

The number of solutions of the Erdős-Straus Equation and sums of k unit fractions

Christian Elsholtz and Stefan Planitzer

Graz University of Technology, Institute of Analysis and Number Theory, Kopernikusgasse 24/II, Graz 8010, Austria
(elsholtz@math.tugraz.at; planitzer@math.tugraz.at)

(MS received 21 February 2018; accepted 20 June 2018)

We prove new upper bounds for the number of representations of an arbitrary rational number as a sum of three unit fractions. In particular, for fixed m there are at most $\mathcal{O}_\epsilon(n^{3/5+\epsilon})$ solutions of $m/n = 1/a_1 + 1/a_2 + 1/a_3$. This improves upon a result of Browning and Elsholtz (2011) and extends a result of Elsholtz and Tao (2013) who proved this when $m = 4$ and n is a prime. Moreover, there exists an algorithm finding all solutions in expected running time $\mathcal{O}_\epsilon(n^\epsilon(n^3/m^2)^{1/5})$, for any $\epsilon > 0$. We also improve a bound on the maximum number of representations of a rational number as a sum of k unit fractions. Furthermore, we also improve lower bounds. In particular, we prove that for given $m \in \mathbb{N}$ in every reduced residue class $e \pmod f$ there exist infinitely many primes p such that the number of solutions of the equation $m/p = 1/a_1 + 1/a_2 + 1/a_3$ is $\gg_{f,m} \exp((5 \log 2 / (12 \operatorname{lcm}(m, f))) + o_{f,m}(1)) \log p / \log \log p$. Previously, the best known lower bound of this type was of order $(\log p)^{0.549}$.

Keywords: unit fractions, Erdős-Straus equation, Diophantine equations

2010 *Mathematics subject classification:* Primary: 11D68,
Secondary: 11D72

1. Introduction

We consider the problem of finding upper bounds for the number of solutions in positive integers a_1, a_2 and a_3 of equations of the form

$$\frac{m}{n} = \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} \tag{1.1}$$

where $m, n \in \mathbb{N}$ are fixed. In the case when $m = 4$ we call equation (1.1) Erdős-Straus equation. The Erdős-Straus conjecture states that this equation has at least one solution for any $n > 1$ (see [12, 16: D11] for classical results concerning the Erdős-Straus equation and several related problems, as well as [15] for a survey of the work of Erdős on Egyptian fractions). Also the more general equation

$$\frac{m}{n} = \sum_{i=1}^k \frac{1}{a_i} \tag{1.2}$$

for $m, n \in \mathbb{N}$ fixed and $a_1, \dots, a_k \in \mathbb{N}$ received some attention. Browning and Elsholtz [5] found upper bounds for the number of solutions of (1.2). For the special case $m = n = 1$ they were able to improve a result of Sándor [27] and proved that there are at most $c_0^{(5/24+\epsilon)2^k}$ representations of 1 as a sum of k unit fractions, for any $\epsilon > 0$ and sufficiently large k . Here $c_0 = \lim_{n \rightarrow \infty} u_n^{2^{-n}} = 1.264\dots$ where $u_1 = 1$ and $u_{n+1} = u_n(u_n + 1)$. On the other hand, Konyagin [21] proved a lower bound of order $\exp(\exp(((\log 2)(\log 3)/3 + o(1))k/\log k))$ for the number of these representations with distinct denominators. While the Erdős-Straus conjecture is about representing certain rational numbers as a sum of just three unit fractions, Martin [24] worked on representations of positive rationals as sums of many unit fractions. In particular, he proved that every positive rational number r has a representation of the form $r = \sum_{s \in S} 1/s$, where the set S contains a positive proportion of the integers less than any sufficiently large real number x .

Chen et.al. [7] dealt with representations of 1 as a sum of k distinct unit fractions where the denominators satisfy certain restrictions (like all of them being odd). Several results on representations of rational numbers as a sum of unit fractions with restrictions on the denominators can be found in the work of Graham [13–15]. Elsholtz [11] proved a lower bound of similar order as the one of Konyagin for the number of representations of 1 as a sum of k distinct unit fractions with odd denominators.

For sums of k unit fractions we adopt the notation of [5] and define $f_k(m, n)$ to be the number of solutions $(a_1, a_2, \dots, a_k) \in \mathbb{N}^k$ of equation (1.2) with $a_1 \leq a_2 \leq \dots \leq a_k$, that is,

$$f_k(m, n) = \left| \left\{ (a_1, a_2, \dots, a_k) \in \mathbb{N}^k : \frac{m}{n} = \frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_k}, a_1 \leq a_2 \leq \dots \leq a_k \right\} \right|.$$

Concerning equation (1.1) with $m = 4$ the results of Elsholtz and Tao [12] show that the number of solutions $f_3(4, n)$ is related to some divisor questions and is on average a power of $\log n$ (at least when n is prime). It even seems possible that for fixed $m \in \mathbb{N}$ and any $\epsilon > 0$ the number of representations of $\frac{m}{n}$ as a sum of k unit fractions is bounded by $\mathcal{O}_{k,\epsilon}(n^\epsilon)$. More details on this are informally and heuristically discussed in § 3. For general m and n the best known upper bound on the number of solutions of (1.1) is due to Browning and Elsholtz [5, theorem 2] who proved an upper bound of order $\mathcal{O}_\epsilon(n^\epsilon(n/m)^{2/3})$. In the case of the Erdős-Straus equation with $n = p$ prime Elsholtz and Tao [12, proposition 1.7] have improved this bound to $\mathcal{O}_\epsilon(p^{3/5+\epsilon})$. It is known that this type of question is easier to study when the denominator is prime.

Our main result will be the following theorem which provides an upper bound on the number of solutions of equation (1.1).

THEOREM 1.1. *For any $m, n \in \mathbb{N}$ and any $\epsilon > 0$ there are at most $\mathcal{O}_\epsilon(n^\epsilon(n^3/m^2)^{1/5})$ solutions of the equation*

$$\frac{m}{n} = \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}$$

in positive integers a_1, a_2 and a_3 .

Note that this improves upon the bound of Browning and Elsholtz in the range $m \ll n^{1/4}$. As a corollary, we get that the Elsholtz-Tao bound for the number of solutions of the Erdős-Straus equation is true for arbitrary denominators $n \in \mathbb{N}$.

COROLLARY 1.2. *The Erdős-Straus equation*

$$\frac{4}{n} = \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}$$

has at most $\mathcal{O}_\epsilon(n^{3/5+\epsilon})$ solutions in positive integers a_1, a_2 and a_3 .

We also prove the following algorithmic version of theorem 1.1 with a matching upper bound for the expected running time¹.

COROLLARY 1.3. *There exists an algorithm with an expected running time of order $\mathcal{O}_\epsilon(n^\epsilon(n^3/m^2)^{1/5})$, for any $\epsilon > 0$, which lists all representations of the rational number m/n as a sum of three unit fractions. Furthermore, all representations of m/n as a sum of $k > 3$ unit fractions may be found in expected time $\mathcal{O}_{\epsilon,k}(n^{2^{k-3}(8/5+\epsilon)-1})$, for any $\epsilon > 0$.*

For sums of k unit fractions, we will prove the following result.

THEOREM 1.4. *We have*

$$f_4(m, n) \ll_\epsilon n^\epsilon \left(\frac{n^{4/3}}{m^{2/3}} + \frac{n^{28/17}}{m^{8/5}} \right)$$

and for any $k \geq 5$

$$f_k(m, n) \ll_\epsilon (kn)^\epsilon \left(\frac{k^{4/3}n^2}{m} \right)^{28/17 \cdot 2^{k-5}}.$$

Keeping in mind that $28/17 = 1.64705\dots$, theorem 1.4 may be compared with the following bounds from [5, theorem 3]:

$$f_4(m, n) \ll_\epsilon n^\epsilon \left(\frac{n^{4/3}}{m^{2/3}} + \left(\frac{n}{m} \right)^{5/3} \right),$$

$$f_k(m, n) \ll_\epsilon (kn)^\epsilon \left(\frac{k^{4/3}n^2}{m} \right)^{5/3 \cdot 2^{k-5}}, \quad \text{for } k \geq 5.$$

A well studied special case of theorem 1.4 concerns representations of 1 as a sum of k unit fractions. Browning and Elsholtz [5] mention several related problems which are studied in the literature and can be improved using better upper bounds on $f_k(m, n)$. We summarize these results in the following corollary.

¹For a definition of expected running time see the proof of this corollary at the end of §5.

COROLLARY 1.5.

(1) For any $\epsilon > 0$, we have that

$$f_k(1, 1) \ll_{\epsilon} k^{7/51 \cdot 2^{k-1} + \epsilon}.$$

(2) Let u_n be the sequence recursively defined by $u_0 = 1$ and $u_{n+1} = u_n(u_n + 1)$ and set $c_0 = \lim_{n \rightarrow \infty} u_n^{2^{-n}}$. Then for $\epsilon > 0$ and $k \geq k(\epsilon)$ we have

$$f_k(1, 1) < c_0^{(7/17 + \epsilon)2^{k-1}}.$$

(3) For $\epsilon > 0$ and $k \geq k(\epsilon)$ the number $S(k)$ of positive integer solutions of the equation

$$1 = \sum_{i=1}^k \frac{1}{a_i} + \frac{1}{\prod_{i=1}^k a_i}$$

is bounded from above by $c_0^{(7/17 + \epsilon)2^k}$.

Proof. The first assertion is an immediate consequence of theorem 1.4. For the proof of the second statement we refer the reader to the proof of theorem 4 in [5]. The only change necessary is plugging in the bound from theorem 1.4 instead of [5, theorem 3] for the last 5 lines of the proof which amounts to just exchanging one exponent. The last statement follows from the first one and the observation that $S(k) \leq f_{k+1}(1, 1)$. □

We note that the number of solutions of the equation $1 = \sum_{i=1}^k 1/a_i + 1/\prod_{i=1}^k a_i$ has applications to problems considered in [4].

Finally, we deal with lower bounds. In [12, theorem 1.8] it is shown that we have

$$f_3(4, n) \geq \exp\left(\left(\log 3 + o(1)\right) \frac{\log n}{\log \log n}\right)$$

for infinitely many $n \in \mathbb{N}$ and that

$$f_3(4, n) \geq \exp\left(\left(\frac{\log 3}{2} + o(1)\right) \log \log n\right)$$

for all integers n in a subset of the positive integers with density 1. The following theorem gives an improvement of these bounds which also give a limitation on improving the upper bounds for the number of solution of the Erdős-Straus equation and in the general case. For comparison we note that $\log 3 = 1.09861 \dots$, $(\log 3)/2 = 0.54930 \dots$ and $\log 6 = 1.79175 \dots$

THEOREM 1.6. For given $m \in \mathbb{N}$ there are infinitely many $n \in \mathbb{N}$ such that

$$f_3(m, n) \geq \exp\left(\left(\log 6 + o_m(1)\right) \frac{\log n}{\log \log n}\right).$$

Furthermore, for given $m \in \mathbb{N}$, there exists a subset \mathcal{M}_1 of the integers with density one, such that for any $n \in \mathcal{M}_1$

$$f_3(m, n) \geq \exp((\log 3 + o_m(1)) \log \log n) \cdot \log \log n \\ \gg (\log n)^{\log 3 + o_m(1)}.$$

For the special case $m = 4$ and for integers n in a set $\mathcal{M}_2 \subset \mathbb{N}$ with density one, the last bound may be improved to

$$f_3(4, n) \geq \exp((\log 6 + o(1)) \log \log n).$$

REMARK 1.7. Previous proofs of lower bounds of similar type as the ones in theorem 1.6 constructed solutions from factorizations of n . We get our improvement from additionally taking into account factorizations of a lot of shifts of n . Hence our proof also shows that there are many values a_1 admitting many pairs (a_2, a_3) . Here, depending on which of the three lower bounds in theorem 1.6 we consider, ‘many’ may either mean $\exp((C + o_m(1)) \log n / \log \log n)$ or $\exp((\tilde{C} + o_m(1)) \log \log n)$, for suitable positive constants C and \tilde{C} .

We may ask if a lower bound on $f_3(m, n)$ of the first type in theorem 1.6 does not only hold for infinitely many positive integers n but also for infinitely many prime denominators p . In [12] there was no lower bound of this type, but it was proved that $f_3(4, p) \gg (\log p)^{0.549}$ for almost all primes. We note that this result implies, using Dirichlet’s theorem on primes, the following corollary.

COROLLARY 1.8. *For every reduced residue class $e \pmod f$, that is, $\gcd(e, f) = 1$, there are infinitely many primes p such that $f_3(4, p) \gg (\log p)^{0.549}$, and $p \equiv e \pmod f$.*

Here we improve this corollary considerably.

THEOREM 1.9. *For every $m \in \mathbb{N}$ and every reduced residue class $e \pmod f$ there are infinitely many primes $p \equiv e \pmod f$ such that*

$$f_3(m, p) \gg_{f,m} \exp\left(\left(\frac{5 \log 2}{12 \operatorname{lcm}(m, f)} + o_{f,m}(1)\right) \frac{\log p}{\log \log p}\right).$$

Here $o_{f,m}(1)$ denotes a quantity depending on f and m which goes to zero as p tends to infinity.

Using results of Harman [19, 20] one might be able to improve the factor $5/12$ in the exponent to 0.4736 .

2. Notation

As usual \mathbb{N} denotes the set of positive integers and \mathbb{P} the set of primes in \mathbb{N} . We denote the greatest common divisor and the least common multiple of n elements $a_i \in \mathbb{N}$ by $\gcd(a_1, a_2, \dots, a_n)$ and $\operatorname{lcm}(a_1, a_2, \dots, a_n)$ or (a_1, a_2, \dots, a_n) and $[a_1, a_2, \dots, a_n]$ for short. For integers $d, n \in \mathbb{N}$ we write $d|n$ if d divides n . We

use the symbols \mathcal{O} , o , \ll and \gg within the contexts of the well known Landau and Vinogradov notations where dependence of the implied constant on certain variables is indicated by a subscript. For any prime $p \in \mathbb{P}$ we define the function $\nu_p : \mathbb{N} \rightarrow \mathbb{N} \cup \{0\}$ to be the p -adic valuation, that is, $\nu_p(n) = a$ if and only if p^a is the highest power of p dividing n . By $\tau(n)$ and $\omega(n)$, as usual, we denote the number of divisors and the number of distinct prime divisors of n . By $\tau(n, m)$, we denote the number of divisors of n coprime to m and $\tau(n, k, m)$, $\omega(n, k, m)$ denote the number of divisors (resp. distinct prime divisors) of n in the residue class $k \pmod m$, where $(k, m) = 1$. Finally, for two coprime integers a and b we denote by $\text{ord}_a(b)$ the least positive integer l , such that $b^l \equiv 1 \pmod a$.

3. Heuristics on $f_k(m, n)$

We now informally discuss why $f_3(m, n) = \mathcal{O}_\epsilon(n^\epsilon)$ can be expected. In fact, as far as we are aware, this was first observed by Roger Heath-Brown (private communication with the first author in 1994). Let us first recall (see e.g. [28, p. 201: theorem 3]) that a fraction m/n with $\text{gcd}(m, n) = 1$ is a sum of two unit fractions $1/a_1 + 1/a_2$ if and only if there exist two distinct, positive and coprime divisors d_1 and d_2 of n such that $d_1 + d_2 \equiv 0 \pmod m$. We may deduce an upper bound of $\mathcal{O}_\epsilon(n^\epsilon)$ for the number of representations of m/n as a sum of two unit fractions. Indeed from

$$\frac{m}{n} = \frac{1}{a_1} + \frac{1}{a_2}, \tag{3.1}$$

by setting $d = (a_1, a_2)$ and $a'_i = a_i/d$ for $i \in \{1, 2\}$, we see that

$$ma'_1a'_2d = n(a'_1 + a'_2).$$

This implies that a'_1, a'_2 are divisors of n , d divides $n(a'_1 + a'_2) < 2n^2$ and any solution (a_1, a_2) of (3.1) uniquely corresponds to a triple (a'_1, a'_2, d) . The number $\sum_{a'_1, a'_2 | n} \tau(n(a'_1 + a'_2))$ of such triples is bounded by $\mathcal{O}_\epsilon(n^\epsilon)$ (see lemma A below).

Studying $m/n = 1/a_1 + 1/a_2 + 1/a_3$ with $a_1 \leq a_2 \leq a_3$ one observes that

$$\frac{1}{a_1} < \frac{m}{n} \leq \frac{3}{a_1}$$

from which $n/m < a_1 \leq 3n/m$ follows. In view of

$$\frac{m}{n} - \frac{1}{a_1} = \frac{ma_1 - n}{na_1} = \frac{1}{a_2} + \frac{1}{a_3} \tag{3.2}$$

there are at most $\mathcal{O}(n/m)$ choices for a_1 , and for given a_1 there are at most $d(na_1) = \mathcal{O}_\epsilon(n^\epsilon)$ divisors of na_1 . This shows that $f_3(m, n) = \mathcal{O}_\epsilon(n^{1+\epsilon}/m)$ is a trivial upper bound. The real question is for how many values of a_1 there can be at least one solution. For increasing a_1 , even if na_1 contains many divisors, the congruence $d_1 + d_2 \equiv 0 \pmod{ma_1 - n}$ should become, on average, more difficult to satisfy if $ma_1 - n \gg n^\epsilon$. Therefore we expect that the number of a_1 contributing at least one solution is $\mathcal{O}_\epsilon(n^\epsilon)$, so that $f_3(m, n) = \mathcal{O}_\epsilon(n^{2\epsilon})$. Moreover equation (3.2) implies that for any given a_1 , the number of solutions is about $\tilde{d}(m, n, a_1)$. Here $\tilde{d}(m, n, a_1)$

counts the number of pairs of coprime divisors d_1, d_2 of na_1 , with $d_1 + d_2 \equiv 0 \pmod{ma_1 - n}$. Therefore $f_3(m, n)$ should be approximately $\sum_{a_1} \tilde{d}(m, n, a_1)$.

Similarly, a completely trivial upper bound on $f_4(m, n)$ is as follows. With $a_1 \leq a_2 \leq a_3 \leq a_4$ it follows that $n/m < a_1 \leq 4n/m$ and hence

$$\frac{ma_1 - n}{na_1} = \frac{m}{n} - \frac{1}{a_1} = \frac{1}{a_2} + \frac{1}{a_3} + \frac{1}{a_4} \leq \frac{3}{a_2}.$$

From those bounds, we easily deduce that $a_2 \leq 12n^2/m$. With

$$\frac{m}{n} - \frac{1}{a_1} - \frac{1}{a_2} = \frac{ma_1a_2 - na_2 - na_1}{na_1a_2} = \frac{1}{a_3} + \frac{1}{a_4},$$

with similar arguments as above, we deduce that $f_4(m, n) = \mathcal{O}_\epsilon(n^{3+\epsilon}/m^2)$. For fixed m the fact that our bound on $f_4(m, n)$ in theorem 1.4 above is better than $\mathcal{O}(n^2)$ shows that for most pairs (a_1, a_2) and moreover, for most choices of $a_2 \in [n/m, 12n^2/m]$ there is no solution of $m/n = 1/a_1 + 1/a_2 + 1/a_3 + 1/a_4$. Here again, as soon as $ma_1a_2 - na_2 - na_1 \gg n^\epsilon$ one should not expect to have two divisors d_1, d_2 of na_1a_2 such that $d_1 + d_2 \equiv 0 \pmod{ma_1a_2 - na_2 - na_1}$. From this reasoning, also $f_k(m, n) = \mathcal{O}_{\epsilon,k}(n^\epsilon)$, for $k \geq 4$ seems to us a reasonable expectation.

The papers [5, 12] studied parametric solutions of the Diophantine equation (1.1). The reason why the result in [12] is superior in the case of n being a prime is that here a full parametric solution (e.g. [26]) is much easier to work with. However, in this paper, we develop parametric solutions of (1.1) and (1.2) from scratch. Some simplified version of this has been used in [10, 12: §11], but there the focus was to generate solutions with many parameters. Here we need to do kind of the opposite, namely to show that every solution comes from a number of parametric families.

The method we introduce should theoretically work for any Diophantine equation as it expresses a k -tuple of integers in a standard form. In practice, it might work favourably if there is some inhomogeneous part as in

$$n = a_1a_2a_3 - a_1 - a_2.$$

For prime values of n in equation (1.1) there are several discussions of parametric solutions in the literature, for example, by Rosati [26] and Aigner [1], see also Mordell’s book [25, Chapter 30]. For composite values n there is no satisfactory treatment in the literature, and § 5 below may be the most detailed study to date.

4. Patterns and relative greatest common divisors

Consider a solution $(a_1, a_2, \dots, a_k) \in \mathbb{N}^k$ with $a_1 \leq a_2 \leq \dots \leq a_k$ of equation (1.2) and set $n_i = (a_i, n)$, $a_i = n_i t_i$ for $i \in \{1, 2, \dots, k\}$. We can thus rewrite equation (1.2) as

$$\frac{m}{n} = \sum_{i=1}^k \frac{1}{n_i t_i}. \tag{4.1}$$

Later, when working on upper bounds for the number of solutions of equation (4.1) for $k \in \{3, 4\}$, we will fix a choice of $(n_1, n_2, \dots, n_k) \in \mathbb{N}^k$. For given $m, n \in \mathbb{N}$ we

call such a choice the *pattern* of a solution of this equation. Note that for solutions corresponding to a given pattern (n_1, n_2, \dots, n_k) we have that $(n/n_i, t_i) = 1$ for all $i \in \{1, 2, \dots, k\}$. As $n_i|n$ the number of distinct patterns is $\mathcal{O}_k(n^\epsilon)$ only.

Also, when dealing with equations of type (4.1) for $k \in \{3, 4\}$ we will make heavy use of the concept of relative greatest common divisors as described by Elsholtz in [9] (for some ad hoc definition see also [10]). Relative greatest common divisors are a useful tool when studying divisibility relations among the t_i in (4.1).

Let $I = \{1, 2, \dots, k\}$ be the index set. Then we define the relative greatest common divisors of the positive integers t_1, t_2, \dots, t_k recursively as follows:

$$x_I = \gcd(t_1, t_2, \dots, t_k)$$

and for any $\{i_1, i_2, \dots, i_{|J|}\} = J \subseteq I, J \neq \emptyset$ we set

$$x_J = \frac{\gcd(t_{i_1}, t_{i_2}, \dots, t_{i_{|J|}})}{\prod_{\substack{J' \subseteq I \\ J \subsetneq J'}} x_{J'}}.$$

For $k \in \{3, 4\}$, we will later identify the elements x_J with $J \subseteq I$ with the elements x_i, x_{ij} and x_{ijk} where $\{i, j, k\} = \{1, 2, 3\}$ in the case when $k = 3$ and with the elements x_i, x_{ij}, x_{ijk} and x_{ijkl} with $\{i, j, k, l\} = \{1, 2, 3, 4\}$ when $k = 4$. With the relative greatest common divisors defined as above we have that

$$t_i = \prod_{\substack{J \subseteq I \\ i \in J}} x_J.$$

A further very useful property of relative greatest common divisors is that $(x_J, x_K) = 1$ if $J \not\subseteq K$ and $K \not\subseteq J$. We prove this property as the following lemma (see also [9, p. 2]).

LEMMA 4.1. *Let $t_1, t_2, \dots, t_k \in \mathbb{N}, J, K \subseteq \{1, 2, \dots, k\}, J, K \neq \emptyset$ and define the corresponding relative greatest common divisors x_J and x_K as above. If $J \not\subseteq K$ and $K \not\subseteq J$ then $(x_J, x_K) = 1$.*

Proof. By assumption $J \not\subseteq K$ and $K \not\subseteq J$ and thus we have that $J \subsetneq J \cup K$ and $K \subsetneq J \cup K$. We suppose that $d = (x_J, x_K) > 1$ and choose an arbitrary prime divisor $p|d$. Set $L = J \cup K, J = \{j_1, j_2, \dots, j_{|J|}\}, K = \{k_1, k_2, \dots, k_{|K|}\}, L = \{l_1, l_2, \dots, l_{|L|}\}$ and write

$$x_J = \frac{(t_{j_1}, t_{j_2}, \dots, t_{j_{|J|}})}{\left(\prod_{\substack{J' \subseteq I \\ J \subsetneq J' \\ L \not\subseteq J'}} x_{J'} \right) \cdot x_L \cdot \left(\prod_{L \subsetneq J'} x_{J'} \right)}$$

$$x_K = \frac{(t_{k_1}, t_{k_2}, \dots, t_{k_{|K|}})}{\left(\prod_{\substack{K' \subseteq I \\ K \subsetneq K' \\ L \not\subseteq K'}} x_{K'} \right) \cdot x_L \cdot \left(\prod_{L \subsetneq K'} x_{K'} \right)}$$

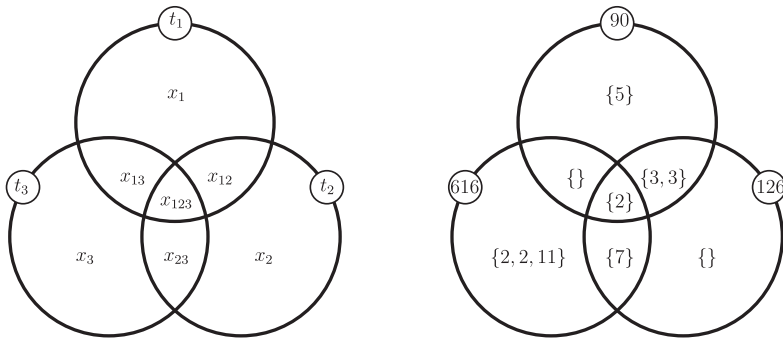


Fig. 1 - B/W online, B/W in print

Figure 1. A visualization of relative greatest common divisors using Venn diagrams. On the left-hand side one sees the general case of three positive integers t_1, t_2 and t_3 and on the right-hand side the situation when $t_1 = 90, t_2 = 126$ and $t_3 = 616$. Empty sets correspond to empty products and we set the corresponding relative greatest common divisor to 1.

With $x_L = (t_{l_1}, t_{l_2}, \dots, t_{l_{|L|}}) / \prod_{\substack{L' \subseteq I \\ L \not\subseteq L'}} x_{L'}$ this simplifies to

$$x_J = \frac{(t_{j_1}, t_{j_2}, \dots, t_{j_{|J|}})}{\left(\prod_{\substack{J' \subseteq I \\ J \not\subseteq J' \\ L \not\subseteq J'}} x_{J'} \right)} \cdot (t_{l_1}, t_{l_2}, \dots, t_{l_{|L|}}), \quad x_K = \frac{(t_{k_1}, t_{k_2}, \dots, t_{k_{|K|}})}{\left(\prod_{\substack{K' \subseteq I \\ K \not\subseteq K' \\ L \not\subseteq K'}} x_{K'} \right)} \cdot (t_{l_1}, t_{l_2}, \dots, t_{l_{|L|}}). \tag{4.2}$$

Let p^α be the highest power of p dividing the greatest common divisor of the terms $(t_{j_1}, t_{j_2}, \dots, t_{j_{|J|}})$ and $(t_{k_1}, t_{k_2}, \dots, t_{k_{|K|}})$. Thus p^α is also the highest power of p such that

$$p^\alpha | ((t_{j_1}, t_{j_2}, \dots, t_{j_{|J|}}), (t_{k_1}, t_{k_2}, \dots, t_{k_{|K|}})) = (t_{l_1}, t_{l_2}, \dots, t_{l_{|L|}}).$$

By definition of the greatest common divisor, without loss of generality, we may suppose that $\nu_p((t_{j_1}, t_{j_2}, \dots, t_{j_{|J|}})) = \alpha$. From equation (4.2), we finally see that $\nu_p(x_J) = 0$, a contradiction to $p|d$. \square

Relative greatest common divisors may be nicely visualized via Venn diagrams (especially when $k \leq 3$). We identify positive integers with the multiset of its prime divisors, that is, each prime p dividing n occurs with multiplicity $\nu_p(n)$ in the multiset. Given the Venn diagram of the multisets corresponding to the integers t_1, \dots, t_k , each area of intersection in the diagram uniquely corresponds to a relative greatest common divisor $x_J, J \subseteq \{1, \dots, k\}$. Figure 1 shows the situation for relative greatest common divisors of three positive integers t_1, t_2 and t_3 .

As mentioned in the beginning of this section relative greatest common divisors were systematically described in [9]. Nonetheless concepts of a similar type date back at least as far as Dedekind [8] who called the relative greatest common divisors of the integers t_1, \dots, t_k the cores (Kerne) of the system (t_1, \dots, t_k) . Dedekind described the construction of these cores explicitly for systems with three and four

elements and developed some theory to describe the cores of systems with more than four elements.

Decompositions similar to relative greatest common divisors also occur when we look for generalizations of the formula

$$[t_1, t_2] = \frac{t_1 t_2}{(t_1, t_2)}, \tag{4.3}$$

where $[t_1, t_2]$ denotes the least common multiple of the integers t_1 and t_2 . A generalization of formula (4.3) to least common multiples and greatest common divisors of k integers t_1, \dots, t_k was found by V.-A. Lebesgue [22, p. 350], who proved that

$$[t_1, t_2, \dots, t_k] = \frac{\prod_{\substack{1 \leq i \leq k \\ i \text{ odd}}} G_i}{\prod_{\substack{1 \leq j \leq k \\ j \text{ even}}} G_j},$$

where the variables G_i denote the product of the greatest common divisors of all choices of subsets of i integers in the set $\{t_1, t_2, \dots, t_k\}$.

5. Sums of three unit fractions

In this section, we deal with equation (4.1) for $k = 3$, i.e. with equations of the form

$$\frac{m}{n} = \frac{1}{n_1 t_1} + \frac{1}{n_2 t_2} + \frac{1}{n_3 t_3}, \tag{5.1}$$

where $n_1 t_1 \leq n_2 t_2 \leq n_3 t_3$, $n_i | n$ and $(n/n_i, t_i) = 1$ for $i \in \{1, 2, 3\}$. In the following we use the concept of relative greatest common divisors introduced in the previous section to get a suitable parametrization of the solutions of (5.1) corresponding to a fixed pattern $(n_1, n_2, n_3) \in \mathbb{N}^3$.

Writing the variables t_i in terms of relative greatest common divisors, equation (5.1) takes the form

$$\frac{m}{n} = \frac{1}{n_1 x_1 x_{12} x_{13} x_{123}} + \frac{1}{n_2 x_2 x_{12} x_{23} x_{123}} + \frac{1}{n_3 x_3 x_{13} x_{23} x_{123}} \tag{5.2}$$

and multiplying out yields

$$m x_1 x_2 x_3 x_{12} x_{13} x_{23} x_{123} = \frac{n}{n_1} x_2 x_3 x_{23} + \frac{n}{n_2} x_1 x_3 x_{13} + \frac{n}{n_3} x_1 x_2 x_{12}. \tag{5.3}$$

The first thing we observe is that we have $x_i = 1$ for all $i \in \{1, 2, 3\}$. This follows from lemma 4.1 and equation (5.3) together with the fact that $x_i | n/n_i$ is possible only if $x_i = 1$ by definition of n_i . We thus can work with the following simplified version of equation (5.3)

$$m x_{12} x_{13} x_{23} x_{123} = \frac{n}{n_1} x_{23} + \frac{n}{n_2} x_{13} + \frac{n}{n_3} x_{12}. \tag{5.4}$$

Next, we introduce the parameters d_{ij} which are defined as $d_{ij} = (n/n_i, n/n_j)$. Again we have that $(x_{ij}, d_{ij}) = 1$ by definition of the n_i and we note that for given m, n and a fixed pattern (n_1, n_2, n_3) also the parameters d_{ij} are fixed.

In what follows, we apply methods developed by Elsholtz and Tao [12, §§ 2 and 3]. The strategy is to derive a system of equations from (5.4) and to make use of divisor relations therein. With the observation of coprimality of d_{ij} and x_{ij} , and using divisibility relations implied by equation (5.4) we may define the following three positive integers

$$w = \frac{\frac{n}{n_1 d_{13}} x_{23} + \frac{n}{n_3 d_{13}} x_{12}}{x_{13}}, y = \frac{\frac{n}{n_1 d_{12}} x_{23} + \frac{n}{n_2 d_{12}} x_{13}}{x_{12}} \text{ and } z = \frac{\frac{n}{n_2 d_{23}} x_{13} + \frac{n}{n_3 d_{23}} x_{12}}{x_{23}}.$$

Later we make use of the product of w and z which is given by

$$\begin{aligned} wz &= \frac{n}{n_1 d_{13}} \frac{n}{n_2 d_{23}} + \frac{x_{12}}{x_{13} x_{23}} \left(\frac{n^2}{n_1 n_3 d_{13} d_{23}} x_{23} + \frac{n^2}{n_2 n_3 d_{13} d_{23}} x_{13} + \frac{n^2}{n_3^2 d_{13} d_{23}} x_{12} \right) \\ &= \frac{n}{n_1 d_{13}} \frac{n}{n_2 d_{23}} + \frac{nx_{12}}{n_3 d_{13} d_{23} x_{13} x_{23}} \left(\frac{n}{n_1} x_{23} + \frac{n}{n_2} x_{13} + \frac{n}{n_3} x_{12} \right) \\ &= \frac{n}{n_1 d_{13}} \frac{n}{n_2 d_{23}} + \frac{nm}{n_3 d_{13} d_{23}} x_{12}^2 x_{123}, \end{aligned}$$

where we used equation (5.4) to get the last equality. We collect the equations just derived in the following list

$$mx_{12}x_{13}x_{23}x_{123} = \frac{n}{n_1} x_{23} + \frac{n}{n_2} x_{13} + \frac{n}{n_3} x_{12} \tag{5.5}$$

$$yx_{12} = \frac{n}{n_1 d_{12}} x_{23} + \frac{n}{n_2 d_{12}} x_{13} \tag{5.6}$$

$$zx_{23} = \frac{n}{n_2 d_{23}} x_{13} + \frac{n}{n_3 d_{23}} x_{12} \tag{5.7}$$

$$mx_{13}x_{23}x_{123} = d_{12}y + \frac{n}{n_3} \tag{5.8}$$

$$mx_{12}x_{13}x_{123} = d_{23}z + \frac{n}{n_1} \tag{5.9}$$

$$wz = \frac{n}{n_1 d_{13}} \frac{n}{n_2 d_{23}} + \frac{nm}{n_3 d_{13} d_{23}} x_{12}^2 x_{123}. \tag{5.10}$$

For proving theorem 1.1 the classical divisor bound will play a crucial role. We will use it in the following form (see [18, theorem 315]).

LEMMA A. Let $d : \mathbb{N} \rightarrow \mathbb{N}$ be the divisor function, that is, $d(n) = \sum_{d|n} 1$. Then for every $\epsilon > 0$ we have

$$d(n) \ll_{\epsilon} n^{\epsilon}.$$

We now have all the tools we need to prove theorem 1.1.

Proof of theorem 1.1. Consider a solution of equation (5.1) for a fixed pattern (n_1, n_2, n_3) . By assumption we have $n_1 t_1 \leq n_2 t_2 \leq n_3 t_3$ and using the

parametrization of the t_i we introduced in equation (5.2) this implies

$$x_{13} \leq \frac{n_2}{n_1}x_{23} \quad \text{and} \quad x_{12} \leq \frac{n_3}{n_2}x_{13}.$$

Using these inequalities in equations (5.6) and (5.7) yields

$$yx_{12} \leq 2\frac{n}{n_1d_{12}}x_{23} \quad \text{and} \quad zx_{23} \leq 2\frac{n}{n_2d_{23}}x_{13}.$$

Dividing by x_{23} and x_{13} respectively and multiplying the last two inequalities we arrive at

$$\frac{yx_{12}}{x_{23}} \frac{zx_{23}}{x_{13}} \leq 4\frac{n^2}{n_1n_2d_{12}d_{23}}.$$

We now intend to obtain a lower bound for $n_1n_2d_{12}d_{23}$. Let $n = \prod_{p \in \mathbb{P}} p^{\nu_p(n)}$ be the prime factorization of n . Then $n_1 = \prod_{p \in \mathbb{P}} p^{\nu_p(n_1)}$ and $n_2 = \prod_{p \in \mathbb{P}} p^{\nu_p(n_2)}$ where $0 \leq \nu_p(n_1), \nu_p(n_2) \leq \nu_p(n)$ for all $p \in \mathbb{P}$. Since

$$d_{12} = \left(\frac{n}{n_1}, \frac{n}{n_2}\right) = \prod_{p \in \mathbb{P}} p^{\nu_p(n) - \max(\nu_p(n_1), \nu_p(n_2))}$$

we have

$$\begin{aligned} n_1n_2d_{12} &= \prod_{p \in \mathbb{P}} p^{\nu_p(n_1) + \nu_p(n_2) + \nu_p(n) - \max(\nu_p(n_1), \nu_p(n_2))} \\ &\geq \prod_{p \in \mathbb{P}} p^{\nu_p(n_1) + \nu_p(n_2) + \nu_p(n) - \nu_p(n_1) - \nu_p(n_2)} = n. \end{aligned}$$

This shows that $n_1n_2d_{12}d_{23} \geq n$ and thus

$$\frac{yx_{12}}{x_{23}} \frac{zx_{23}}{x_{13}} \ll n.$$

By assumption we have that n_1t_1 is the smallest denominator in equation (5.1). This implies that

$$\frac{m}{n} \leq \frac{3}{n_1t_1} \quad \text{and thus} \quad t_1 \leq \frac{3n}{mn_1} \ll \frac{n}{m}.$$

The bound in theorem 1.1 can finally be derived from the following inequality

$$y \cdot z \cdot x_{12}x_{13} \cdot (x_{12}x_{13})^2 = \frac{yx_{12}}{x_{23}} \frac{zx_{23}}{x_{13}} (x_{12}x_{13}x_{123})^2 \ll \frac{n^3}{m^2}. \tag{5.11}$$

This implies that at least one of the factors $y, z, x_{12}x_{13}$ and $x_{12}x_{123}$ is bounded by $\mathcal{O}((n^3/m^2)^{1/5})$.

If this is the case for y then by lemma A and equation (5.8) we have at most $\mathcal{O}_\epsilon(n^\epsilon)$ choices for the parameters x_{13}, x_{23} and x_{123} for every choice of y . The parameter x_{12} is then uniquely determined by (5.5).

Similarly, if z is the bounded parameter use lemma A and equation (5.9) to see that there are at most $\mathcal{O}_\epsilon(n^\epsilon)$ choices for the parameters x_{12} , x_{13} and x_{123} for every choice of z . Again the remaining parameter x_{23} is uniquely determined by (5.5).

Suppose that $x_{12}x_{13} \ll (n^3/m^2)^{1/5}$. By lemma A for every fixed choice of $x_{12}x_{13}$ we may choose the factors x_{12} and x_{13} in at most $\mathcal{O}_\epsilon(n^\epsilon)$ ways. For each of those choices lemma A and equation (5.7) imply that there are at most $\mathcal{O}_\epsilon(n^\epsilon)$ choices for the parameter x_{23} . As before the remaining parameter x_{123} is then fixed by (5.5).

Finally, we need to consider the case when $x_{12}x_{123}$ is the bounded factor. As in the previous case for any fixed choice of $x_{12}x_{123}$ we have at most $\mathcal{O}_\epsilon(n^\epsilon)$ choices for the factors x_{12} and x_{123} . Since equation (5.1) has no solutions for $m > 3n$ we have that $m \ll n$ and using equation (5.10) we see that for any fixed choice of x_{12} and x_{123} we have at most $\mathcal{O}_\epsilon(n^\epsilon)$ choices for the parameters w and z . With z , x_{12} and x_{123} fixed, x_{13} is uniquely determined by (5.9). The last parameter x_{23} is again uniquely determined by (5.5).

In any case, we have a bounded number of applications of the divisor bound from lemma A, say it was applied at most l times. Furthermore, the largest denominator on the right-hand side of equation (5.1) is bounded by $\mathcal{O}(n^4)$ (see also the proof of corollary 1.3 below), thus there exists a fixed positive constant c , such that the size of integers to which we apply the divisor bound is bounded by $\mathcal{O}(n^c)$. Setting $\tilde{\epsilon} = c\epsilon$ we hence have at most $\mathcal{O}_\epsilon(n^{\tilde{\epsilon}}(n^3/m^2)^{1/5})$ choices for the parameters x_{12} , x_{13} , x_{23} and x_{123} which uniquely determine a solution of (5.1) if n_1 , n_2 and n_3 are fixed. Note that this bound is independent of the concrete choice of the parameters n_i and again by lemma A we have at most $\mathcal{O}_\epsilon(n^{3\epsilon})$ choices for the pattern (n_1, n_2, n_3) . Theorem 1.1 now follows by redefining the choice of ϵ . □

Finally, we prove corollary 1.3.

Proof of corollary 1.3. The proof of theorem 1.1 suggests an algorithm for computing all decompositions of a rational number m/n as a sum of three unit fractions. The running time of this algorithm depends on the quality of algorithms used for integer factorization. In [23] a probabilistic algorithm is analysed which finds all prime factors of a given integer in expected running time $\exp((1 + o(1))\sqrt{\log n \log \log n})$ for $n \rightarrow \infty$, which is clearly $\mathcal{O}_\epsilon(n^\epsilon)$. Lenstra and Pomerance [23, § 12] point out, that here the term probabilistic means that the algorithm is allowed to call a random number generator which outputs 0 or 1 each with probability 1/2. The term expected running time refers to averaging over the output of the random number generator only and not over the input n . Hence the expected running time is also valid for each individual n .

As a consequence, using an algorithm of this type, all decompositions of m/n as a sum of three unit fractions can be found by carrying out the following steps. Factorize the integer n and compute all possible patterns (n_1, n_2, n_3) . For any of these $\mathcal{O}_\epsilon(n^\epsilon)$ patterns it follows from the calculations in the proof of theorem 1.1, that the implied constant in inequality (5.11) may be chosen as $C := (36/n_1^2 d_{23})$. For all choices of integers y , z , $x_{12}x_{13}$ and $x_{12}x_{123} \in [1, C^{1/5}(n^3/m^2)^{1/5}]$ we determine the integers x_{12} , x_{13} , x_{23} and x_{123} via factoring $x_{12}x_{13}$, $x_{12}x_{123}$ and a small number of integers mentioned in formulae (5.5)-(5.10). All in all, this leads to an algorithm of expected running time $\mathcal{O}_\epsilon(n^\epsilon(n^3/m^2)^{1/5})$.

As for representations of the form

$$\frac{m}{n} = \sum_{i=1}^k \frac{1}{a_i} \tag{5.12}$$

with $k > 3$ we enumerate all possible choices for the denominators a_i , $1 \leq i \leq k - 3$, and apply our algorithm for finding representations as sum of three unit fractions to determine all choices for the remaining three denominators, that is, we solve

$$\frac{m}{n} - \sum_{i=1}^{k-3} \frac{1}{a_i} = \frac{1}{a_{k-2}} + \frac{1}{a_{k-1}} + \frac{1}{a_k}. \tag{5.13}$$

We suppose the denominators a_i in equation (5.12) are given in increasing order and prove upper bounds for the size of a_i , $1 \leq i \leq k$. In particular, we use an induction argument to show that $a_i \leq \alpha_i n^{2^{i-1}}$ where the finite sequence $(\alpha_i)_{1 \leq i \leq k}$ is recursively defined by $\alpha_1 = k$ and $\alpha_i = (k - i + 1) \prod_{j < i} \alpha_j$ for $2 \leq i \leq k$. For $i = 1$ this bound follows easily from the following inequality

$$\frac{m}{n} = \frac{1}{a_1} + \dots + \frac{1}{a_k} \leq \frac{k}{a_1}$$

which leads to $a_1 \leq \frac{kn}{m} \leq kn$. If we suppose the bound holds for a_i , with a similar argument we get

$$\frac{m}{n} - \frac{1}{a_1} - \dots - \frac{1}{a_i} = \frac{1}{a_{i+1}} + \dots + \frac{1}{a_k} \leq \frac{(k-i)}{a_{i+1}}.$$

The last inequality together with the induction hypothesis for $j < i + 1$ implies

$$\begin{aligned} a_{i+1} &\leq (k-i) \frac{n \prod_{j < i+1} a_j}{m \prod_{j < i+1} a_j - n \sum_{j < i+1} \prod_{\substack{l < i+1 \\ l \neq j}} a_l} \\ &\leq (k-i)n \prod_{j < i+1} a_j \leq \alpha_{i+1} n^{2^i}. \end{aligned}$$

By definition α_i is a polynomial in k of degree 2^i with leading coefficient 1. Furthermore, the denominator of the rational number on the left-hand side of equation (5.13) is of size at most $n \prod_{i=1}^{k-3} a_i \ll_k n^{2^{k-3}}$. By the aforementioned result, we can compute all decompositions as a sum of three unit fractions of this number in time $\mathcal{O}_{\epsilon,k}(n^{2^{k-3}(3/5+\epsilon)})$. We have to compute these representations for at most $\prod_{i=1}^{k-3} a_i \ll_k n^{2^{k-3}-1}$ rational numbers which leads to an upper bound of

$$\mathcal{O}_{\epsilon,k}(n^{2^{k-3}(8/5+\epsilon)-1})$$

for the running time. □

REMARK 5.1. The procedure for computing representations as a sum of k unit fractions as described in the proof of corollary 1.3 could lead to a speedup for calculations similar to those in [2]. In the calculations above the size of the numerator

of the rational number on the left hand side of equation (5.13), which we denote by m'/n' , was not taken into account. We note that also the proof of the upper bound for $f_3(m, n)$ by Browning and Elsholtz [5, theorem 2] may be similarly turned into an algorithm of running time $\mathcal{O}_\epsilon(n^\epsilon(n/m)^{2/3})$. In practice one would check dynamically if $m' \ll (n')^{1/4}$ before computing the representations as a sum of three unit fractions of $\frac{m'}{n'}$. If this is the case, the algorithm described in the first part of the proof of corollary 1.3 should be applied, if $m' \gg (n')^{1/4}$ the method of [5] should be used.

6. Sums of k unit fractions

In this section, we will prove theorem 1.4. Browning and Elsholtz used an induction argument on their bound for the quantity $f_3(m, n)$ to get bounds for $f_k(m, n)$ for $k \geq 4$. Using their arguments directly on our result from theorem 1.1 would lead to worse upper bounds than those of Browning and Elsholtz. The reason is that our bound for $f_3(m, n)$ is weaker than the one in [5] when m is large.

As in [5, § 4] the proof of theorem 1.4 will be based on the observation that from equation (4.1) it follows that

$$f_k(m, n) \leq \sum_{n/m < n_1 t_1 \leq kn/m} f_{k-1}(mn_1 t_1 - n, n_1 t_1 n),$$

which, after introducing the parameter $u = mn_1 t_1 - n$, becomes

$$f_k(m, n) \leq \sum_{\substack{0 < u \leq (k-1)n \\ m|u+n}} f_{k-1}\left(u, \frac{n(u+n)}{m}\right). \tag{6.1}$$

The improvement in theorem 1.4 stems from extending the method of Browning and Elsholtz by applying the following new idea. In the case of $k = 4$ we do not consider the sum on the right-hand side of (6.1) as a whole but we split the sum into two parts. In the first part we collect the values of u where $0 < u \leq n^\delta$ for some $0 < \delta < 1$ which will be chosen later. This sum will be small since it contains few summands.

The second part will consist of all summands where $u > n^\delta$. This corresponds to $n_1 t_1 > (n + n^\delta)/m$ which will force $n_2 t_2$ and $n_3 t_3$ to be small.

The following lemma B is [5, theorem 2].

LEMMA B. For any $\epsilon > 0$, we have

$$f_3(m, n) \ll_\epsilon n^\epsilon \left(\frac{n}{m}\right)^{2/3}.$$

In the proof of theorem 1.4 below we make use of lemma B rather than theorem 1.1. Furthermore, we will use a lifting procedure which was first used by Browning and

Elsholtz [5] to lift upper bounds of the form

$$f_5(m, n) \ll_{\epsilon} n^{\epsilon} \left(\frac{n^2}{m}\right)^c \tag{6.2}$$

to upper bounds for $f_k(m, n)$ for $k > 5$. For possible future use we write this procedure up in the following lemma and work through the original proof by Browning and Elsholtz with an arbitrary exponent $c > 1$ in (6.2).

LEMMA C. *Suppose that there exists $c > 1$ such that*

$$f_5(m, n) \ll_{\epsilon} n^{\epsilon} \left(\frac{n^2}{m}\right)^c .$$

Then for any $k \geq 5$ we have

$$f_k(m, n) \ll_{\epsilon} (kn)^{\epsilon} \left(\frac{k^{4/3}n^2}{m}\right)^{c2^{k-5}} .$$

Proof. We will inductively show that for $k \geq 5$ there exists Θ_k depending on k such that we have

$$f_k(m, n) \ll_{\epsilon} (kn)^{\epsilon} \left(\frac{k^{\Theta_k}n^2}{m}\right)^{c2^{k-5}} \tag{6.3}$$

and we note that this is certainly true for $k = 5$ by assumption. The proof works in three steps.

1. Establish an upper bound where the implied constant is allowed to depend on k .

For $k \geq 5$ we want to have a bound of the form

$$f_k(m, n) \ll_{k, \epsilon} n^{\epsilon} \left(\frac{n^2}{m}\right)^{c2^{k-5}} \tag{6.4}$$

where the implied constant is allowed to depend on k . An upper bound of this type may easily be achieved via (6.1). Indeed this bound holds true for $k = 5$ by assumption and assuming its existence for $f_k(m, n)$ we find for $f_{k+1}(m, n)$

$$\begin{aligned} f_{k+1}(m, n) &\ll \sum_{\substack{0 < u \leq kn \\ m | u+n}} f_k\left(u, \frac{n(u+n)}{m}\right) \ll_{k, \epsilon} n^{\epsilon} \left(\frac{n^2}{m}\right)^{c2^{k-4}} \sum_{u=1}^{\infty} \frac{1}{uc2^{k-5}} \\ &\ll_{k, \epsilon} n^{\epsilon} \left(\frac{n^2}{m}\right)^{c2^{k-4}} , \end{aligned}$$

where we used that $c > 1$.

2. Use inequality (6.1) and split the sum into two parts.

For the upper bound where the implied constant is independent of k we again suppose it to be true for $f_k(m, n)$ with $k \geq 5$ and inductively prove it to hold for $f_{k+1}(m, n)$. Using inequalities (6.1) and (6.3) we get

$$\begin{aligned} f_{k+1}(m, n) &\ll \sum_{\substack{0 < u \leq kn \\ m | u+n}} f_k \left(u, \frac{n(u+n)}{m} \right) \\ &\ll \sum_{\substack{0 < u \leq (L-1)n \\ m | u+n}} f_k \left(u, \frac{n(u+n)}{m} \right) + \sum_{\substack{(L-1)n < u \leq kn \\ m | u+n}} f_k \left(u, \frac{n(u+n)}{m} \right) \\ &\ll (kn)^\epsilon k^{\Theta_k c 2^{k-5}} \left(\frac{n^2}{m} \right)^{c 2^{k-4}} \times \\ &\quad \left(\sum_{0 < u \leq (L-1)n} \frac{1}{u^{c 2^{k-5}}} L^{c 2^{k-4}} + \sum_{(L-1)n < u \leq kn} \frac{1}{u^{c 2^{k-5}}} (k+1)^{c 2^{k-4}} \right). \end{aligned}$$

Since $c 2^{k-5} > 1$ the infinite sums over $\frac{1}{u^{c 2^{k-5}}}$ converge. For the first sum, we use that the sum is bounded by a constant for the second sum we use the following more accurate bound

$$\sum_{(L-1)n < u \leq kn} \frac{1}{u^{c 2^{k-5}}} \leq \sum_{u=L}^{\infty} \frac{1}{u^{c 2^{k-5}}} \ll \int_L^{\infty} \frac{1}{u^{c 2^{k-5}}} du \ll L^{1-c 2^{k-5}}.$$

Together with the fact that $(a + b)^\alpha \geq a^\alpha + b^\alpha$ for $a, b > 0$ and $\alpha > 1$ this shows that

$$\begin{aligned} f_{k+1}(m, n) &\ll_\epsilon ((k+1)n)^\epsilon (k+1)^{\Theta_k c 2^{k-5}} \left(\frac{n^2}{m} \right)^{c 2^{k-4}} \left(L^{c 2^{k-4}} + \left(\frac{k+1}{L^{1/2-(c 2^{k-4})^{-1}}} \right)^{c 2^{k-4}} \right) \\ &\ll_\epsilon ((k+1)n)^\epsilon (k+1)^{\Theta_k c 2^{k-5}} \left(\frac{n^2}{m} \right)^{c 2^{k-4}} \left(L + \frac{k+1}{L^{1/2-(c 2^{k-4})^{-1}}} \right)^{c 2^{k-4}}. \end{aligned}$$

3. Optimizing for L and determining an upper bound for Θ_k .

By the bound we derived in step 1 we may suppose that $k \geq \max\{(\log(2/3(c\epsilon)^{-1}))/\log 2 + 4, ((1 + \sqrt{5})/2)^{1/\epsilon} - 1\}$. With $L = (k+1)^{2/3}$ we get

$$\begin{aligned} f_{k+1}(m, n) &\ll_\epsilon ((k+1)n)^\epsilon (k+1)^{\Theta_k c 2^{k-5}} \left(\frac{n^2}{m} \right)^{c 2^{k-4}} (k+1)^{2/3 \cdot c 2^{k-4}} \left(1 + L^{(c 2^{k-4})^{-1}} \right)^{c 2^{k-4}} \\ &\ll_\epsilon (k+1)^{\epsilon(1+c 2^{k-3})} n^\epsilon (k+1)^{c 2^{k-4}(\Theta_k/2+2/3)} \left(\frac{n^2}{m} \right)^{c 2^{k-4}}. \end{aligned}$$

With $\Theta_{k+1} = \Theta_k/2 + 2/3$ and an appropriate choice of ϵ this implies

$$f_{k+1} \ll_{\epsilon} ((k+1)n)^{\epsilon} \left(\frac{(k+1)\Theta_{k+1}n^2}{m} \right)^{c2^{(k+1)-5}}$$

Since for $\Theta_5 \leq 4/3$ the sequence recursively defined by $\Theta_{k+1} = \Theta_k/2 + 2/3$ monotonically increases towards its limit $4/3$ we eventually get for any $k \geq 5$:

$$f_k(m, n) \ll_{\epsilon} (kn)^{\epsilon} \left(\frac{k^{4/3}n^2}{m} \right)^{c2^{k-5}}. \quad \square$$

Proof of theorem 1.4. In the following $\delta < 1$ is a fixed constant to be chosen at the end of the proof. We start with proving bounds on $f_4(m, n)$ and we write $f_4(m, n) = f_4^{(1)}(m, n) + f_4^{(2)}(m, n)$. Here $f_4^{(1)}(m, n)$ counts those solutions to equation (4.1) with $n_1t_1 \leq (n + n^\delta)/m$ and $f_4^{(2)}(m, n)$ those with $n_1t_1 > (n + n^\delta)/m$. From (6.1), we have that

$$\begin{aligned} f_4(m, n) &= f_4^{(1)}(m, n) + f_4^{(2)}(m, n) \leq \sum_{\substack{0 < u \leq n^\delta \\ m|u+n}} f_3\left(u, \frac{n(u+n)}{m}\right) + f_4^{(2)}(m, n) \\ &= S_1 + f_4^{(2)}(m, n). \end{aligned}$$

We use the following estimate (uniform in $a \in \mathbb{Z}$)

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod q}} n^{-\Theta} = \frac{x^{1-\Theta}}{(1+\Theta)q} + \mathcal{O}_{\Theta}(1). \tag{6.5}$$

To bound the sum S_1 we use (6.5) and lemma B to get

$$S_1 \ll_{\epsilon} n^{\epsilon} \left(\frac{n^2}{m} \right)^{2/3} \sum_{\substack{0 < u \leq n^\delta \\ m|u+n}} \frac{1}{u^{2/3}} \ll_{\epsilon} n^{\epsilon} \left(\frac{n^2}{m} \right)^{2/3} \left(\frac{n^{\delta/3}}{m} + 1 \right). \tag{6.6}$$

Next, we prove that

$$f_4^{(2)}(m, n) \ll_{\epsilon} n^{\epsilon} \frac{n^{(12-4\delta)/5}}{m^{8/5}}.$$

Since there are at most $\mathcal{O}_{\epsilon}(n^{\epsilon})$ distinct patterns (n_1, n_2, n_3, n_4) it suffices to prove this bound for all solutions counted by $f_4^{(2)}(m, n)$ corresponding to a fixed pattern. To get an upper bound for the contribution of $f_4^{(2)}(m, n)$ we thus suppose that (n_1, n_2, n_3, n_4) is fixed and note that the fact that $4n/m \geq n_1t_1 > (n + n^\delta)/m$

implies the following upper bound for n_2t_2 :

$$\frac{3}{n_2t_2} \geq \frac{mn_1t_1 - n}{nn_1t_1} \geq \frac{mn^\delta}{4n^2}.$$

Therefore we have

$$n_2t_2 \ll \frac{n^{2-\delta}}{m}. \tag{6.7}$$

We use again relative greatest common divisors and write a representation of m/n as a sum of four unit fractions as

$$\begin{aligned} \frac{m}{n} = & \frac{1}{n_1x_1x_{12}x_{13}x_{14}x_{123}x_{124}x_{134}x_{1234}} + \frac{1}{n_2x_2x_{12}x_{23}x_{24}x_{123}x_{124}x_{234}x_{1234}} \\ & + \frac{1}{n_3x_3x_{13}x_{23}x_{34}x_{123}x_{134}x_{234}x_{1234}} + \frac{1}{n_4x_4x_{14}x_{24}x_{34}x_{124}x_{134}x_{234}x_{1234}}. \end{aligned}$$

It is again easy to see that $x_1 = x_2 = x_3 = x_4 = 1$ and multiplying out the last equation yields

$$\begin{aligned} & mx_{12}x_{13}x_{14}x_{23}x_{24}x_{34}x_{123}x_{124}x_{134}x_{234}x_{1234} \\ & = \frac{n}{n_1}x_{23}x_{24}x_{34}x_{234} + \frac{n}{n_2}x_{13}x_{14}x_{34}x_{134} + \frac{n}{n_3}x_{12}x_{14}x_{24}x_{124} + \frac{n}{n_4}x_{12}x_{13}x_{23}x_{123}. \end{aligned} \tag{6.8}$$

From equation (6.8) we see that the quantity

$$z_{34} = \frac{(n/n_3)x_{12}x_{14}x_{24}x_{124} + (n/n_4)x_{12}x_{13}x_{23}x_{123}}{x_{34}}$$

is an integer and we use

$$z_{34}x_{34} = \frac{n}{n_3}x_{12}x_{14}x_{24}x_{124} + \frac{n}{n_4}x_{12}x_{13}x_{23}x_{123}. \tag{6.9}$$

By (6.7) and $4n/m \geq n_1t_1 > (n + n^\delta)/m$ we have

$$(t_1t_2)^4 = (x_{12}x_{13}x_{14}x_{123}x_{124}x_{134}x_{1234})^4(x_{12}x_{23}x_{24}x_{123}x_{124}x_{234}x_{1234})^4 \ll \frac{n^{12-4\delta}}{m^8}, \tag{6.10}$$

and we write

$$\begin{aligned} & (x_{12}x_{13}x_{14}x_{123}x_{124}x_{134}x_{1234})^4(x_{12}x_{23}x_{24}x_{123}x_{124}x_{234}x_{1234})^4 = \\ & (x_{12}x_{13}x_{14}x_{23}x_{24}x_{123}x_{124}x_{1234})(x_{12}x_{13}x_{23}x_{24}x_{123}x_{124}x_{134}x_{234}x_{1234}) \times \\ & (x_{12}x_{14}x_{23}x_{24}x_{123}x_{124}x_{134}x_{234}x_{1234})(x_{12}x_{13}x_{14}x_{24}x_{123}x_{124}x_{134}x_{234}x_{1234}) \times \\ & (x_{12}^4x_{13}x_{14}x_{23}x_{123}^4x_{124}^4x_{134}x_{234}x_{1234}^4). \end{aligned} \tag{6.11}$$

We show that each of the five factors in brackets on the right-hand side of the last equation corresponds to at most $\mathcal{O}_\epsilon(n^\epsilon)$ solutions of (6.8), where ϵ is an arbitrarily

small positive number. First we note that all factors are of polynomial size in n and by lemma A, given one of these factors, we have $\mathcal{O}_\epsilon(n^\epsilon)$ choices for all the x_{ij} , x_{ijk} and x_{1234} appearing as sub-factors.

Given positive integer constants C_0, C_1, C_2 and C_3 of size polynomial in n , we count the number of integer solutions (A, B) of the equation

$$C_0AB = C_1A + C_2B + C_3. \tag{6.12}$$

Rewriting this equation in the form

$$(C_0A - C_2)(C_0B - C_1) = C_0C_3 + C_1C_2$$

we see that the number of solutions (A, B) is bounded by $\mathcal{O}_\epsilon(n^\epsilon)$. For the second to the fifth factor on the right-hand side of (6.11) exactly two parameters are missing to uniquely determine a solution of (6.8). All of these factors miss the parameter x_{34} . The second one additionally misses x_{14} , the third one x_{13} , the fourth one x_{23} and the last one x_{24} . In all of these cases equation (6.8) provides an instance of (6.12) where the variables A and B correspond to the two missing parameters (the term containing both missing parameters on the right-hand side of (6.8) may be shifted to the left-hand side).

In the first factor on the right-hand side of (6.11) three parameters are missing. From equation (6.9) we see that we have at most $\mathcal{O}_\epsilon(n^\epsilon)$ choices for the parameter x_{34} . To see the same bound for the parameters x_{134} and x_{234} we use again that equations of type (6.12) can be factorized.

Since by (6.10) at least one of the factors on the right-hand side of (6.11) is $\mathcal{O}(n^{(12-4\delta)/5}/m^{8/5})$ we have that

$$f_4^{(2)}(m, n) \ll_\epsilon n^\epsilon \frac{n^{(12-4\delta)/5}}{m^{8/5}}. \tag{6.13}$$

Again we note that in the considerations above the divisor bound from lemma A was applied a bounded number of times and the bound in (6.13) follows upon redefining the choice of ϵ . Choosing $\delta = 16/17$ in (6.6) and (6.13) we get

$$f_4(m, n) \ll n^\epsilon \left(\frac{n^{4/3}}{m^{2/3}} + \frac{n^{28/17}}{m^{8/5}} \right). \tag{6.14}$$

To bound $f_5(m, n)$ we again use (6.1) and (6.5) and get

$$f_5(m, n) \ll n^\epsilon \sum_{\substack{0 < u \leq 4n \\ m|u+n}} \left(\left(\frac{n^2}{m} \right)^{4/3} \frac{1}{u^{2/3}} + \left(\frac{n^2}{m} \right)^{28/17} \frac{1}{u^{8/5}} \right) \ll n^\epsilon \left(\frac{n^2}{m} \right)^{28/17}. \tag{6.15}$$

Setting $c = 28/17$ in lemma C yields the bound in theorem 1.4. □

7. Lower bounds

Proof of theorem 1.6. To prove the first bound we are going to extend an idea used in the proof of [5, theorem 1]. As before, we use highly composite denominators

$n \in \mathbb{N}$, but here we show that there are many values a_1 with many corresponding pairs (a_2, a_3) giving a solution of

$$\frac{m}{n} = \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}.$$

To prove our lower bound for $f_3(m, n)$ we consider the set

$$\mathcal{N} = \left\{ mn' : n' = \prod_{i=1}^r p_i \right\},$$

where p_i is the i -th prime. In choosing the denominators $n \in \mathcal{N}$ we reduce the problem to finding many solutions of the equation

$$\frac{1}{n'} = \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}.$$

We set $a_1 = n' + d$, where d is any divisor of n' , and are left with

$$\frac{1}{n'} - \frac{1}{n' + d} = \frac{1}{n'(n'/d + 1)} = \frac{1}{a_2} + \frac{1}{a_3}.$$

For two divisors d_1 and d_2 of n' with $(d_1, d_2) = 1$ we have

$$\frac{1}{n'(n'/d + 1)} = \frac{1}{n'((n'/d + 1)/d_1)(d_1 + d_2)} + \frac{1}{n'((n'/d + 1)/d_2)(d_1 + d_2)}. \tag{7.1}$$

We note that for two pairs of divisors d_1, d_2 and d'_1, d'_2 with $(d_1, d_2) = 1$ and $(d'_1, d'_2) = 1$ it follows that

$$\frac{n'(n'/d + 1)}{d_1}(d_1 + d_2) = \frac{n'(n'/d + 1)}{d'_1}(d'_1 + d'_2) \Leftrightarrow \frac{d_1}{d_2} = \frac{d'_1}{d'_2}.$$

Since d_1 and d_2 as well as d'_1 and d'_2 are coprime we get $d_1 = d'_1$ and $d_2 = d'_2$. This implies that each pair (d_1, d_2) with $d_1 < d_2$ gives a unique solution of equation (7.1). Furthermore, for any choice of d, d_1, d_2 it follows that

$$n' + d < \frac{n'(n'/d + 1)}{d_2}(d_1 + d_2),$$

which altogether implies that by counting all possible choices for d, d_1, d_2 we get a lower bound for twice the value of $f_3(1, n')$.

Choosing n' as in the construction of the set \mathcal{N} , we have $2^{\omega(n')}$ choices for the divisor d and using the binomial theorem there are

$$\sum_{i=0}^{\omega(n')} \binom{\omega(n')}{i} \sum_{j=0}^{\omega(n')-i} \binom{\omega(n')-i}{j} = \sum_{i=0}^{\omega(n')} \binom{\omega(n')}{i} 2^{\omega(n')-i} = 3^{\omega(n')}$$

choices for the divisors d_1 and d_2 . As a consequence of the prime number theorem it is known that $\omega(n') \sim \log n'/\log \log n'$ and hence, for $n \in \mathcal{N}$

$$\begin{aligned} f_3(m, n) = f_3(1, n') &\geq \frac{1}{2} 2^{\omega(n')} 3^{\omega(n')} \geq \exp\left((\log 6 + o(1)) \frac{\log n'}{\log \log n'}\right) \\ &\geq \exp\left((\log 6 + o_m(1)) \frac{\log n}{\log \log n}\right). \end{aligned}$$

For the second bound we modify the idea used in the proof of [12, theorem 1.8]. For fixed $m \in \mathbb{N}$, as a consequence of the Turán-Kubilius inequality (see e.g. [29, p. 434]) we get that the set

$$\mathcal{M}_1 = \bigcap_{\substack{k \leq m \\ (k, m) = 1}} \left\{ n \in \mathbb{N} : \omega(n, k, m) = \left(\frac{1}{\varphi(m)} + o(1)\right) \log \log n \right\}$$

is a set with density one, that is, $\lim_{x \rightarrow \infty} |\{n \in \mathcal{M}_1 : n \leq x\}|/x = 1$.

For any $n \in \mathcal{M}_1$ we write $m/n = m'/n'$ with $(m', n') = 1$ and note that $\omega(n, k, m) = \omega(n', k, m')$ for all k with $(k, m) = 1$. By construction of the set \mathcal{M}_1 and since n' is coprime to m' , we find $(1/\varphi(m) + o(1)) \log \log n$ prime divisors p of n' in the residue class $-n' \pmod{m'}$. For any of these prime divisors we have

$$\frac{m'}{n'} - \frac{1}{(n' + p)/m'} = \frac{p}{n'(n' + p)/m'} = \frac{1}{n'(n'/p + 1)/m'}$$

where $(n'/p + 1)/m'$ is an integer. Again, by construction of the set \mathcal{M}_1 , for the number of prime factors of n' we have

$$\omega(n') \geq \omega(n) - \omega(m) = (1 + o_m(1)) \log \log n.$$

For two coprime divisors d_1 and d_2 of n' we construct decompositions of $1/(n'(n'/p + 1)/m')$ as a sum of two unit fractions as in (7.1). As above we see that for any prime divisor p of n' in the residue class $-n' \pmod{m'}$ there are at least $3^{\omega(n')}$ such decompositions and all of them are distinct.

Altogether this implies that for any $n \in \mathcal{M}_1$

$$\begin{aligned} f(m, n) &\geq \left(\frac{1}{\varphi(m)} + o(1)\right) 3^{\omega(n')} \cdot \log \log n \geq \left(\frac{1}{\varphi(m)} + o(1)\right) 3^{\omega(n) - \omega(m)} \cdot \log \log n \\ &\geq \exp((\log 3 + o_m(1)) \log \log n) \cdot \log \log n. \end{aligned}$$

Finally, we prove the improved lower bound on $f_3(4, n)$. To do so, we set

$$\mathcal{M}_2 = \left(\bigcap_{i \in \{1,3\}} \{n \in \mathbb{N} : \frac{\tau(n,4)}{4} \leq \tau(n,i,4)\} \right) \cap \{n \in \mathbb{N} : \omega(n) = (1 + o(1)) \log \log n\} \cap \{n \in \mathbb{N} : \tau(n) \geq (\log n)^{\log 2 + o(1)}\}.$$

The first two sets with $i = 1$ and $i = 3$ in the intersection in the definition of \mathcal{M}_2 have density 1 by [17, theorem 5]. For the third and the fourth set this is true by the Turán-Kubilius inequality (again see e.g. [29, p. 434]). Hence the set \mathcal{M}_2 has density 1 and we investigate what happens for n in a certain residue class modulo 4.

If $n \equiv 0 \pmod 4$, then $4/n = 1/(n/4)$ and for any divisor d of $n/4$ we have

$$\frac{1}{n/4} - \frac{1}{n/4 + d} = \frac{1}{(n/4)(n/(4d) + 1)}.$$

Since $\omega(n/4) \geq \omega(n) - 1$, with the same arguments as above, we conclude that the number of representations of $1/(n/4)(n/(4d) + 1)$ as a sum of two unit fractions is at least of order $3^{\omega(n/4)} = 3^{(1+o(1)) \log \log n}$. From $\tau(n) = \prod_{p|n} (\nu_p(n) + 1)$ we easily deduce that $\tau(n/4) \geq (1/3)\tau(n)$. Altogether, we thus get

$$f_3(4, n) \geq \frac{1}{3} \tau(n) 3^{\omega(n/4)} \geq \exp((\log 6 + o(1)) \log \log n).$$

If $n \equiv 2 \pmod 4$, then $n/2$ is odd and the same is true for all $\tau(n/2) = (1/2)\tau(n)$ divisors of $n/2$. We have $4/n = 2/(n/2)$ and for any divisor d of $n/2$

$$\frac{2}{n/2} - \frac{1}{n/2 + d/2} = \frac{1}{(n/2)((n/(2d) + 1)/2)}.$$

As above we get

$$f_3(4, n) \geq \tau\left(\frac{n}{2}\right) 3^{\omega(n)-1} \geq \exp((\log 6 + o(1)) \log \log n).$$

Finally, if $n \equiv r \pmod 4$ for $r \in \{1, 3\}$, we have $\tau(n, 4) = \tau(n)$ and by construction of the set \mathcal{M}_2 , we have more than $\tau(n)/4$ divisors d of n in the residue class $-r \pmod 4$. Again, for any of these divisors we have

$$\frac{4}{n} - \frac{1}{(n+d)/4} = \frac{1}{n((n/d+1)/4)}.$$

Applying the arguments used previously one more time, we find

$$f_3(4, n) \geq \frac{\tau(n)}{4} 3^{\omega(n)} \geq \exp((\log 6 + o(1)) \log \log n)$$

also in this case. □

REMARK 7.1. The difference in the constants in the exponential functions of the lower bounds on $f(m, n)$ and $f(4, n)$ for sets of integers with density one in

theorem 1.6 is basically due to cancellation effects when dealing with general m . In particular, we deal with $m/n = m'/n'$, where $(m', n') = 1$, and we would need to have good control of the number of divisors of n' in the residue class $-n' \pmod{m'}$ to get the $\log 6$ exponent also in the general case. However, if we do not ask about a lower bound holding for a set of density one within the positive integers, but for a set of integers of density one within the set \mathcal{S} of positive integers coprime to a given $m \in \mathbb{N}$, we may achieve the $\log 6$ exponent. To do so we replace the set \mathcal{M}_1 with

$$\mathcal{M}'_1 = \left(\bigcap_{\substack{1 \leq i \leq m \\ (i,m)=1}} \{n \in \mathbb{N} : \tau(n, i, m) \geq \frac{\tau(n)}{2\varphi(m)}\} \right) \cap \{n \in \mathbb{N} : \omega(n) = (1 + o(1)) \log \log n\} \cap \{n \in \mathbb{N} : \tau(n) \geq (\log n)^{\log 2 + o(1)}\} \cap \mathcal{S}.$$

Now we may use results from [17, theorem 5] as well as Turán-Kubilius like previously and get that \mathcal{M}'_1 has density one in \mathcal{S} . Instead of constructing the first denominator via shifts in prime factors of n we may use arbitrary divisors of n in this case, which leads to the improvement mentioned above.

Proof of theorem 1.9. We consider solutions corresponding to the pattern $(1, p, p)$. In equation (1.1) we suppose that a_1 is the denominator with $(a_1, p) = 1$ and we write $a_1 = t_1$, $a_2 = pt_2$ and $a_3 = pt_3$. We use the parametrization via relative greatest common divisors of the t_i and applying lemma 4.1 it is easy to see, that $x_1 = x_2 = x_3 = 1$ in this case. Hence we are looking for infinitely many primes $p \equiv e \pmod f$ such that for given $m \in \mathbb{N}$ the equation

$$\frac{m}{p} = \frac{1}{x_{12}x_{13}x_{123}} + \frac{1}{px_{12}x_{23}x_{123}} + \frac{1}{px_{13}x_{23}x_{123}} \tag{7.2}$$

has many solutions. Multiplying equation (7.2) by the common denominator we get

$$mx_{12}x_{13}x_{23}x_{123} = px_{23} + x_{13} + x_{12}.$$

Setting $x_{12} + x_{13} = kx_{23}$, $M = \text{lcm}(m, f)$ and $x_{12} = M/m$ we deduce that

$$M \left(kx_{23} - \frac{M}{m} \right) x_{123} = p + k.$$

The residue class $(f - e) \equiv -e \pmod f$ splits into the residue classes $(f - e) + if \pmod M$, for $0 \leq i \leq m/(m, f) - 1$. Note, that $\text{gcd}(f, m/(m, f)) = 1$ hence the integers $i \cdot f$ for $0 \leq i \leq m/(m, f) - 1$ are a full system of residues modulo $m/(m, f)$. In particular there exists a $0 \leq j \leq m/(m, f) - 1$ such that $(f - e) + jf \equiv 1 \pmod{m/(m, f)}$. We set $k = (f - e) + jf$ and with $(e, f) = 1$ we altogether see that $(M, k) = 1$.

Now let $Q = \prod_{i=1}^r q_i$ where q_i is the i -th prime with $q_i \equiv -M/m \pmod k$ and $q_i > M$. Note that $\text{gcd}(M, Q) = 1$.

With $r = \lfloor \log t / (\varphi(k)C \log \log t) \rfloor$ we find that Q is of order $t^{1/C + o_f, m(1)}$. We now use Linnik’s theorem on primes in arithmetic progressions. As the modulus is very

smooth we can use an exponent of $C = 12/5 + o(1)$, due to Chang [6, corollary 11]. Hence we may find a prime p of order $M^C t^{1+o_{f,m}(1)}$ with

$$p \equiv -k \pmod{QM}.$$

This congruence implies that $p + k$ is divisible by the primes q_1, \dots, q_r and together with $k = (f - e) + jf$, we deduce that $p \equiv e \pmod{f}$ and $p + k \equiv 0 \pmod{M}$.

Let $l \in \mathbb{N}_0$ and S be a subset of size $l \operatorname{ord}_k(-M/m) + 1$ of the prime factors of Q . Hence $x_{23} = (\prod_{q \in S} q + M/m)/k$ is an integer and we set $x_{123} = (p + k)/(M \prod_{q \in S} q)$. We observe that any of these choices leads to a different solution of (7.2). To see this we look at the denominator $a_2 = px_{12}x_{23}x_{123}$ of the second fraction on the right-hand side of this equation. Suppose that two sets S and S' would lead to the same denominator a_2 . With $x_{12} = M/m$ this would imply the existence of $x_{23} \neq x'_{23}$ such that

$$p \frac{M}{m} x_{23} \frac{p + k}{M(kx_{23} - M/m)} = p \frac{M}{m} x'_{23} \frac{p + k}{M(kx'_{23} - M/m)}$$

from which we derive that

$$\frac{x_{23}}{x'_{23}} = \frac{kx_{23} - M/m}{kx'_{23} - M/m} = \frac{\prod_{q \in S} q}{\prod_{q' \in S'} q'}.$$

If $q \in S$ would divide x_{23} then q would also divide M/m , which is impossible by construction of Q . We hence have that $\prod_{q \in S} q / \prod_{q' \in S'} q' = 1$ and thus $S = S'$.

To count the number of solutions we get with the above construction, we make use of a formula which can be found in [3, theorem 1], for example, and which states

$$\sum_{i \geq 0} \binom{n}{iu} = \frac{1}{u} \sum_{j=0}^{u-1} (1 + \xi_u^j)^n, \tag{7.3}$$

where $\xi_u = \exp(2\pi i/u)$. Note that for the term corresponding to $j = 0$ in the sum on the right-hand side of (7.3) we get 2^n while for all other j we have $|1 + \xi_u^j| < 2$. Hence we deduce

$$\sum_{i \geq 0} \binom{n}{iu} = \frac{2^n}{u} (1 + o_u(1)).$$

The number of choices of the parameter x_{23} is

$$\begin{aligned} \sum_{i \geq 0} \binom{r}{i \operatorname{ord}_k(-M/m) + 1} &= \sum_{i \geq 0} \frac{r}{i \operatorname{ord}_k(-M/m) + 1} \binom{r-1}{i \operatorname{ord}_k(-M/m)} \\ &\geq \sum_{i \geq 0} \binom{r-1}{i \operatorname{ord}_k(-M/m)} \\ &= \frac{2^{r-1}}{\operatorname{ord}_k(-M/m)} (1 + o_{f,m}(1)). \end{aligned}$$

Plugging in $r = \lfloor \log t / (\varphi(k)C \log \log t) \rfloor$ and using that $p \leq M^C t^{1+o_{f,m}(1)}$ we get a lower bound of

$$\begin{aligned}
 f_3(m, p) &\gg_{f,m} \exp \left(\left(\frac{\log 2}{C\varphi(k)} + o_{f,m}(1) \right) \frac{\log t}{\log \log t} \right) \\
 &\gg_{f,m} \exp \left(\left(\frac{5 \log 2}{12 \operatorname{lcm}(m, f)} + o_{f,m}(1) \right) \frac{\log p}{\log \log p} \right).
 \end{aligned}
 \tag{7.4}$$

□

REMARK 7.2. The best-known exponent for Linnik’s theorem takes care of the worst case modulus and is 5 by work of Xylouris [30]. Chang’s result [6, corollary 11] considers smooth moduli (as in our situation) and allows for the better exponent $12/5 + o(1)$. Harman investigated, in connection with constructing Carmichael numbers, what happens if one is allowed to avoid a small set of exceptional moduli. In this situation, he improved the exponent to $1/0.4736$ (see [20, theorem 1.2] and [19] for some more explanation). As in our situation, we choose the modulus MQ , and hence can avoid ‘bad’ factors, it seems possible that theorem 1.9 can also be proved with a factor of 0.4736 instead of $5/12 = 0.4166\dots$ in the exponent of the lower bound on $f_3(m, p)$.

REMARK 7.3. If we consider the case $m = 4, f = 4$ and $e \in \{1, 3\}$ in theorem 1.9, we can explicitly compute k in the first line of (7.4). We simply have $k = 3$ if $e = 1$ and $k = 1$ if $e = 3$ hence we arrive at the lower bounds

$$f_3(4, p) \gg \exp \left((0.1444 + o(1)) \frac{\log p}{\log \log p} \right)$$

if $e = 1$ and

$$f_3(4, p) \gg \exp \left((0.2888 + o(1)) \frac{\log p}{\log \log p} \right)$$

if $e = 3$.

Acknowledgements

The authors acknowledge the support of the Austrian Science Fund (FWF): W1230. Furthermore, we would like to thank Igor Shparlinski and Glyn Harman for drawing our attention to the papers [6, 19, 20].

References

- 1 A. Aigner. Brüche als Summe von Stammbrüchen. *J. Reine Angew. Math.* **214/215** (1964), 174–179.
- 2 R. Arce-Nazario, F. Castro and R. Figueroa. On the number of solutions of $\sum_{i=1}^{11} 1/x_i = 1$ in distinct odd natural numbers. *J. Number Theory* **133** (2013), 2036–2046.
- 3 A. T. Benjamin, B. Chen and K. Kindred. Sums of evenly spaced binomial coefficients. *Math. Mag.* **83** (2010), 370–373.
- 4 L. Brenton and R. Hill. On the Diophantine equation $1 = \sum 1/n_i + 1/\prod n_i$ and a class of homologically trivial complex surface singularities. *Pacific J. Math.* **133** (1988), 41–67.
- 5 T. D. Browning and C. Elsholtz. The number of representations of rationals as a sum of unit fractions. *Illinois J. Math.* **55** (2011), 685–696.

- 6 M.-C. Chang. Short character sums for composite moduli. *J. Anal. Math.* **123** (2014), 1–33.
- 7 Y.-G. Chen, C. Elsholtz and L.-L. Jiang. Egyptian fractions with restrictions. *Acta Arith.* **154** (2012), 109–123.
- 8 R. Dedekind. Über Zerlegungen von Zahlen durch ihre größten gemeinsamen Teiler. In *Gesammelte mathematische Werke, Zweiter Band* (eds. R. Fricke, E. Noether and Ö. Ore). pp. 103–147 (Friedr. Vieweg & Sohn Akt.-Ges.: Braunschweig, 1931) Available online at <http://resolver.sub.uni-goettingen.de/purl?PPN235693928>.
- 9 C. Elsholtz. *Sums of k Unit Fractions*, Shaker Verlag, Aachen, 1999, Phd Thesis, Technische Universität Darmstadt (1998), 109 pages.
- 10 C. Elsholtz. Sums of k unit fractions. *Trans. Amer. Math. Soc.* **353** (2001), 3209–3227.
- 11 C. Elsholtz. Egyptian fractions with odd denominators. *Q. J. Math.* **67** (2016), 425–430.
- 12 C. Elsholtz and T. Tao. Counting the number of solutions to the Erdős–Straus equation on unit fractions. *J. Aust. Math. Soc.* **94** (2013), 50–105.
- 13 R. L. Graham. On finite sums of reciprocals of distinct n th powers. *Pacific J. Math.* **14** (1964), 85–92.
- 14 R. L. Graham. On finite sums of unit fractions. *Proc. London Math. Soc. (3)* **14** (1964), 193–207.
- 15 R. L. Graham. Paul Erdős and Egyptian fractions. In *Erdős Centennial, Bolyai Soc. Math. Stud.* (eds. L. Lovász, I. Z. Ruzsa and V. T. Sós), vol. 25, pp. 289–309 (Budapest: János Bolyai Math. Soc., 2013).
- 16 R. K. Guy. *Unsolved problems in number theory*, 3rd edn (New York: Springer-Verlag, 2010).
- 17 R. R. Hall. On the distribution of divisors of integers in residue classes (mod k). *J. Number Theory* **2** (1970), 168–188.
- 18 G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*, 6th edn (Oxford: Oxford University Press, 2008).
- 19 G. Harman. On the number of Carmichael numbers up to x . *Bull. London Math. Soc.* **37** (2005), 641–650.
- 20 G. Harman. Watt’s mean value theorem and Carmichael numbers. *Int. J. Number Theory* **4** (2008), 241–248.
- 21 S. V. Konyagin. Double exponential lower bound for the number of representations of unity by Egyptian fractions. *Math. Notes* **95** (2014), 277–281, Translation of *Mat. Zametki* **95** (2014), no. 2, 312–316.
- 22 V.-A. Lebesgue. Extrait des exercices d’analyse numérique, *Nouvelles annales de mathématiques 1^{re} série* **8** (1849), 347–353.
- 23 H. W. Lenstra, Jr. and C. Pomerance. A rigorous time bound for factoring integers. *J. Amer. Math. Soc.* **5** (1992), 483–516.
- 24 G. Martin. Dense Egyptian fractions. *Trans. Amer. Math. Soc.* **351** (1999), 3641–3657.
- 25 L. J. Mordell. *Diophantine equations. Pure and applied mathematics*, vol. 30 (London and New York: Academic Press, 1969).
- 26 L. A. Rosati. Sull’equazione diofantea $4/n = 1/x_1 + 1/x_2 + 1/x_3$. *Boll. Un. Mat. Ital. (3)* **9** (1954), 59–63.
- 27 C. Sándor. On the number of solutions of the Diophantine equation $\sum_{i=1}^n 1/x_i = 1$. *Period. Math. Hungar.* **47** (2003), 215–219.
- 28 B. M. Stewart. *Theory of numbers*, 2nd edn (New York, London: The Macmillan Company, Collier-Macmillan Limited, 1964).
- 29 G. Tenenbaum. *Introduction à la théorie analytique et probabiliste des nombres*, 3rd edn (Paris: Éditions Belin, 2008).
- 30 T. Xylouris. *Über die Nullstellen der Dirichletschen L -Funktionen und die kleinste Primzahl in einer arithmetischen Progression*, Universität Bonn, PhD thesis (2011).