On the exceptional set for binary Egyptian fractions

Jing-Jing Huang and Robert C. Vaughan

Abstract
For fixed integer $a \geq 3$, we study the binary Diophantine equation $\frac{a}{n} = \frac{1}{x} + \frac{1}{y}$ and in particular the number $E_a(N)$ of $n \leq N$ for which the equation has no positive integer solutions in $x$ and $y$. The asymptotic formula

$$E_a(N) \sim C(a) \frac{N(\log \log N)^{2^{m-1}-1}}{(\log N)^{1-1/2m}},$$

as $N$ goes to infinity, is established in this article, and this improves the best result in the literature dramatically. The proof depends on a very delicate analysis of a certain combinatorial property of the underlying group $(\mathbb{Z}/a\mathbb{Z})^*$ and $m$ depends in a subtle way on the factorization of $a$.

1. Introduction

Let $a$ be a fixed positive integer. We consider the binary Diophantine equation

$$\frac{a}{n} = \frac{1}{x} + \frac{1}{y},$$

and denote by $R(n; a)$ the number of pairs of positive integer solutions $(x, y)$ satisfying equation (1). A good deal is now known about the average behaviour of $R(n; a)$. See [1, 6, 7] for details. In this paper, we are concerned with the number of $n$ such that (1) is not soluble in positive integers $x$ and $y$ and to this end we define

$$\mathcal{E}_a = \{n \in \mathbb{N} : R(n; a) = 0\}.$$ 

Clearly, both $\mathcal{E}_1$ and $\mathcal{E}_2$ are empty. When $a \geq 3$ the structure of $\mathcal{E}_a$ is more delicate and of great interest. In this paper, we investigate the asymptotic size of $\mathcal{E}_a$. Thus, we define

$$\mathcal{E}_a(N) = \{n \in \mathcal{E}_a : n \leq N\}$$

and

$$E_a(N) = \#\mathcal{E}_a(N).$$

In 1985, Hofmeister and Stoll [5] proved that the set $\mathcal{E}_a$ has asymptotic density 0, and more precisely that

$$E_a(N) \ll_a \frac{N}{(\log N)^{1/\varphi(a)}}.$$ 

For $a = 5$ and $a \geq 7$ this bound is far from the truth. Their method is based on the observation that if (1) is insoluble, then $n$ is not divisible by any prime of the form $p = -1 \pmod{a}$. Thus, a
simple application of Selberg's upper bound sieve gives the stated bound. However, when \( a = 5 \) or \( a \geq 7 \) the bulk of the \( n \) deficient in such prime factors nevertheless have a representation.

In 1998, in his thesis, Elsholtz [3] showed that \( E_a(N) \ll_a N/\sqrt{\log N} \) for all \( a \geq 1 \) and that \( E_a(N) \gg_a N/\log N \) for \( a \equiv 0, 3 \pmod{4} \) and stated without proof the case \( a \) prime of Theorem 1.1. After the proof of Theorem 1.1 appeared on arXiv he showed us an unpublished manuscript in which a proof of the case \( a = 13 \) is outlined. His proof is based on an elementary argument which essentially exhausts all the possible factorizations of the elements in \( \mathcal{E}_{13} \). Unfortunately, the convoluted nature of his argument makes it unlikely that it would generalize.

The aim of this paper is to obtain the asymptotic behaviour of \( E_a(N) \) in all cases, and this is done in Theorem 1.1. The composite case is considerably more complicated than the prime case.

The main innovation in this paper is to take a group theoretic perspective and this leads to a combinatorial framework to study the factorizations of the elements in \( \mathcal{E}_a \) for \( a \geq 3 \). The analytic tools of this paper are rather routine and have been known for a long time. Whilst refinements are possible we do not pursue this as it would obscure the main thrust of the paper. We therefore rest that part of the argument on the work of others.

**Theorem 1.1.** For fixed \( a \geq 3 \), let \( 2^{\gamma_0} p_1^{\gamma_1} p_2^{\gamma_2} \ldots p_k^{\gamma_k} \) be the canonical decomposition of \( a \) and define \( m \) and \( \delta \) by

\[
2^m \| \gcd(\delta, p_1 - 1, p_2 - 1, \ldots, p_k - 1)
\]

and

\[
\delta = \begin{cases} 
0 & \text{when } \gamma_0 \leq 1, \\
2 & \text{when } \gamma_0 \geq 2.
\end{cases}
\]

Then we have

\[
E_a(N) \sim C(a) \frac{N(\log \log N)^{2^{m-1}+1}}{(\log N)^{1/2^m}},
\]

where \( C(a) \) is a positive constant depending only on \( a \).

In order to establish this theorem we need first to investigate the underlying structure of \( \mathcal{E}_a \), and we embark on this in Section 2. For general \( a \) this involves a novel group theoretic argument. The case when \( a \) is a prime power is somewhat easier to understand and, having established some preliminary lemmata in Subsection 2.1, we consider this case in Subsection 2.2. This then leads into a discussion of the general case in Subsection 2.3.

In Section 3 the main analytic input is introduced, and it is convenient to base this on an arithmetical application of a theorem of Delange. Delange’s theorem is a refinement of the Wiener–Ikehara theorem and is qualitative in nature. In particular it does not give an explicit error term. By using instead a method allied to that leading to the strongest known unconditional error term in the prime number theorem it would be possible to give a quantitative error term in Lemma 3.2 of a similar quality. However whilst this would be quite routine in nature there would be many detailed complications and more importantly the extra effort would not lead to any further illumination of the central problem of this paper in that a greater loss in the error term appears at a later stage of our argument. We are happy to leave this approach as an exercise to the reader.

The proof of the main theorem is completed in Section 4 through a suitable combination of Sections 2 and 3.

The referee has raised the question as to whether lower order main terms can be extracted from our argument. Based on the ideas in this paper, a plausible speculation is that \( E_a(N) \)
behaves something like

$$\sum_{\substack{H < G \, \text{maximal} \, \, \text{subgroup} \, \, \text{of} \\ \mathbb{Z}/a\mathbb{Z}^*}} \sum_{j_H = 0}^{k_H} C_{H,j_H}(a) \frac{N(\log \log N)^{ij_H}}{(\log N)^{1-1/[G:H]}} \left( \frac{e^{-i\pi/[G:H]}}{\pi} I(N; 1/[G:H], j_H) \right),$$

where $G = (\mathbb{Z}/a\mathbb{Z})^*$.

$$I(N; \theta, j) = \int_0^{\log N} \alpha^{-\theta} e^{-\alpha} f(-\alpha/\log N) \left( 1 - \frac{\log \alpha + i\pi}{\log \log N} \right)^j d\alpha,$$

$f(z)$ is analytic in the neighbourhood of 0 but depends on $G$, $H$ and $j$, and the quality of the error in this approximation would also depend on our knowledge of the zero–free region for the relevant Dirichlet $L$-functions. The extraction of leading terms from the integral is certainly possible, but there is a playoff between the upper endpoint of integration above and the error term which in turn predicates against the number of leading terms which can be extracted. As an aside one can say that this is reminiscent of expanding the logarithmic integral, as occurs in the prime number theorem, in negative powers of the logarithm. The error term in so doing is worse than the standard error term in the prime number theorem. The analytic machinery to obtain the above expression would be rather more complicated than that which we employ here, would involve the use of loop integrals about the point 1 and is very unlikely to reveal anything very interesting.

In Section 2 we essentially figure out how far $i_H$ should go when $H$ is a maximal subgroup of $(\mathbb{Z}/a\mathbb{Z})^*$ such that $-1 \not\in H$. For general $H$, this would require a version of Lemma 2.4 for the general group $\mathbb{Z}/M\mathbb{Z}$, $M \in \mathbb{N}$. See the comments after Lemma 2.4 for more details.

Throughout this paper, we reserve the letters $p$, $q$ and $r$ for prime numbers and calligraphic letters for sets and sequences. In particular, if $A \subseteq \mathbb{N}$ we denote by $A(N)$ the subset of $A$ with elements less than or equal to $N$ and $|A(N)|$ denotes the cardinality of $A(N)$. We also use Vinogradov’s ‘$\ll$’ notation, namely when we write $f(x) \ll g(x)$ we mean $|f(x)| \leq Cg(x)$ for some absolute constant $C$ and sufficiently large $x$. And accordingly, the notation $\ll_a$ means that the implicit constant defined above depends on the parameter $a$.

2. The structure of $\mathcal{E}_a$

2.1. Some elementary lemmata

It is more convenient to work with the notation $\mathcal{E}_a^*$, $\mathcal{E}_a^*(N)$ and $E_a^*(N)$, defined as follows:

$$\mathcal{E}_a^* = \{ n \in \mathcal{E}_a : (n, a) = 1 \},$$

and $\mathcal{E}_a^*(N)$ and $E_a^*(N)$ can be defined accordingly. Then we have immediately the following.

**Lemma 2.1.** We have

$$\mathcal{E}_a(N) = \bigcup_{d|a} \mathcal{E}_a^* N/d$$

and

$$E_a(N) = \sum_{d|a} E_a^* N/d.$$
The starting point of our argument is the following elementary lemma. This lemma has been discovered multiple times, see, for instance, Rav [8].

**Lemma 2.2.** Equation (1) with \((a, n) = 1\) is soluble in positive integers if and only if there exists a pair of coprime factors \(u\) and \(v\) of \(n\) such that \(a|u + v\).

**Proof.** If (1) is soluble, then we rewrite it as \(axy = n(x + y)\), let \((x, y) = l\) and write \(x = ul\) and \(y = vl\) with \((u, v) = 1\). Thus, \(aluv = n(u + v)\). Then, as \((a, n) = 1\) and \((uv, u + v) = 1\), we have \(uv|n\) and \(a|u + v\).

In the opposite direction, we write \(u + v = a\alpha'\) and \(n = u\beta\gamma\), so that \(a/n = 1/a'\beta\gamma n'\). \(\square\)

This lemma suggests that the solubility of (1) depends solely on the residue classes of factors of \(n\) modulo \(a\), and hence depends on the residue classes of prime factors of \(n\) modulo \(a\), which naturally leads our discussion to the distribution of prime factors of \(n\) in the multiplicative group \((\mathbb{Z}/a\mathbb{Z})^*\).

### 2.2. The case that \(a\) is a power of odd prime

We consider the case where \(a = p^\gamma\) is a power of odd prime in this subsection and come back to the general case later. This strategy fits with both the motivational purpose and the presentational purpose. Let \(G\) denote the cyclic group \((\mathbb{Z}/a\mathbb{Z})^*\) of reduced residue classes modulo \(a = p^\gamma\), and let \(H\) be the maximal subgroup of \(G = (\mathbb{Z}/a\mathbb{Z})^*\) with cardinality \(|H|\) being odd; namely, \(H\) is the maximal subgroup of \(G\) such that \(\bar{1} \notin H\) and clearly such a group is unique. Here and throughout this article 7 means the residue class \(i \pmod{a}\), if there is no ambiguity about the modulus \(a\) in the context. Now let \(\phi(a) = 2^m d\) with \(d\) being an odd number. If we fix a primitive root \(g\) modulo \(a\), then

\[G = \{g, g^2, g^3, \ldots, g^{2^m d}\}\]  

(2)

and

\[H = \{g^{2^m}, g^{2^m 2^m}, g^{3 2^m}, \ldots, g^{d 2^m}\},\]  

(3)

by which one readily verifies that \(\bar{1} \notin H\) since \(g^{\phi(a)/2} \equiv -1 \pmod{a}\). Hence, we have the index \(|G : H| = 2^m\) and \(|H| = \phi(a)/2^m = d\).

Essentially, the structure of \(E_a^*\) is that any \(n \in E_a^*\) can have arbitrarily many prime factors lying in the residue classes in \(H\) but can have at most a bounded number of prime factors lying outside \(H\). It is this observation that renders the counting function of \(E_a^*\) susceptible to an analytic argument.

**Lemma 2.3.** Let \(\mathbb{P}\) denote the set of prime numbers. Then we have the following inclusion relation of sets:

\[\{n \in \mathbb{N} : p|n, p \in \mathbb{P} \Rightarrow p \in H\} \subseteq E_a^*\]

**Proof.** For any \(n\) on the left-hand side, and for any pair of coprime positive integers \(u\) and \(v\) with \(uv|n\) we have \(\bar{u}, \bar{v} \in H\) in light of the fact that \(H\) is a group. Since \(\bar{1} \notin H\), we have \(\bar{u} \notin H\) and hence \(\bar{uv} \neq \bar{v}\), in other words \(a \nmid u + v\). Now Lemma 2.3 follows from Lemma 2.2. \(\square\)

The next lemma is central to our understanding of the structure of \(E_a^*\).
LEMMA 2.4. Let \( m \geq 1 \) and \( \mathcal{G} \) denote the additive group \( \mathbb{Z}/(2^m\mathbb{Z}) \), let \( \{e_j\}_1 \) be a sequence with \( t \) non-zero elements of \( \mathcal{G} \) (that is, repeated elements are allowed in \( \{e_j\} \)), and form the subset of \( \mathcal{G} \):

\[
S = \left\{ \sum_{j=1}^t \delta_j e_j : \delta_j \in \{-1, 0, 1\} \right\}.
\]

(i) If \( t \geq 2^{m-1} \), then \( \frac{2^m-1}{2} \in S \). Namely, as long as the length of \( \{e_j\} \) is at least \( 2^{m-1} \), for whatever choices of the elements \( e_j \), one can always find a partial sum, as in the definition of \( S \), such that it is equal to \( \frac{2^m-1}{2} \).

(ii) If \( t = 2^{m-1} - 1 \), then the corresponding set \( S \) does not contain \( \frac{2^m-1}{2} \) if and only if the sequence \( \{e_j\}_1 \) satisfies \( e_j \equiv \pm e \pmod{2^m} \) for each \( j \) and some fixed \( e \in (\mathbb{Z}/(2^m\mathbb{Z}))^* \).

Proof. We prove (i) first. Note that if one of the \( e_j \) satisfies \( e_j \equiv 0 \pmod{2^{m-1}} \), then (i) is automatically true by choosing the \( \delta_j \) for that particular \( j \) to be 1 and all the others to be 0. So, without loss of generality, we assume that none of the \( e_j \) for \( 1 \leq j \leq t \) is divisible by \( 2^{m-1} \). The proof is by induction on \( m \). The initial case \( m = 1 \) is trivial. Thus, we can suppose that \( m \geq 1 \) and that the conclusion is true for \( m \). Consider \( \mathcal{G} = \mathbb{Z}/(2^{m+1}\mathbb{Z}) \) and a sequence \( \{e_j\}_1^{2^m} \subseteq \mathcal{G} \). Note that none of the \( e_j \) is congruent to 0 modulo \( 2^m \) and hence, by the induction assumption, we know that there exist \( \delta_j \in \{-1, 0, 1\} \) for \( 1 \leq j \leq 2^m \) such that

\[
s_1 := \sum_{j=1}^{2^m-1} \delta_j e_j \equiv 2^{m-1} \pmod{2^m}
\]

and

\[
s_2 := \sum_{j=2^{m-1}+1}^{2^m} \delta_j e_j \equiv 2^{m-1} \pmod{2^m}.
\]

Choose \( u_i \) so that \( s_i = 2^m + u_i 2^m \) for \( i \in \{1, 2\} \). Then, by considering separately the cases when the \( u_i \) are of the same or differing parity, it follows that either \( s_1 + s_2 \) or \( s_1 - s_2 \) is congruent to \( 2^m \) modulo \( 2^{m+1} \). This establishes (i). The proof of (ii) is similar but a little more elaborate. If there is an \( e \in (\mathbb{Z}/(2^m\mathbb{Z}))^* \) such that \( e_j \equiv \pm e \pmod{2^m} \) for every \( j \), then, regardless of the choice of \( \delta_j \), we have \( \sum_{j=1}^t \delta_j e_j \equiv \pm u \pmod{2^m} \) where \( |u| \leq 2^{m-1} - 1 \). Thus, \( \frac{2^m-1}{2} \not\in S \) and we can assume this henceforward. As before, we argue by induction on \( m \). When \( m = 1 \), we have \( t = 0 \) and \( S \) is empty so the conclusion is trivial. When \( m = 2 \), we have \( t = 1 \) and \( 2^{m-1} = 2 \), so \( e_1 \neq 0 \) or 2 \( \pmod{4} \) and we are done. Now suppose that the conclusion holds for a given value of \( m \geq 2 \) and consider the case with \( m \) replaced by \( m + 1 \). That is, we suppose that \( \frac{2^m}{2} \) is not contained in \( S \) and will deduce that there is an \( e \in (\mathbb{Z}/(2^{m+1}\mathbb{Z}))^* \) such that each \( e_j \) satisfies \( e_j \equiv \pm e \pmod{2^{m+1}} \). We now form the partial sums

\[
s_1 := \sum_{j=1}^{2^{m-1}-1} \delta_j e_j
\]

and

\[
s_2 := \sum_{j=2^{m-1}}^{2^m-1} \delta_j e_j.
\]
By (i) and the inductive hypothesis, if there is no \( e \) such that \( e_j \equiv \pm e \pmod{2^m} \) for \( 1 \leq j \leq 2^{m-1} - 1 \), where \( e \in (\mathbb{Z}/(2^m)\mathbb{Z})^* \), then there is a choice of the \( \delta_j \) such that

\[
s_2 \equiv 2^{m-1} \pmod{2^m}
\]

and

\[
s_1 \equiv 2^{m-1} \pmod{2^m}.
\]

Thus if there is no such \( e \), then as before one of \( s_1 \pm s_2 \equiv 2^m \pmod{2^{m+1}} \), which we have expressly excluded. Thus, there is such an \( e \). Moreover, we can repeat the argument with every permutation of the \( e_j \). Thus, we can conclude that there is an \( e \) such that \( e_j \equiv \pm e \pmod{2^m} \) for \( 1 \leq j \leq 2^{m-1} - 1 \), where \( e \in (\mathbb{Z}/(2^m)\mathbb{Z})^* \). In other words, either

\[
e_j \equiv \pm e
\]

or

\[
\pm(e + 2^m) \pmod{2^{m+1}}.
\]

Now, we may conclude that either all the \( e_j \) are congruent to \( \pm e \) or they are congruent to \( \pm(e + 2^m) \), because if, say, \( e_1 \equiv \pm e \pmod{2^{m+1}} \) and \( e_2 \equiv \pm(e + 2^m) \pmod{2^{m+1}} \), then either \( e_1 + e_2 \) or \( e_1 - e_2 \) is \( 2^m \pmod{2^{m+1}} \), contradicting \( 2^m \not\in S \).

A weaker version of the lemma, in which one replaces the exact lower bound \( 2^{m-1} \) of \( t \) in part (i) by the crude bound \((2^m - 1)(2^{m-1} - 1) + 1 \) would follow by a direct application of the pigeonhole principle. An extension of this lemma to general modulus (not necessarily a power of 2) could be formulated and then proved by Kneser’s theorem (see [4, Chapter 1]), which is, however, not of direct relevance to the purpose of this memoir. Nevertheless, it would be of essence if one desires to establish the lower order terms for the asymptotics in Theorem 1.1.

Having established the necessary preliminaries, we are poised to reveal the structure of \( \mathcal{E}_a^* \) when \( a = p^r \) is a power of an odd prime.

**Lemma 2.5.** Let \( \mathcal{P} \) be the sequence of prime factors of \( n \), counted with multiplicity, and let \( \mathcal{T} \) be the subsequence of prime \( r \in \mathcal{P} \) with \( r \not\in H \). Then denote by \( t \) the length of \( \mathcal{T} \). Considering the projection map: \( \mathbb{Z} \to \mathbb{Z}/a\mathbb{Z} \), suppose the image of the sequence \( \mathcal{P} \) contains \( H \).

(i) If \( t \geq 2^{m-1} \), then \( n \not\in \mathcal{E}_a^* \).

(ii) If \( t = 2^{m-1} - 1 \), then \( n \in \mathcal{E}_a^* \) if and only if every prime factor in \( \mathcal{T} \) is congruent to \( g^{e'} \) modulo \( p^r \) for a fixed primitive root \( g \pmod{p^r} \), and for some \( e' \) such that \( e' \equiv \pm e \pmod{2^m} \) with \( e \) being a fixed odd number.

**Proof.** Recall that \( G \) and \( H \) are given by (2) and (3), respectively, for a fixed primitive root \( g \) modulo \( a \). Write \( T = \{r_j\}_1^t \). Let the sequence \( \{e_j\} \) be such that \( g^{e_j} \equiv r_j \pmod{a} \). By the assumption \( r_j \not\in H \), we know \( e_j \not\equiv 0 \pmod{2^m} \), for \( 1 \leq j \leq t \). Let \( \mathcal{G} = \mathbb{Z}/2^m\mathbb{Z} \). Now \( \{e_j\} \) can be viewed as a sequence of non-zero elements in \( \mathcal{G} \). Clearly, we see that Lemma 2.4 comes into play here. More precisely, when \( t \geq 2^{m-1} \), there exist \( \delta_j \in \{-1, 0, 1\} \) such that

\[
\sum_{j=1}^t \delta_j e_j \equiv 2^{m-1} \pmod{2^m}.
\]

This is equivalent to

\[
\sum_{j=1}^t \delta_j e_j \equiv b2^{m-1} \pmod{2^m d},
\]
for some odd number $b$ such that $1 \leq b \leq d$. Hence,

$$
\sum_{j=1}^{t} \delta_j c_j + (d - b)2^{m-1} \equiv 2^{m-1}d \pmod{2^m d}.
$$

Translating this using multiplicative language, we know that

$$
g^{(d-b)/2} \cdot 2^m \prod_{j=1}^{t} (g'v)^{\delta_j} \equiv g^{2m/d} \equiv -1 \pmod{a}.
$$

By assumption, there exists $q \in \mathcal{P}$ such that $q \equiv g^{(d-b)/2} \cdot 2^m \pmod{a}$. On the other hand, $g'^j \equiv r_j \pmod{a}$ and $q \prod_{j=1}^{s} r_j | n$. Hence, there exist two coprime divisors $u$ and $v$ of $n$, such that $u/v \equiv -1 \pmod{a}$ namely $u + v \equiv 0 \pmod{a}$. By Lemma 2.2, we know $n \notin \mathcal{E}_a$. This proves part (i).

For part (ii), the necessity of the condition follows by exactly the same argument as above, keeping in mind that Lemma 2.4 still plays an important role. Now, in order to prove the sufficiency, we just need to reverse the above argument and argue by contradiction. (Note that the condition $\mathcal{P}$ contains $H$ is not needed in this direction.)

Our next task naturally is to extend Lemma 2.5 to general modulus. We will see how one can carry the arguments here to the general case only with some mild difficulties in the next subsection.

2.3. The case for general $a$

Now we treat the general case $a = 2^{\gamma_0} p_1^{\gamma_1} p_2^{\gamma_2} \ldots p_k^{\gamma_k}$. Of course by the Chinese remainder theorem we have the group isomorphism

$$(\mathbb{Z}/a\mathbb{Z})^\ast \simeq (\mathbb{Z}/2^{\gamma_0}\mathbb{Z})^\ast \times (\mathbb{Z}/p_1^{\gamma_1}\mathbb{Z})^\ast \times \ldots \times (\mathbb{Z}/p_k^{\gamma_k}\mathbb{Z})^\ast.$$

As before, we still denote this group by $G$. Here, all the groups $(\mathbb{Z}/p^\ast)$ are cyclic when $p$ is an odd prime, but in general $(\mathbb{Z}/2^{\gamma_0}\mathbb{Z})^\ast$ is not except when $\gamma_0 \leq 2$. For instance, $(\mathbb{Z}/2\mathbb{Z})^\ast$ is trivial and $(\mathbb{Z}/4\mathbb{Z})^\ast$ is isomorphic to $(\mathbb{Z}/2\mathbb{Z}, +)$. In particular, there is no difference between the cases $\gamma_0 = 0$ and $1$ because they exert no influence on $G$. While, when $\gamma_0 \geq 3$, $(\mathbb{Z}/2^{\gamma_0}\mathbb{Z})^\ast$ is a product of two cyclic groups with generators $-1$ (mod $2^{\gamma_0}$) and $5$ (mod $2^{\gamma_0}$), respectively, namely

$$(\mathbb{Z}/2^{\gamma_0}\mathbb{Z})^\ast \simeq \langle -1 \rangle \times \langle 5 \rangle.$$

Apparently, $|\langle -1 \rangle| = 2$ and $|\langle 5 \rangle| = 2^{\gamma_0-2}$.

Here, we still want to find a maximal subgroup $H$ of $G$ such that $-1$ (mod $a$) $\notin H$. However, the issue here is that such subgroups of $G$ might not be unique. They can be easily constructed as follows. Let $G_i = (\mathbb{Z}/p_i^{\gamma_i}\mathbb{Z})^\ast$, for $0 \leq i \leq k$ and let $H_i$ be the maximal subgroup of $G_i$ such that $-1$ (mod $p_i^{\gamma_i}$) $\notin H_i$. As we remarked before, $H_1, H_2, \ldots, H_k$ are unique but $H_0$ is not in general. In fact, $H_0$ is trivial if $\gamma_0 \leq 2$ and is one of the two subgroups of index 2 in the ambient group $G_0$ if $\gamma_0 \geq 3$. Recall our discussion in the cyclic case, hence $[G_i : H_i] = 2^{m_i}$ for some positive integer $m_i$ and for all $1 \leq i \leq k$. Moreover,

$$[G_0 : H_0] = \begin{cases} 1 & \text{when } \gamma_0 \leq 1, \\ 2 & \text{when } \gamma_0 \geq 2. \end{cases}$$

Now choose $m$ such that

$$m = \begin{cases} \min_{1 \leq i \leq k} m_i & \text{when } \gamma_0 \leq 1, \\ 1 & \text{when } \gamma_0 \geq 2. \end{cases}$$
namely

\[ 2^m | \gcd(\delta, p_1 - 1, p_2 - 1, \ldots, p_k - 1), \]

where

\[ \delta = \begin{cases} 
0 & \text{when } \gamma_0 \leq 1, \\
2 & \text{when } \gamma_0 \geq 2.
\end{cases} \]

By definition, we have \( m \geq 1 \). The subgroup \( H \) as described above is one of the following groups with index \( [G : H] = 2^m \):

\[ H_0 \times G_1 \times \ldots \times G_k, G_0 \times H_1 \times \ldots \times G_k, \ldots \]

in which we just replace the \( i \)th component of \( G \) by \( H_i \) for \( 0 \leq i \leq k \). We write \( \phi(a) = 2^m d \) and hence \( |H| = \phi(a)/2^m = d \). Note that \( d \) is not necessarily odd in general.

It is routine to prove the following lemma (see the proof of Lemma 2.3).

**Lemma 2.6.** We have the following inclusion relation of sets:

\[ \{ n \in \mathbb{N} : p | n, p \in \mathbb{P} \Rightarrow p \in H \} \subseteq \mathcal{E}_a^*. \]

The next lemma is crucial for our arguments.

**Lemma 2.7.** Let \( H' \) be a subset of \( G \) with cardinality \( |H'| \geq d \). And suppose for each \( h' \in H' \), there are at least \( \phi(a) \) many (counted with multiplicity) prime factors \( q \) of \( n \) satisfying \( q \equiv h' \pmod{a} \). Then \( n \notin \mathcal{E}_a^* \) unless \( H' = H \) for some subgroup \( H \) defined above.

**Proof.** The proof still relies on Lemma 2.2. Actually, by Lemma 2.2, if we can find a divisor of \( n \) that is congruent to \(-1 \) (mod \( a \)), then \( n \notin \mathcal{E}_a^* \). Now, our argument goes roughly as follows: the fact that \( n \) has sufficiently many prime factors lying in sufficiently many different reduced residue classes in \( G \), forces \( n \) to have at least one divisor lying in the residue class \(-1 \) (mod \( a \)) unless \( H' \) is one of the above subgroups of \( G \). To make this statement rigorous, let \( H'' \) be the set of all the residue classes of divisors of \( n \) in \( G \). Then \( \mathcal{I} \in H'' \) and for any two elements \( h_1, h_2 \in H' \), we have \( h_1^{-1} = h_1^{-\phi(a)-1} \in H'' \) and \( h_1 h_2 \in H'' \). This means that \( H'' \) contains the subgroup \( (H') \) generated by the elements in \( H' \), and in particular, this subgroup has cardinality at least \( |H'| \geq d \). However, our \( H \) is maximized such that \( -\mathcal{I} \notin H \), which implies either that \( -\mathcal{I} \in (H') \) and hence \( -\mathcal{I} \notin H'' \), or that \( H' \) itself is a maximal subgroup such that \( -\mathcal{T} \notin H' \). In the former case, we have \( n \notin \mathcal{E}_a^* \) by Lemma 2.2 and in the latter case, we know by Lemma 2.6 that \( n \notin \mathcal{E}_a^* \).

Now we need an analogue of Lemma 2.5 for the general case. Here, we need to pay special attention to the power of 2 dividing \( a \). When \( \gamma_0 \geq 2 \), we have \( m = 1 \), which is the ‘worst’ case in the sense that then \( E_a(N) \) has the maximal order of magnitude.

**Lemma 2.8.** Suppose that \( H = G_0 \times G_1 \times \ldots \times H_i \times \ldots \times G_k \) is a subgroup of \( G \) defined as above. Let \( P \) be the sequence of prime factors of \( n \) (counted with multiplicity). And let \( T \) be the subsequence of prime \( r \) in \( P \) with \( r \notin H \). Then denote by \( t \) the length of \( T \). Considering the projection map: \( \mathbb{Z} \rightarrow \mathbb{Z}/a\mathbb{Z} \), suppose the image of the sequence \( P \) contains \( H \).

(i) If \( t \geq 2^{m-1} \), then \( n \notin \mathcal{E}_a^* \).
(ii) If \( t = 2^{m-1} - 1 \) and \( m \geq 2 \) (in this case, \( \gamma_0 \leq 1 \) and hence \( G_0 \) is trivial and, in particular, our \( H_i \) here cannot be \( H_0 \)), then \( n \in \mathcal{E}_a^* \) if and only if every prime factor in \( T \) is congruent
to \( g^{e'} \mod p_i^{\gamma_i} \) for a fixed primitive root \( g \mod p_i^{\gamma_i} \), and for some \( e' \) such that \( e' \equiv \pm e \mod 2^{m} \) with \( e \) being a fixed odd number.

Proof. Generally speaking, the arguments in the proof of Lemma 2.5 still work here. Nevertheless, one needs to make some changes accordingly. It is trivial to verify the conclusions when \( m = 1 \). So, without loss of generality, we assume \( m \geq 2 \) hence \( 1 \leq i \leq k \).

Let \( T = \{ r_j \} \) and fix a primitive root \( g \mod p_i^{\gamma_i} \). Let the sequence \( \{ e_j \} \) be such that \( g^{e_j} \equiv r_j \mod p_i^{\gamma_i} \). By the assumption \( r_j \mod a \notin H \), namely, \( r_j \mod p_i^{\gamma_i} \notin H \) we know \( e_j \not\equiv 0 \mod 2^m \), for \( 1 \leq j \leq t \). Let \( G = \mathbb{Z}/2^m\mathbb{Z} \). Now, \( \{ e_j \} \) can be viewed as a sequence of non-zero elements in \( G \). Hence, by Lemma 2.4, when \( t \geq 2^{m-1} \), there exist \( \delta_j \in \{ -1, 0, 1 \} \) such that

\[
\sum_{j=1}^{t} \delta_j e_j \equiv 2^{m-1} \mod 2^m.
\]

After writing \( \phi(p_i^{\gamma_i}) = 2^m d_i \) with \( d_i \) odd, this is equivalent to

\[
\sum_{j=1}^{t} \delta_j e_j \equiv b2^{m-1} \mod 2^m d_i,
\]

for some odd number \( b \) such that \( 1 \leq b \leq d_i \). Hence,

\[
\sum_{j=1}^{t} \delta_j e_j + (d_i - b)2^{m-1} \equiv 2^{m-1}d_i \equiv \mod 2^m d_i.
\]

Translating this using multiplicative language, we obtain

\[
g^{(d_i - b)/2} \cdot 2^m \prod_{j=1}^{t} (g^{e_j})^{\delta_j} \equiv g^{2^{m-1}d_i} \equiv -1 \mod p_i^{\gamma_i}.
\]

By assumption, there exists \( q \in \mathcal{P} \) such that

\[
\begin{cases}
q \equiv - \prod_{j=1}^{t} r_j^{-\delta_j} \mod p_i^{\gamma_i} & (1 \leq l \leq k, l \neq i), \\
q \equiv g^{(d_i - b)/2} \cdot 2^m \mod p_i^{\gamma_i}
\end{cases}
\]

simultaneously. Hence, by the Chinese remainder theorem, we obtain

\[
q \prod_{j=1}^{t} r_j^{-\delta_j} \equiv -1 \mod a,
\]

so that there exist two coprime divisors \( u \) and \( v \) of \( n \), such that \( u/v \equiv -1 \mod a \), namely, \( u + v \equiv 0 \mod a \). Again by Lemma 2.2, we know \( n \notin \mathcal{E}_a^* \).

Part (ii) can be proved similarly (see the comment in the proof of Lemma 2.5).

3. The analytic inputs

We need the following generalization of Ikehara’s Tauberian theorem, which is due to Delange (see [2], see also Tenenbaum [9, Theorem 7.15]). This extends Ikehara’s theorem to the case of a singularity of mixed type, involving algebraic and logarithmic poles. As usual, we use \( \sigma \) to denote the real part of the complex number \( s \), and we define \( l(s) = \log(1/(s - 1)) \) for \( \sigma > 1 \) by taking \( l(2) = 0 \) and then defining \( l(s) \) by continuous variation along the line segment joining 2 to \( s \).
Lemma 3.1 (Delange, 1954). Let $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ be a Dirichlet series with non-negative coefficients, converging for $\sigma > 1$. Suppose that $f(s)$ is holomorphic at all points of the line $\sigma = 1$ other than $s = 1$ and that, in the neighbourhood of this point and for $\sigma > 1$, we have

$$f(s) = (s-1)^{-\omega-1} \sum_{j=0}^{t} g_j(s) \left( \log \left( \frac{1}{s-1} \right) \right)^j + g(s),$$

where $\omega$ is some real number, and the $g_j(s)$ and $g(s)$ are functions holomorphic at $s = 1$, the number $g_1(1)$ being non-zero.

(i) If $\omega$ is not a negative integer, then we have, as $x \to \infty$,

$$\sum_{n \leq x} a_n \sim \frac{g_1(1)}{\Gamma(\omega + 1)} x (\log x)^\omega (\log \log x)^t.$$

(ii) If $\omega = -m - 1$ for a non-negative integer $m$ and if $t \geq 1$, then we have, as $x \to \infty$,

$$\sum_{n \leq x} a_n \sim (-1)^m m! g_1(1) x (\log x)^\omega (\log \log x)^{t-1}.$$

The following lemma is the key analytic ingredient of this paper. Essentially it plays the role of a sieve, but the upshot is that it produces asymptotics, not just an upper bound as almost all sieves do.

Lemma 3.2. Suppose that $a$ is a positive integer; let $B = \{b_1, \ldots, b_w\}$ be a subset of $(\mathbb{Z}/a\mathbb{Z})^*$ with $w \geq 0$ elements and let $C = \{c_j\}_1^t$ be a sequence of length $t$ with elements in $(\mathbb{Z}/a\mathbb{Z})^*$ (elements could be repeated). And suppose further that $B$ and $C$ do not share common elements. Now let $\mathbb{P}$ denote the set of primes and define

$$A = A(B, C) = \{ q_1 q_2 \ldots q_l r_1 r_2 \ldots r_t : q_i \in \mathbb{P}, r_j \in \mathbb{P}, q_i \in B, r_j = c_j, l \geq 0 \}. $$

(i) If $w \geq 1$, then we have, as $x \to \infty$,

$$|A(x)| \sim C(a, B, t) \frac{x (\log_\log x)^t}{(\log x)^{1-w/\phi(a)}}.$$  

(ii) If $w = 0$ and $t \geq 1$, then we have, as $x \to \infty$,

$$|A(x)| \sim C(a, t) \frac{x (\log x)^{t-1}}{\log x}.$$  

The constants $C(a, B, t)$ and $C(a, t)$ are positive and do not depend on the choices of the $c_j$.

Proof. Let

$$a_n = \begin{cases} 
1, & n \in A, \\
0, & n \notin A. 
\end{cases}$$

The set $A(B, C)$ has a multiplicative structure, and this leads naturally to the following Dirichlet series:

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} = \prod_{q \in \mathbb{P}} \left( 1 - \frac{1}{q^s} \right)^{-1} \prod_{j=1}^{t} \sum_{r \in \mathbb{P}} \frac{1}{r^s}, \quad (4)$$

which converges absolutely and locally uniformly in the region $\sigma > 1$. 

When $D(s)$ is a Dirichlet series that converges absolutely and locally uniformly for $\sigma > \sigma_0$, has an analytic continuation for $\sigma > \sigma_1$, is non-zero for $\sigma > \sigma_2$ and satisfies $\lim_{\sigma \to \infty} D(\sigma) = 1$, we define $D(\sigma)^\alpha$ for $\sigma > \max(\sigma_1, \sigma_2)$ and an arbitrary complex number $\alpha$ by $\exp(\alpha \log D(\sigma))$, where we choose the principal value of $\log D(\sigma)$ for some suitably large $\sigma_3$ and then define $\log D(\sigma)$ by continuous variation along the line segment from $\sigma_3$ to $\sigma$.

Let $e(\chi) = (1/\phi(a)) \sum_{\chi \mod a} \chi(q) \overline{\chi(b)}$. Then, by the orthogonality of Dirichlet characters, the product over $q$ on the right-hand side of (4) is

$$\prod_{b \in B} \prod_{q} (1 - 1/q^s)^{-e(\chi)} = \prod_{b \in B} \prod_{\chi \mod a} (L(s, \chi) g_1(s, \chi))^{\chi(b)/\phi(a)},$$

where

$$g_1(s, \chi) = \prod_{q} \frac{(1 - \chi(q)/q^s)}{(1 - 1/q^s)\chi(q)},$$

which converges absolutely when $\sigma > \frac{1}{2}$, and hence has no zeros in that region. Thus, $g_1(s, \chi)^{\chi(b)/\phi(a)}$ is a well-defined analytic function when $\sigma > \frac{1}{2}$.

Now the above product can be further rearranged as

$$L(s, \chi_0)^{\omega/\phi(a)} g_1(s), \quad (5)$$

where $\omega$ is the cardinality of $B$, $\chi_0$ is the principal character modulo $a$ and

$$g_1(s) = \prod_{b \in B} \prod_{\chi \neq \chi_0 \mod a} (L(s, \chi) g_1(s, \chi))^{\chi(b)/\phi(a)}.$$

In particular,

$$g_1(1) \neq 0.$$

Note that $g_1(1)$ may depend on the choice of $B$.

It is well known that $L(s, \chi)$ has no zeros with $\sigma \geq 1$ and has an analytic continuation to the whole complex plane. Moreover, when $\chi$ is non-principal, it is entire and when $\chi$ is a principal character $\chi_0$, it has a simple pole at $s = 1$ and $(s - 1)L(s, \chi_0)$ is entire. Thus, when $\chi$ is non-principal,

$$L(s, \chi)^{\chi(b)/\phi(a)}$$

is analytic in the region $\sigma \geq 1$ and hence so is $g_1(s)$.

On the other hand, again by the orthogonality of Dirichlet characters the sum over $r$ on the right-hand side of (4) is

$$\frac{1}{\phi(a)} \sum_{\chi \mod a} \chi(-c_j) \sum_{p} \frac{\chi(p)}{p^s}.$$

Now, it is readily verified that, when $\sigma > 1$, we have

$$\log L(s, \chi) = - \sum_{p} \log \left(1 - \frac{\chi(p)}{p^s}\right) = \sum_{p} \frac{\chi(p)}{p^s} + \sum_{p} \sum_{k=2}^{\infty} \frac{\chi(p^k)}{kp^{ks}}.$$

The second sum on the right-hand side converges locally uniformly when $\sigma > \frac{1}{2}$. Thus,

$$\sum_{p} \frac{\chi(p)}{p^s} = \log L(s, \chi) + h(s, \chi),$$
where \( h(s, \chi) \) is holomorphic for \( \sigma > \frac{1}{2} \). Note that \( \log L(s, \chi) \) is analytic on the line \( \sigma = 1 \) except when \( \chi = \chi_0 \), when it has a logarithmic singularity at the point \( s = 1 \). Hence,

\[
\sum_{p \neq \tau} \frac{1}{p^s} = \frac{1}{\phi(a)} \log L(s, \chi_0) + h(s, c_j),
\]

where \( h(s, c_j) \) is an analytic function of \( s \) for \( \sigma \geq 1 \). Therefore,

\[
\prod_{j=1}^{t} \sum_{p \neq \tau} \frac{1}{p^s} = \frac{1}{\phi(a)^t} \sum_{j=0}^{t} (\log L(s, \chi_0))^j h_j(s),
\]

where the \( h_j(s) \) are analytic when \( \sigma \geq 1 \) and \( h_t(1) = 1 \).

Now, on combining (4)–(6), we have

\[
f(s) = \frac{g_1(s)}{\phi(a)^t} L(s, \chi_0)^{w/\phi(a)} \sum_{j=0}^{t} (\log L(s, \chi_0))^j h_j(s).
\]

We have \( L(s, \chi_0) = \zeta(s) \prod_p (1 - p^{-s}) \) and the Riemann zeta function \( \zeta(s) \) has a simple pole at \( s = 1 \) with residue 1 at \( s = 1 \). Thus,

\[
L(s, \chi_0) = \frac{\phi(a) g_2(s)}{a(s-1)},
\]

where \( g_2(s) \) is an entire function with \( g_2(1) = 1 \). On plugging this into the above expression for \( f(s) \), the asymptotic formulæ of Lemma 3.2 follow from Lemma 3.1. Note that we apply part (i) of Lemma 3.1 when \( w \geq 1 \) and part (ii) when \( w = 0 \) and \( t \geq 1 \). That the constants \( C(a, B, t) \) and \( C(a, t) \) are positive follows by observing first that, by Lemma 3.1, they are non-zero and then that the left-hand side of the asymptotic formula is non-negative.

4. Proof of Theorem 1.1

The main analytic tool in the proof of Theorem 1.1 is Lemma 3.2 and we will apply it to the various sets from Subsection 2.3. Recall the definitions of the groups \( G \) and \( H \) and of the numbers \( m \) and \( d \) from Subsection 2.3. We denote by \( \mathcal{H} \) the set of all subgroups \( H \) defined in Subsection 2.3 for general \( a \). Now, as was defined in Lemma 3.2, we form the set

\[
\mathcal{A}(H, C),
\]

where \( H = G_0 \times G_1 \times \ldots \times H_i \times \ldots \times G_k \in \mathcal{H} \),

and \( C = \{c_j\}_1^t \) is a sequence of length \( t = 2^{m-1} - 1 \) with elements in \( G \). Moreover, for a fixed primitive root \( g \) (mod \( p \)) and a fixed odd number \( e \), we have \( c_j \equiv g^{e'} \) (mod \( p \)) for some \( e' \) with \( e' \equiv \pm e \) (mod \( 2^m \)). For a fixed \( H \in \mathcal{H} \), there are only finitely many possibilities (\( d^2 2^{t+m-1} \) actually) for \( C \). Lemma 3.2 immediately implies the following lemmas:

**Lemma 4.1.**

\[
|\mathcal{A}(H, C)(N)| \sim C(H, C) \frac{N (\log \log N)^{2^{m-1}-1}}{(\log N)^{1-1/2^m}}.
\]

Now, we need to show that the intersection of any two distinct such sets, \( \mathcal{A}(H^1, C^1) \) and \( \mathcal{A}(H^2, C^2) \), is a relatively small set.
Lemma 4.2. We have

\[
|\mathcal{A}(H^1, C^1) \cap \mathcal{A}(H^2, C^2)(N)| \ll_a \frac{N(\log \log N)^{2m-2}}{(\log N)^{1-1/4m}}.
\]

Proof. If \(H^1\) and \(H^2\) are the same, then \(C^1\) and \(C^2\) differ in at least one element. Hence, the intersection is empty. So without loss of generality, we can assume that \(H^1\) and \(H^2\) are not the same. Then \(H^1 \cap H^2\) is a subgroup of \(G\) with index \(4^m\). Also note the relation

\[
\mathcal{A}(H^1, C^1) \cap \mathcal{A}(H^2, C^2) \subseteq \mathcal{A}(H^1 \cap H^2, C^1 \cup C^2),
\]

where \(C^1 \cup C^2\) is the union of the sequences \(C^1\) and \(C^2\) and hence is of length \(2^m - 2\). Then the desired conclusion follows from Lemma 3.2.

Now set

\[
U = \bigcup_{H \in \mathfrak{H}} \bigcup_{C} \mathcal{A}(H, C),
\]

where the union runs through all \(H \in \mathfrak{H}\) and the corresponding sequences \(C\) for \(H\) as defined above. We know that \(U \subseteq E^*_a\) from Lemma 2.8.

Lemma 4.3. We have

\[
E^*_a(N) - |U(N)| \ll_a \begin{cases} 
\frac{N(\log \log N)^{2m-1-2}}{(\log N)^{1-1/2m}} & \text{when } m \geq 2, \\
\frac{N(\log \log N)^{\phi(a)}\phi(a)}{(\log N)^{1-1/2m+1/\phi(a)}} & \text{when } m = 1.
\end{cases}
\]

Proof. We let \(\mathcal{W}(n)\) be the set of residue classes modulo \(a\) in which there are at least \(\phi(a)\) (counted with multiplicity) prime factors of \(n\). By Lemma 2.7, we know that

(i) if \(|\mathcal{W}(n)| \geq d + 1\), then \(n \notin E^*_a\);
(ii) if \(|\mathcal{W}(n)| = d\), then \(n \notin E^*_a\) unless \(\mathcal{W}(n) = H\) for some subgroup \(H\) of \(G\) as above.

Let

\[
\mathcal{N}(i) = \{ n \in E^*_a : |\mathcal{W}(n)| = i \},
\]

for \(0 \leq i \leq \phi(a)\). From the above discussion, we know \(\mathcal{N}(i)\) is empty as long as \(i > d\). Hence,

\[
E^*_a = \bigcup_{i=0}^{d} \mathcal{N}(i).
\]

Firstly observe that, by Lemma 3.2, we have

\[
\left| \left( \bigcup_{i=0}^{d-1} \mathcal{N}(i) \right)(N) \right| \ll_a \frac{N(\log \log N)^{\phi(a)}\phi(a)}{(\log N)^{1-1/2m+1/\phi(a)}}
\]

Now, if \(m = 1\), then we have \(\mathcal{N}(d) = U\) by part (i) of Lemma 2.8, and if \(m \geq 2\), then we have by Lemmas 2.7, 2.8 and 3.2 that

\[
|\left( \mathcal{N}(d) \right)(N)| - |U(N)| \ll_a \frac{N(\log \log N)^{2m-1-2}}{(\log N)^{1-1/2m}}.
\]

Therefore Lemma 4.3 follows by putting the above conclusions together.
Here we bound the error term rather crudely, following from Lemma 2.7. Actually, it can be refined substantially by a generalization of Lemma 2.4, which is, however, not pertinent to the purpose of the current paper.

Now Theorem 1.1 follows from Lemmas 2.1, 4.1, 4.2 and 4.3. It should be noted that the leading constant $C(a)$ appearing in Theorem 1.1 can be traced back explicitly in our arguments, but is inevitably messy, would require some non-trivial expenditure of effort and would not give any further insights into our problem.

References

1. E. Croot, D. Dobbs, J. Friedlander, A. Hetzel, and F. Pappalardi, ‘Binary Egyptian fractions’, *J. Number Theory* 84 (2000) 63–79.
2. H. Delange, ‘Généralisation du Théorème de Ikeda’, *Ann. Sci. École Norm. Sup.* (3) 71 (1954) 213–242.
3. C. Elsholtz, ‘Sums of $k$ unit fractions’, PhD Thesis, Technische Universität Darmstadt, Darmstadt, 1998.
4. H. Halberstam and K. F. Roth, *Sequences*, 2nd edn (Springer, New York, 1983).
5. G. Hofmeister and P. Stoll, ‘Note on Egyptian fractions’, *J. reine angew. Math.* 362 (1985) 141–145.
6. J.-J. Huang and R. C. Vaughan, ‘Mean value theorems for binary Egyptian fractions’, *J. Number Theory* 131 (2011) 1641–1656.
7. J.-J. Huang and R. C. Vaughan, ‘Mean value theorems for binary Egyptian fractions II’, *Acta Arith.* 155 (2012) 287–296.
8. Y. Rav, ‘On the representation of rational numbers as a sum of a fixed number of unit fractions’, *J. Reine Angew. Math.* 222 (1966) 207–213.
9. G. Tenenbaum, *Introduction to analytic and probabilistic number theory* (Cambridge University Press, Cambridge, 1995).

Jing-Jing Huang               Robert C. Vaughan  
Department of Mathematics      Department of Mathematics  
University of Toronto          McAllister Building  
Bahen Centre, Room 6103       Pennsylvania State University  
40 St. George St.              University Park, PA 16802  
Toronto, Ontario               USA  
Canada M5S 2E4                 rvaughan@math.psu.edu  

huang@math.toronto.edu