Proceedings of the Royal Society of Edinburgh, **155**, 2285–2295, *2025*
DOI:10.1017/prm.2024.42

Unit fractions with shifted prime denominators

Thomas F. Bloom
Mathematical Institute, Woodstock Road, Oxford, OX2 6GG, United Kingdom (bloom@maths.ox.ac.uk)

(Received 20 May 2023; accepted 11 March 2024)

We prove that any positive rational number is the sum of distinct unit fractions with denominators in $\{p-1 : p \text{ prime}\}$. The same conclusion holds for the set $\{p-h : p \text{ prime}\}$ for any $h \in \mathbb{Z}\backslash\{0\}$, provided a necessary congruence condition is satisfied. We also prove that this is true for any subset of the primes of relative positive density, provided a necessary congruence condition is satisfied.

Keywords: number theory; unit fractions; prime numbers

2020 *Mathematics Subject Classification:* 11D68; 11L20

The study of decompositions of rational numbers into sums of distinct unit fractions (often called 'Egyptian fractions') is one of the oldest topics in number theory (see [**2**] for further background and many related problems on Egyptian fractions). It is elementary to prove that such a decomposition is always possible (for example a simple greedy algorithm suffices). In this paper we explore a natural variant that imposes restrictions on the denominators in these decompositions.

QUESTION 1. *For which $A \subseteq \mathbb{N}$ is it true that every positive rational number can be written as $\sum_{n\in B} 1/n$ for some finite $B \subset A$?*

A trivial necessary condition is that the set contains multiples of every prime; for example, the set of all odd numbers does not have this property (it cannot represent $1/2$). The condition $\sum_{n\in A} 1/n = \infty$ is also clearly necessary. Both of these conditions together are not sufficient – indeed, it is easy to see that there is no solution to $1 = \sum_{p\in A} 1/p$ where A is any finite set of primes, even though $\sum_{p\leqslant N} 1/p \sim \log\log N$.

An early seminal paper on this topic is by Graham [**6**], who proved a general result that implies, for example, that such a decomposition always exists when A is the set of all primes *and* squares. Motivated by a conjecture of Sun [**12**, Conjecture 4.1], Eppstein [**4**] developed an alternative elementary method, which implies such a decomposition always exists when A is the set of 'practical numbers' (those n such that all $m \leqslant n$ can be written as the sum of distinct divisors of n).

A variant of question 1 can be asked even when there are trivial obstructions. For example, Graham [**6**] has shown that every rational number x can be written

© The Author(s), 2024. Published by Cambridge University Press on behalf of The Royal Society of Edinburgh. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

as the sum of distinct unit fractions with square denominators, subject to the obviously necessary condition that $x \in [0, \pi^2/6 - 1) \cup [1, \pi^2/6)$, and further showed that every rational number x with square-free denominator can be written as the sum of distinct unit fractions with square-free denominators.

A natural candidate of number theoretic interest, for which there exist no obvious obstructions to any rational decomposition, and for which the methods of [6] and [4] are not applicable, is the set of shifted primes $\{p - 1 : p \text{ prime}\}$. That such a restricted Egyptian fraction decomposition always exists was conjectured by Sun [12, Conjecture 4.1] (see also [13, Conjecture 8.17] and [11] for some numerical data). In this paper we use the method of [1] to prove this conjecture: any positive rational $x > 0$ has a solution (indeed, infinitely many) to

$$x = \frac{1}{p_1 - 1} + \cdots + \frac{1}{p_k - 1}$$

where $p_1 < \cdots < p_k$ are distinct primes. We also prove a similar result with denominators $p_i - h$ for any (fixed) $h \neq 0$, although for $|h| > 1$ there are some trivial congruence obstructions – for example, since no subset of $\{p + 2 : p \text{ prime}\}$ has lowest common multiple divisible by 8 the fraction $1/8$ cannot be represented as the sum of distinct unit fractions of the shape $1/p + 2$.

We deduce this existence result from the following more general result, showing that any shifted set of primes, all divisible by q, of 'positive upper relative logarithmic density' contains a decomposition of $1/q$. (Recall that $\sum_{p \leqslant N} 1/p \sim \log\log N$, and so it is natural to consider $\sum_{\substack{p \leqslant N \\ p \in A}} 1/p$ divided by $\log\log N$ as a measure of the size of A.)

THEOREM 0.1. *Let $h \in \mathbb{Z}\backslash\{0\}$ and $q \geqslant 1$ be such that $(|h|, q) = 1$. If A is a set of primes congruent to $h \pmod q$ such that*

$$\limsup_{N \to \infty} \frac{\sum_{p \in A \cap [1,N]} \frac{1}{p}}{\log\log N} > 0$$

then there exists a finite $S \subset A$ such that

$$\frac{1}{q} = \sum_{p \in S} \frac{1}{p - h}.$$

A simple application of partial summation produces the following version with (relative) upper logarithmic density replaced by (relative) lower density.

COROLLARY 0.2. *Let $h \in \mathbb{Z}\backslash\{0\}$ and $q \geqslant 1$ be such that $(|h|, q) = 1$. If A is a set of primes congruent to $h \pmod q$ with positive relative lower density, that is,*

$$\liminf_{N \to \infty} \frac{|A \cap [1, N]|}{N/\log N} > 0,$$

then there exists a finite $S \subset A$ such that

$$\frac{1}{q} = \sum_{p \in S} \frac{1}{p - h}.$$

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

We remark that (unlike the statement for unrestricted sets of integers, see [**1**, Theorem 2]) the stronger version of Corollary 0.2 with the $\liminf_{N\to\infty}$ replaced by $\limsup_{N\to\infty}$ is false - for example, if $A[N]$ is the set of primes in $[N/2, N]$ then $\sum_{p\in A[N]} 1/p \ll 1/\log N$, and hence if $A = \cup_k A[N_k]$ where $N_k = 2^{k^C}$ for some large absolute constant $C > 0$ then $\sum_{n\in A} 1/n < 1$, and hence certainly we cannot find a finite $S \subset A$ such that $\sum_{n\in S} 1/n = 1$, and yet

$$\limsup_{N\to\infty} \frac{|A \cap [1, N]|}{N/\log N} \geqslant 1/2.$$

We now show how Theorem 0.1 implies the headline result: any positive rational number (subject to the necessary congruence conditions) can be written as the sum of distinct unit fractions with shifted prime denominators.

Corollary 0.3. *Let $h \in \mathbb{Z}$ and $x = r/q \in \mathbb{Q}_{>0}$ be such that $(|h|, q) = 1$. There are distinct primes $p_1, \dots, p_k$ such that*

$$x = \frac{1}{p_1 - h} + \cdots + \frac{1}{p_k - h}.$$

Proof. By Dirichlet's theorem (see for example [**10**, Corollary 4.12]) if A is the set of primes congruent to $h \pmod q$, then

$$\sum_{n\in A\cap[1,N]} \frac{1}{n} \geqslant \left(\frac{1}{\phi(q)} + o(1)\right) \log\log N.$$

Trivially the same must hold for $A\backslash B$, for any finite set B. In particular by r repeated applications of Theorem 0.1 (first to A, then $A\backslash S_1$, and so on) we can find r disjoint finite sets $S_1, \dots, S_r \subset A$ such that

$$\frac{1}{q} = \sum_{p\in S_i} \frac{1}{p - h}$$

for $1 \leqslant i \leqslant r$. It follows that

$$x = \frac{r}{q} = \sum_{p\in \bigcup S_i} \frac{1}{p - h}$$

as required. □

We prove Theorem 0.1 with an application of the author's earlier work [**1**] (which in turn is a stronger form of an argument of Croot [**3**]). Loosely speaking, the main result of [**1**] shows that we can solve $1 = \sum 1/n_i$ with $n_i \in A$ whenever A satisfies

(i) $\sum_{n\in A} 1/n \to \infty$,

(ii) every $n \in A$ is 'friable' (or 'smooth'), in that if a prime power q divides n then $q \leqslant n^{1-\delta(n)}$ for some $0 < \delta(n) = o(1)$,

(iii) every $n \in A$ has 'small divisors', and

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

(iv) every $n \in A$ has $\approx \log\log n$ many distinct prime divisors.

To prove Theorem 0.1, therefore, it suffices to show that the set $\{p - h/q : p \in A\}$ has these properties. Fortunately, there has been a great deal of study of the arithmetic properties of shifted primes, and so using classical techniques from analytic number theory we are able to find a large subset of our original set A satisfying all four properties.

For experts in analytic number theory we add that in establishing the necessary number theoretic facts about shifted primes we have followed the simplest path, forgoing many of the more elaborate refinements possible. The main observation of this paper is that the inputs required to the method of [**1**] are mild enough to be provable for the shifted primes using (a crude form of) existing technology.

To minimize technicalities we have proved only a qualitative form of Theorem 0.1. In principle a (very weak) quantitative version could be proved with the same methods, along similar lines to [**1**, Theorem 3], but this would complicate the presentation significantly.

Finally, the methods and main results of [**1**] have now been formally verified using the Lean proof assistant, in joint work with Bhavik Mehta.[1] This formalization has not been extended to the present work, but since the proof of Theorem 0.1 uses the main result of [**1**] as its primary ingredient (combined with classical number theory) it can be viewed as 'partially formally verified'.

In § 1 we prove Theorem 0.1 assuming certain number theoretic lemmas. In § 2 we prove these lemmas.

1. Proof of Theorem 0.1

Our main tool is the following slight variant of [**1**, Proposition 1] (which is identical to the below except that the exponent of c is replaced by $1/\log\log N$).

PROPOSITION 1.1. *Let $c \in (0, 1/4)$ and N be sufficiently large (depending only on c). Suppose $A \subset [N^{1-c}, N]$ and $1 \leqslant y \leqslant z \leqslant (\log N)^{1/500}$ are such that*

(i) $\sum_{n\in A} 1/n \geqslant 2/y + (\log N)^{-1/200}$,

(ii) *every $n \in A$ is divisible by some d_1 and d_2 where $y \leqslant d_1$ and $4d_1 \leqslant d_2 \leqslant z$,*

(iii) *every prime power q dividing some $n \in A$ satisfies $q \leqslant N^{1-4c}$, and*

(iv) *every $n \in A$ satisfies*

$$\frac{99}{100}\log\log N \leqslant \omega(n) \leqslant 2\log\log N.$$

There is some $S \subseteq A$ such that $\sum_{n\in S} 1/n = 1/d$ for some $d \in [y, z]$.

Proof. The proof is identical to that of [**1**, Proposition 1], except that in the final part of the proof we choose $M = N^{1-c}$. Observe that the inputs to that proof, namely [**1**, Proposition 2, Proposition 3, and Lemma 7], are valid for any $M \in$

[1] The formalized proof can be found at https://github.com/b-mehta/unit-fractions.

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

$(N^{3/4}, N)$. It remains to check the 'friable' hypothesis, for which we require that if $n \in A$ and q is a prime power with $q \mid n$ then, for some small absolute constant $c' > 0$,

$$q \leqslant c' \min\left(\frac{M}{z}, \frac{M}{(\log N)^{1/100}}, \frac{M^3}{N^{2-4/\log\log N}(\log N)^{2+1/50}}\right).$$

For N sufficiently large (depending only on c and c') the right-hand side is $> N^{1-4c}$, and so hypothesis (iii) suffices. □

It is convenient to recast this in a slightly different form.

PROPOSITION 1.2. *Let $\delta, \epsilon > 0$ and suppose y is sufficiently large, depending on δ and ϵ, and $y \leqslant w \leqslant z$. If N is sufficiently large (depending on $\delta, \epsilon, y, w, z$) and $A \subset [2, N]$ is such that for all $n \in A$*

(i) *if a prime power q divides n then $q \leqslant n^{1-\epsilon}$,*

(ii) $|\omega(n) - \log\log n| \leqslant \log\log n/1000$,

(iii) *n is divisible by some $d_1 \in [y, w)$,*

(iv) *n is divisible by some $d_2 \in [4w, z)$, and*

(v) $\sum_{n\in A} 1/n \geqslant \delta \log\log N$,

then there exists $S \subseteq A$ such that $\sum_{n\in S} 1/n = 1/d$ for some $d \leqslant z$.

Proof. For $i \geqslant 0$ let $N_i = N^{(1-\epsilon/4)^i}$, and let $A_i = A \cap (N_{i+1}, N_i]$. Note that $N_i < 2$ for $i \geqslant C \log\log N$, where C is some sufficiently large constant depending only on ϵ. Since $\sum_{n\leqslant \log\log N} 1/n \ll \log\log\log N$ it follows by the pigeonhole principle that there must exist some i such that with $A' = A_i$ and $N' = N_i \gg \log\log N$ we have

$$\sum_{n\in A'} \frac{1}{n} \gg_{\delta,\epsilon} 1$$

and $A' \subset ((N')^{1-\epsilon/4}, N']$. It suffices to verify that the assumptions of Proposition 1.1 are satisfied by A', with $c = \epsilon/4$. We have already verified the first assumption (assuming y and N are sufficiently large; note that since $N' \gg \log\log N$ this ensures that N' is also sufficiently large). The second assumption of Proposition 1.1 is ensured by hypotheses (iii) and (iv).

For the third assumption, note that by hypothesis (i) if $n \in A'$ is divisible by a prime power q then

$$q \leqslant n^{1-\epsilon} \leqslant (N')^{1-\epsilon}$$

as required. Finally the fourth assumption follows from hypothesis (ii) and noting that for all $n \in [(N')^{1-\epsilon/4}, N']$ we have

$$\log\log n = \log\log N' + O_\epsilon(1),$$

and the $O_\epsilon(1)$ term is $\leqslant \log\log N'/500$, say, provided we take N sufficiently large. □

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

To prove Theorem 0.1 we want to apply Proposition 1.2 to $B = \{p - h/q : p \in A\}$. To verify the hypotheses we will require the following number-theoretic lemmas. We were unable to find these exact statements in the literature, so have included proofs in the following section, but the proofs are all elementary and cover well-trodden ground.

LEMMA 1.3. *For any $\epsilon > 0$ and $h \in \mathbb{Z}\backslash\{0\}$ the relative density of primes p such that $n = p - h$ is divisible by a prime power $q > n^{1-\epsilon}$ is $O_h(\epsilon)$.*

LEMMA 1.4. *For any $\delta > 0$ and $h \in \mathbb{Z}\backslash\{0\}$ the relative density of primes p such that $n = p - h$ has*

$$|\omega(n) - \log\log(n)| \geqslant \delta \log\log n$$

is 0.

LEMMA 1.5. *For any $h \in \mathbb{Z}\backslash\{0\}$, if $4 \leqslant y < z$ the relative density of primes p such that $n = p - h$ is not divisible by any primes $q \in [y, z]$ is $O_h(\log y/\log z)$.*

We will now show how these lemmas, combined with Proposition 1.2, imply Theorem 0.1.

Proof of Theorem 0.1. By assumption there is some $\delta > 0$ and infinitely many N such that

$$\sum_{p \in A \cap [1,N]} \frac{1}{p} \geqslant 4\delta \log\log N.$$

Let $B = \{p - h/q : p \in A\} \subset \mathbb{N}$, so that there must exist infinitely many N such that

$$\sum_{n \in B \cap [1,N]} \frac{1}{n} \geqslant 3\delta \log\log N.$$

Let $\epsilon = c\delta$ where $c > 0$ is some small absolute constant to be determined later. Let y be sufficiently large in terms of δ (so that Proposition 1.2 can apply) and $w \leqslant z$ be determined shortly, and let $B' \subseteq B$ be the set of those $n \in B$ such that

(i) if a prime power r divides n then $r \leqslant n^{1-\epsilon}$,

(ii) $|\omega(n) - \log\log n| \leqslant \log\log n/1000$,

(iii) n is divisible by some prime $p_1 \in [y, w)$, and

(iv) n is divisible by some prime $p_2 \in [4w, z)$.

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

If B_0 is the set of $m = p - h$ which are divisible by some prime power $r > m^{1-2\epsilon}$ then by Lemma 1.3 we have, for all large X,

$$|B_0 \cap [1, X]| \ll \epsilon \frac{X}{\log X},$$

and hence, since for all large primes p we have $(p-h)^{1-2\epsilon} \leqslant (p-h/q)^{1-\epsilon}$, the set B_1 of those $n \in B$ which fail the first condition satisfies, for all large X,

$$|B_1 \cap [1, X]| \ll \epsilon \frac{X}{\log X},$$

whence by partial summation, for all large N,

$$\sum_{n \in B_1 \cap [1,N]} \frac{1}{n} \ll \epsilon \log\log N.$$

By a similar argument (recalling that q is some fixed constant, and so $\omega(p - h/q) = \omega(p-h) + O(1)$ and $\log\log(p - h/q) = \log\log(p-h) + O(1)$), Lemma 1.4 implies that the sum of reciprocals from those $n \in B \cap [1, N]$ which fail the second condition is $o(\log\log N)$. Similarly, by Lemma 1.5 we can choose w and z (depending only δ) such that for all large N the sum of reciprocals from those $n \in B \cap [1, N]$ which fail either condition (iii) or (iv) is $\leqslant \delta \log\log N$. Therefore, there exist infinitely many N such that (provided ϵ is a small enough multiple of δ)

$$\sum_{n \in B' \cap [1,N]} \frac{1}{n} \geqslant 2\delta \log\log N.$$

Fix such an N and let $B'' = B' \cap [1, N]$. All of the conditions from Proposition 1.2 are now satisfied for B'', and hence there exists some $S_1 \subseteq B''$ and $d_1 \leqslant z$ such that $\sum_{n \in S_1} 1/n = 1/d_1$.

We now apply Proposition 1.2 again to $B'' \backslash S_1$, and continue this process $k = \lceil z \rceil^2$ many times, producing some disjoint $S_1, \ldots, S_k$ and associated $d_1, \ldots, d_k \leqslant z$ where $\sum_{n \in S_i} 1/n = 1/d_i$ for $1 \leqslant i \leqslant k$. Notice that the conditions of Proposition 1.2 remain satisfied for each $B'' \backslash \cup_{i \leqslant j} S_i$ for $j \leqslant k$, since

$$\sum_{n \in \cup_{i \leqslant j} S_i} \frac{1}{n} \leqslant k \ll z^2 < \delta \log\log N,$$

assuming N is sufficiently large, since z depends on δ only.

By the pigeonhole principle there must exist some $d \leqslant z$ and $i_1, \ldots, i_d$ such that $d_{i_j} = d$ for $1 \leqslant j \leqslant d$, and hence $S = \cup_{1 \leqslant j \leqslant d} S_{i_j}$ satisfies

$$\sum_{n \in S} \frac{1}{n} = d \cdot \frac{1}{d} = 1$$

as required. □

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

2. Number theoretic ingredients

It remains to prove Lemmas 1.3, 1.4, and 1.5, which we will do in turn.

2.1. Friability of shifted primes

There has been a great deal of work on shifted primes with only small prime divisors. Often the focus is on an existence result, finding the smallest possible $\delta > 0$ such that there exist infinitely many shifted primes $p-1$ with no prime divisors $> p^{\delta}$. We refer to [9] for recent progress on this and references to earlier work. Our focus is a little different: we are content with a very high friability threshold, but we need to show that almost all shifted primes are this friable. For the regime of friability that we are interested even the original elementary methods of Erdős [5] suffice

Proof of Lemma 1.3. This is only a slight generalization of [5, Lemma 4]. It suffices to show that, for all $\epsilon > 0$ and large N, the number of $p \leqslant N$ such that $p-h$ is divisible by some prime power q with $q > N^{1-\epsilon}$ is

$$\ll_h \epsilon \frac{N}{\log N}.$$

We first note that trivially for any q the number of $p \leqslant N$ such that $p-h$ is divisible by q is certainly $O_h(N/q)$, and hence the count of those $p-h$ divisible by some non-prime prime power $q > N^{1-\epsilon}$ is

$$\ll_h N \sum_{k \geqslant 2} \sum_{N^{1-\epsilon} \leqslant m^k \leqslant N} \frac{1}{m^k} \ll N^{\epsilon} \ll \epsilon \frac{N}{\log N}$$

for all large N. It remains to bound the count of those $p \leqslant N$ such that $p-h$ is divisible by some prime $q > N^{1-\epsilon}$. Such $p-h$ we can write uniquely (assuming N is large enough depending on h) as $p-h = qa$ for some $a \leqslant 2N^{\epsilon}$ and $q > N^{1-\epsilon}$ prime. A simple application of Selberg's sieve (for example [8, Theorem 3.12]) yields that, for any fixed $a \geqslant 1$ and $h \neq 0$ the number of primes $q \leqslant x$ such that $aq+h$ is also prime is

$$\ll_h \frac{a}{\phi(a)} \frac{x}{(\log x)^2}.$$

Since $q \leqslant N/a + O_h(1)$, the number of $p \leqslant N$ such that $p-h = qa$ is

$$\ll \frac{1}{\phi(a)} \frac{N}{(\log N)^2}.$$

Summing over all $a \leqslant 2N^{\epsilon}$ the total count is

$$\ll_h \frac{N}{(\log N)^2} \sum_{a \leqslant 2N^{\epsilon}} \frac{1}{\phi(a)} \ll \epsilon \frac{N}{\log N}$$

as required, using the fact that $\sum_{a \leqslant M} 1/\phi(a) \ll \log M$. □

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

2.2. Number of prime divisors of shifted primes

We need to know that $\omega(n) \sim \log\log n$ for almost all $n \in \{p-h : p \text{ prime}\}$. This is in fact the typical behaviour of $\omega(n)$ for a generic integer n, and we expect the same behaviour when restricting n to the random-like sequence of shifted primes. Indeed, just like $\omega(n)$ itself, $\omega(p-h)$ satisfies an Erdős–Kac theorem, that is, $\omega(p-h)$ behaves like a normal distribution with mean $\log\log(p-h)$ and standard deviation $\sqrt{\log\log(p-h)}$. This was established by Halberstam [7], although a simple variance bound suffices for our application here.

Proof of Lemma 1.4. It suffices to show that, for all $\delta > 0$ and large N, if A is the set of $p \leqslant N$ such that $|\omega(p-h) - \log\log(p-h)| > \delta \log\log(p-h)$, then

$$|A| \ll \frac{N}{(\log N)(\log\log N)}.$$

Let $A_1 = A \cap [1, N^{1/2}]$ and $A_2 = A \backslash A_1$. We can trivially bound $|A_1| \ll N^{1/2}$, and for $p \in A_2$ we have $\log\log(p-h) = \log\log N + O(1)$, whence for large enough N if $p \in A_2$ we have

$$|\omega(p-h) - \log\log N| > \frac{\delta}{2}\log\log N.$$

By [7, Theorem 3], however, we have

$$\sum_{p \leqslant N} |\omega(p-h) - \log\log N|^2 \ll \pi(N)\log\log N,$$

and hence

$$|A_2|\,(\log\log N)^2 \ll_\delta \pi(N)\log\log N,$$

and the result now follows from Chebyshev's estimate $\pi(N) \ll N/\log N$. □

2.3. Shifted primes with small divisors

For Lemma 1.5 we need to show that there are few shifted primes remaining after we remove all multiples of primes $p \in [y, z]$, which is a classic upper bound sieve problem. Since the information we require is very weak even the simplest sieve suffices: the following is proved as [8, Theorem 1.1].

Lemma 2.1 Sieve of Eratosthenes-Legendre. *Let A be a finite set of integers and $\mathcal{P}$ a finite set of primes. Let $z \geqslant 2$ and $P(z) = \prod_{\substack{p \in \mathcal{P} \\ p < z}} p$. For any $d \geqslant 1$ let $A_d = \{n \in A : d \mid n\}$. Suppose that $f(d)$ is a multiplicative function and $X > 1$ is such that for all $d \mid P(z)$ we have*

$$|A_d| = f(d)X + R_d.$$

Then

$$\#\{n \in A : (n, P(z)) = 1\} \ll X \prod_{\substack{p \in P \\ p < z}} (1 - f(p)) + \sum_{d \mid P(z)} |R_d|\,.$$

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

For the required sieve input we will use the following classic result on the distribution of primes within arithmetic progressions (which is proved, for example, as [**10**, Corollary 11.21]). Recall that $\pi(N; d, h)$ is the number of primes $p \leqslant N$ such that $p \equiv h \pmod d$.

THEOREM 2.2 Siegel–Walfisz. *There is a constant $c > 0$ such that, for all sufficiently large N, for all $h \in \mathbb{Z}$ and $1 \leqslant d \leqslant \log N$ with $(|h|, d) = 1$ we have*

$$\pi(N; d, h) = \frac{\mathrm{li}(N)}{\phi(d)} + O(N \exp(-c\sqrt{\log N})).$$

Proof of Lemma 1.5. Fix $4 \leqslant y \leqslant z$ and let $P = \prod_{\substack{y \leqslant q \leqslant z \\ q \nmid h}} q$ (where q is restricted to primes). It suffices to show that, for all large N,

$$\#\{p - h \leqslant N : (p - h, P) = 1\} \ll_h \frac{\log y}{\log z}\mathrm{li}(N).$$

We will apply Lemma 2.1 with $A = \{p - h : p \leqslant N\}$,

$$\mathcal{P} = \{p \in [y, z] : p \nmid h\},$$

$f(d) = 1/\phi(d)$, and $X = \mathrm{li}(N)$, noting that by Theorem 2.2 whenever $(d, h) = 1$ and $d \leqslant \log N$

$$|A_d| = \pi(N; d, h) = \frac{\mathrm{li}(N)}{\phi(q)} + O(x \exp(-c\sqrt{\log x})).$$

It follows that

$$\#\{p - h \leqslant N : (p - h, P) = 1\} \ll \mathrm{li}(N) \prod_{\substack{y \leqslant q \leqslant z \\ q \nmid h}} \left(1 - \frac{1}{q - 1}\right) + 2^z N \exp(-c\sqrt{\log N}).$$

The conclusion now follows provided we choose N large enough so that $z \ll \sqrt{\log \log N}$, say, and using Mertens' estimate that

$$\prod_{p \leqslant w} (1 - 1/p) \asymp \frac{1}{\log w}.$$

□

Acknowledgements
The author is funded by a Royal Society University Research Fellowship. We would like to thank Greg Martin for a helpful conversation about friable values of shifted primes and remarks on an earlier version of this paper.

References

1 T. F. Bloom, *On a density conjecture about unit fractions*, submitted.
2 T. F. Bloom and C. Elsholtz, *Egyptian fractions* NAW 5/23 nr. 4 December 2022.

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press

3 E. S. Croot. On a coloring conjecture about unit fractions. *Ann. Math. (2)* **157** (2003), 545–556.

4 D. Eppstein. Egyptian fractions with denominators from sequences closed under doubling. *J. Integer Seq.* **24** (2021), 21.8.8..

5 P. Erdős. On the normal number of prime factors of $p-1$ and some related problems concerning Euler's ϕ-function. *Quart. J. Math. Oxford Ser.* **6** (1935), 205–213.

6 R. L. Graham. On finite sums of unit fractions. *Proc. London Math. Soc. (3)* **14** (1964), 193–207.

7 H. Halberstam. On the distribution of additive number theoretic functions (III). *J. London Math. Soc.* **31** (1956), 15–27.

8 H. Halberstam and H. E. Richert. *Sieve Methods* (New York: Academic Press, 1974).

9 J. D. Lichtman, *Primes in arithmetic progressions to large moduli and shifted primes without large prime factors*, eprint arXiv:2211.09641.

10 H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory I. Classical Theory* (Cambridge University Press, Cambridge, 2007).

11 Z.-W. Sun, *A conjecture on unit fractions involving primes* http://maths.nju.edu.cn/zwsun/UnitFraction.pdf.

12 Z.-W. Sun, *Conjectures on Representations Involving Primes* In: M. Nathanson (ed) Combinatorial and Additive Number Theory II. Springer Proceedings in Mathematics & Statistics, Vol 220 (Springer, Cham).

13 Z.-W. Sun. *New Conjectures in Number Theory and Combinatorics* (Harbin: Harbin Institute of Technology Press, 2021).

https://doi.org/10.1017/prm.2024.42 Published online by Cambridge University Press