization for review. Team credit does not assert that each member manually wrote every proof step. No independent human review, prize acceptance, or entitlement to payment is asserted.

## A The precise Liu–Sawhney input

Here is Proposition 5.2 of [1], restated with renamed parameters. For $W \subseteq [1, N]$, let $W_q = \{n \in W : q \mid n\}$ and let $Q_W = \operatorname{lcm}(W)$. There is an absolute $C$ such that, for fixed $d \in (0, 1)$, $e \in (0, 1/10)$ and sufficiently large $N$, the following holds. Suppose $\eta \geq 1/\log N$ and

$$N^{.99} \leq S \leq K_0 \leq M_0 \leq N/10^4, \qquad K_0 \leq M_0 \exp(-(\log N)^{1-d}),$$

$$S \le \min\left\{ \frac{M_0^2}{CN}, \frac{\eta M_0 K_0^2}{N^2 (\log N)^3} \right\}.$$

Define

$$G = \max\left\{ \frac{\eta}{(\log N)^d (\log\log N)^3}, \frac{\eta^2 (\log N)^{1-2d}}{\log(N/M_0)^2 (\log\log N)^5} \right\}.$$

Assume

$$\frac{G^2}{(\log N)^{2e}(\log(N/M_0) + (\log N)^{1-d})} \ge C.$$

Let $W \subseteq [M_0, N]$ satisfy: every prime power dividing an element of $W$ is at most $S$; $\Omega(n) \le 5 \log\log N$ for $n \in W$; and $q\mathcal{R}(W_q) \ge \eta$ for every prime power occurring in $W$. For every positive integer $x$ with

$$(1 + 1/\log N) x/Q_W \le \mathcal{R}(W) \le (\log N) x/Q_W,$$

there is $W' \subseteq W$ such that $\mathcal{R}(W') = x/Q_W$.

In the dense-expansion lemma all parameters other than $N$ are fixed or explicitly specified functions of $N$. In particular, the proof does not use the inconsistent displayed sign of $K$ in the arXiv version's proof of Theorem 1.3; it uses the proposition's stated $K_0 \le M_0 \exp(-(\log N)^{1-d})$ condition directly.

## References


[1] Y. P. Liu and M. Sawhney, *On further questions regarding unit fractions*, arXiv:2404.07113v1 (2024), Proposition 5.2 and Lemmas 2.2–2.3. International Mathematics Research Notices (2026), rnaf382; published article.

[2] P. Erdős and R. L. Graham, *Old and new problems and results in combinatorial number theory* (1980), p. 43; original scan.

[3] T. F. Bloom, Erdős Problem 319.

[4] `g-01234`, *A Candidate Proof That $c(N)/N \to 1$ in Erdős Problem 319*, circulation draft, 16 July 2026; pinned manuscript.

[5] `peilinliu66-dev`, *JSP-000265: complete Erdős 319 fixed-sign density-one Lean proof*, submission opened 3 October 2026; awards PR 4669; selected proof revision.