THE ERDÖS–MOSER EQUATION \(1^k + 2^k + \cdots +(m-1)^k = m^k\)
REVISITED USING CONTINUED FRACTIONS

YVES GALLOT, PIETER MOREE, AND WADIM ZUDILIN

Abstract. If the equation of the title has an integer solution with \(k \geq 2\), then \(m > 10^{9.9} \times 10^9\). This was the current best result and proved using a method due to L. Moser (1953). This approach cannot be improved to reach the benchmark \(m > 10^{10^7}\). Here we achieve \(m > 10^{10^6}\) by showing that \(2k/(2m-3)\) is a convergent of \(\log 2\) and making an extensive continued fraction digits calculation of \((\log 2)/N\), with \(N\) an appropriate integer. This method is very different from that of Moser. Indeed, our result seems to give one of very few instances where a large scale computation of a numerical constant has an application.

1. Introduction

In this paper we are interested in non-trivial integer solutions, that is, solutions with \(k \geq 2\), of the equation

\[
1^k + 2^k + \cdots + (m-2)^k + (m-1)^k = m^k.
\]

Conjecturally such solutions do not exist. For \(k = 1\), one has clearly the solution \(1+2 = 3\) (and no further ones). From now on we will assume that \(k \geq 2\). Moser [29] showed in 1953 that if \((m,k)\) is a solution of (1), then \(m > 10^{10^6}\) and \(k\) is even (for an easy reproof see Moree [28]). His result has since then been improved upon. Butske et al. [6] have shown by computing, rather than estimating, certain quantities in Moser’s original proof that \(m > 1.485 \cdot 10^{9.321}155\). By proceeding along these lines this bound cannot be improved on substantially. Butske et al. [6, p. 411] expressed the hope that new insights will eventually make it possible to reach the more natural benchmark \(10^{10^7}\).

Using that \(\Sigma_k(m) = 1^k + 2^k + \cdots + m^k \leq \int_1^m t^k dt\) and \(\Sigma_k(m+1) > \int_0^m t^k dt\) we obtain that \(k+1 < m < 2(k+1)\). This shows that the ratio \(k/m\) is bounded. By a more elaborate reasoning along these lines Krzysztofek [20] obtained that \(k+2 < m < \frac{3}{2}(k+1)\). This implies that \(k \geq 4\) and hence

\[
k + 2 < m < 2k.
\]

Received by the editor July 7, 2009 and, in revised form, April 9, 2010.
2010 Mathematics Subject Classification. Primary 11D61, 11Y65; Secondary 11A55, 11B83, 11K50, 11Y60, 41A60.

This research was carried out while the third author was visiting the Max Planck Institute for Mathematics (MPIM) and the Hausdorff Center for Mathematics (HCM) and was financially supported by these institutions. He and the second author thank the MPIM and HCM for providing such a nice research environment.
Dividing both sides of (1) by $m^k$ one sees that for every integer $m \geq 2$, (1) has precisely one real solution $k$. It is known that $\lim_{m \to \infty} k/m = \log 2$ and we show here that in fact the behaviour of $k$ as a function of $m$ can be determined in a much more explicit way (Theorem 1 and Section 2).

Moree et al. [27], using properties of the Bernoulli numbers and polynomials (an approach initiated in Urbanowicz [31]), showed that $N_1 = \text{lcm}(1, 2, \ldots, 200) | k$. Kellner [19] in 2002 showed that also all primes $200 < p < 1000$ have to divide $k$. Actually, Moree et al. [27, p. 814] proved a slightly stronger result and on combining this with Kellner’s, one obtains that $N_2 | k$ with

$$N_2 = 2^8 \cdot 3^5 \cdot 5^4 \cdot 7^3 \cdot 11^2 \cdot 13^2 \cdot 17^2 \cdot 19^2 \cdot \prod_{23 \leq p \leq 997} p > 5.7462 \cdot 10^{427}.$$ 

For some further references and information on the Erdős–Moser equation we refer to the book by Guy [14, D7].

In this paper we attack (1) using the theory of continued fractions. This approach was first explored in 1976 by Best and te Riele [3] in their attempt to solve the related conjecture of Erdős [11] that there are infinitely many pairs $(m, k)$ such that $\sum k (m) \geq m^k$ and $2(m - 1)^k < m^k$. In this context they also gave the following variant of one of their results (without proof), namely, (3) with $O(m^{-2})$ replaced with $o(m^{-1})$. The proof we give here uses the same circle of ideas as used by Best and te Riele. It seems that after their work, continued fractions in the Erdős–Moser context have been completely ignored. We hope the present paper makes clear that this is unjustified.

**Theorem 1.** For integer $m > 0$ and real $k > 0$ satisfying equation (1), we have the asymptotic expansion

$$k = \log 2 \left( m - \frac{3}{2} - \frac{c_1}{m} + O\left( \frac{1}{m^2} \right) \right) \quad \text{as } m \to \infty,$$

with $c_1 = \frac{25}{12} - 3 \log 2 \approx 0.00389 \ldots$. Moreover, if $m > 10^9$, then

$$\frac{k}{m} = \log 2 \left( 1 - \frac{3}{2m} - \frac{C_m}{m^2} \right), \quad \text{where} \quad 0 < C_m < 0.004.$$

**Corollary 1.** If $(m, k)$ is a solution of (1) with $k \geq 2$, then $2k/(2m - 3)$ is a convergent $p_j/q_j$ of $\log 2$ with $j$ even.

**Corollary 2.** The number of solutions $m \leq x$ of (1), as $x$ tends to infinity, is at most $O(\log x)$.

The equation (1) seems to be a sole example of an exponential Diophantine equation in just two unknowns for which even the finiteness of solutions is not yet established. The best result in this direction is given by Corollary 2 which is an immediate consequence of the exponential growth of $p_j$ as a function of $j$ and Corollary 1.

Corollary 1 is not the only result which relates convergents to solutions of Diophantine equations. For example, if $(x_0, y_0)$ is a positive solution to Pell’s equation $x^2 - dy^2 = \pm 1$, with $d$ a positive square-free integer, then $x_0/y_0$ is a convergent of the continued fraction expansion of $\sqrt{d}$. On the other hand, in our situation the number in question, $\log 2$, is transcendental and its continued fraction expansion is expected to be sufficiently “generic” (unlike that of quadratic irrationals).
Corollary 1 naturally leads us to investigate common factors of \( k \) and \( 2m - 3 \). This can be done using the method of Moser, but is not in the literature, as before there was no special reason for considering \( 2m - 3 \).

A key role in this arithmetic study is played by the congruence

\[
\sum_{j=1}^{l-1} j^r \equiv \begin{cases} 
0 \pmod{\frac{1}{2}} & \text{if } r > 1 \text{ is odd,} \\
\sum_{p|l, p|l-1} \frac{1}{p^r} \pmod{1} & \text{otherwise.}
\end{cases}
\]

This identity can be proved using the Von Staudt–Clausen theorem; for alternative proofs see, e.g., Carlitz [7] or Moree [25]. Its relevance for the study of (1) was first pointed out by Moree [26].

Given \( N \geq 1 \), put

\[
P(N) = \{ p : p - 1 \mid N \} \cup \{ p : 3 \text{ is a primitive root modulo } p \}.
\]

By a classical result of Hooley [16] it follows, assuming the Generalized Riemann Hypothesis (GRH), that \( P(N) \) has a natural density \( A \), with \( A = 0 \).

The main idea of this paper is, in essence, to make use of the fact that the convergents \( p_j/q_j \) of \( \log 2 \) have no reason to also satisfy \( N_2 \mid p_j \). The first piece of information comes from asymptotic analysis and the latter piece from arithmetic. Analysis and arithmetic give rise to conditions on the solutions that “do not feel each other”, and this is exploited in our main result:

**Theorem 2.** Let \( N \geq 1 \) be an arbitrary integer. Let

\[
\frac{\log 2}{2N} = [a_0, a_1, a_2, \ldots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots}}
\]

be the (regular) continued fraction of \( (\log 2)/(2N) \), with \( p_i/q_i = [a_0, a_1, \ldots, a_i] \) its \( i \)-th partial convergent.

Suppose that the integer pair \((m, k)\) with \( k \geq 2 \) satisfies (1) with \( N \mid k \). Let \( j = j(N) \) be the smallest integer such that:

1. \( j \) is even;
2. \( a_{j+1} \geq 180N - 2 \);
3. \( (q_j, 6) = 1 \); and
4. \( \nu_p(q_j) = \nu_p(3^{p-1} - 1) + \nu_p(N) + 1 \) for all primes \( p \in P(N) \) dividing \( q_j \).

Then \( m > q_j/2 \).

Computing many partial quotients (that is, continued fraction digits) of \( \log 2 \) is closely related to computing \( \log 2 \) with many digits of accuracy. Indeed, it is a well-known result of Lochs that for a generic number knowing it accurately
up to \( n \) decimal digits implies that we can compute about 0.97\( n \) (where 0.97 \( \approx \) \((\log 2)/(\log 10)/\pi^2\)) continued fraction digits accurately. For example, knowing 1000 decimal digits of \( \pi \) allows one to compute 968 continued fraction digits.

It seems a hopeless problem to prove anything about \( \mathbb{E}(\log q_j(N)) \), the expected value of \( \log q_j(N) \) produced by the result. However, metric theory of continued fractions offers some hope of proving a non-trivial lower bound for \( \mathbb{E}(\log q_j(N)(\xi)) \), where we require conditions \( (a) \), \( (b) \), \( (c) \) and \( (d) \) to be satisfied, but we replace \((\log 2)/(2N)\) by a “generic” \( \xi \in [0,1] \setminus \mathbb{Q} \). In this context recall the result of Lévy [21] that, for such a \( \xi \),

\[
\lim_{j \to \infty} \frac{\log q_j(\xi)}{j} = \frac{\pi^2}{12 \log 2} \approx 1.18.
\]

The Gauss–Kuz’min statistics assert that, for a generic \( \xi \), the probability that a given term in its continued fraction expansion is at least \( 1/\pi \) is \( 1/p \).

From Theorem 2. Being able to compute the convergents of \( (\log 2)/(2N) \) arbitrarily far, we would expect (taking \( N = N_2 \)) to show that \( m > 10^{10^{450}} \). With the current computer technology computing sufficiently many convergents is the bottleneck. Taking this into consideration we would expect to get

\[
m > 10^{0.515r},
\]

from Theorem 2 where \( r \) is the number of convergents we can compute accurately and 0.515 is the base 10 logarithm of Lévy’s constant (6). Note that the fact that \( N_2 \) has many divisors gives us some flexibility and increases the likelihood of the
heuristics to be applicable. Indeed, our numerical experimenting agrees well with our heuristic considerations (see Section 4). In early 2009, A. Yee and R. Chan 32 reached \( r > 31 \cdot 10^9 \) for \( \log 2 \). On the other hand, Y. Kanada and his team 18 had already computed \( \pi \) to over 1.24 trillion decimal digits in 2002, using formulae of the same complexity as those used for the computation of \( \log 2 \) (see [2, Chapter 3] for details). Thus, given the present computer (im)possibilities, one could hope to show (with a lot of effort!) that \( m > 10^{10^{12}} \).

Applying Theorem 2 with \( N = 2^8 \cdot 3^5 \cdot 5^3 \) or \( N = 2^8 \cdot 3^5 \cdot 5^4 \), and invoking the result of Moree et al. 27 that \( N \mid k \), we obtain the following.

**Theorem 3.** If an integer pair \((m, k)\) with \( k \geq 2 \) satisfies (1) then \( m > 2.7139 \cdot 10^{1667658416} \).

As an application we can show that \( \omega(m - 1) \geq 33 \), this improves on the result of Brenton and Vasiliu 5, who have shown that \( \omega(m - 1) \geq 26 \), where \( \omega \) denotes the number of distinct prime divisors; see Section 5.1 for further details.

The fact \( N_2 \mid k \) naively implies that \( k \) is of size \( 10^{427} \) (at least), which is much smaller than Moser’s \( 10^{10^{6}} \). However, in this paper we show that the fact actually yields that \( k > 10^{10^9} \) (and likely even \( k > 10^{10^{600}} \)) — a modestly small number dividing \( k \) leads to a huge lower bound for \( k \). Thus, on revisiting [27] after 16 years, its main result is seen to be far more powerful than the second author thought at that time.

In the three following sections we prove Theorems 1, 2 and 3, respectively. Finally, Section 5 is devoted to discussing some problems related to the Erdős–Moser equation.

2. Asymptotic dependence of \( k \) in terms of \( m \)

Our proof of Theorem 1 makes use of the following lemma.

**Lemma 1.** For any real \( k > 0 \), we have

\[
(1 - y)^k = e^{-ky} \left(1 - \frac{k}{2} y^2 \frac{k}{3} y^3 + \frac{k(k-2)}{8} y^4 + \frac{k(5k-6)}{30} y^5 + O(y^6)\right) \quad \text{as } y \to 0.
\]

Moreover, for \( k > 8 \) and \( 0 < y < 1 \), the following inequality holds:

\[
e^{-ky} \left(1 - \frac{k}{2} y^2 \frac{k}{3} y^3 + \frac{k(k-2)}{8} y^4 + \frac{k(5k-6)}{30} y^5 - \frac{k^3}{6} y^6\right) < (1 - y)^k < e^{-ky} \left(1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k-2)}{8} y^4 + \frac{k^2}{2} y^5\right).
\]

**Proof.** As for the asymptotic relation in (7), we simply develop the Taylor expansion of \((1 - y)^k e^{ky}\) up to \( y^5 \). Unfortunately, estimates coming from the classical forms for the remainder are not sufficient enough to derive a sharp dependence on \( k \) as in (8) for the last term. Therefore, we need more drastic methods to quantify the asymptotics in (7) when \( 0 < y < 1 \).

First, note that

\[
(1 - y)e^y = 1 - \sum_{n=2}^{\infty} \frac{n-1}{n!} y^n
= 1 - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{8} - \frac{y^5}{30} - \cdots, \quad 0 < y < 1.
\]
Since all coefficients, starting from \( n = 2 \), in this power series are negative and their sum is exactly \(-1\), for these values of \( y \) we have the inequality

\[
1 - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{8} - \frac{y^5}{30} - \frac{y^6}{120} < (1-y)e^y < 1 - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{8}.
\]

The quantities

\[
x_1 = \frac{y^2}{2} + \frac{y^3}{3} + \frac{y^4}{8}
\]

and

\[
x_2 = \frac{y^2}{2} + \frac{y^3}{3} + \frac{y^4}{8} + \frac{y^5}{30} + \frac{y^6}{120},
\]

which appear in (9), lie between 0 and 1 for \( 0 < y < 1 \).

Our next ingredient is Gerber's generalization of the Bernoulli inequality (see also Alzer [1]). It states that the remainder after \( k \) terms of the (possibly divergent) binomial series for \((1+x)^a\) (\(a, x \) real with \(-1 < x\)) has the same sign as the first neglected term. In particular, we have for real \( k > 2 \) and \( 0 < x < 1 \),

\[
(1-x)^k < 1 - kx + \frac{k(k-1)}{2} x^2,
\]

and for real \( k > 3 \) and \( 0 < x < 1 \),

\[
(1-x)^k > 1 - kx + \frac{k(k-1)}{2} x^2 - \frac{k(k-1)(k-2)}{6} x^3.
\]

Using the right inequality in (12) and taking \( x = x_1 \) in (11) we obtain, for \( k > 2 \),

\[
(1-y)^k e^{ky} < 1 - k\left(\frac{y^2}{2} + \frac{y^3}{3} + \frac{y^4}{8}\right) + \frac{k(k-1)}{2}\left(\frac{y^2}{2} + \frac{y^3}{3} + \frac{y^4}{8}\right)^2
\]

\[
= 1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k-2)}{8} y^4 + k(k-1)y^5\left(\frac{1}{6} + \frac{17}{144} y + \frac{1}{24} y^2 + \frac{1}{128} y^3\right)
\]

\[
< 1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k-2)}{8} y^4 + \frac{385}{1152} k(k-1)y^5,
\]

implying the upper estimate in (8). In the same vein, the application of the left identity in (9) and of (12) with \( x = x_2 \) results, for \( k > 3 \), in

\[
(1-y)^k e^{ky} > 1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k-2)}{8} y^4 + \frac{k(5k-6)}{30} y^5
\]

\[
- ky^6 \sum_{n=0}^{12} (a_n k^2 + b_n k + c_n)y^n,
\]

where the polynomials \( p_n(k) = a_n k^2 + b_n k + c_n \), \( n = 0, 1, \ldots, 12 \), all have positive leading coefficients \( a_n \); moreover, \( p_n(k) > 0 \) for \( k > 3 \) and \( n = 2, 3, \ldots, 12 \), \( p_1(k) = \frac{1}{24} k^2 - \frac{14}{60} k + \frac{17}{120} > 0 \) for \( k > 4 \), and \( p_0(k) = \frac{1}{48} k^2 - \frac{13}{72} k + \frac{121}{720} > 0 \) for \( k > 8 \). Using this positivity of the polynomials we can continue the inequality in (14) for \( k > 8 \).
as follows:

\[
(1 - y)^k e^{ky} > 1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k - 2)}{8} y^4 + \frac{k(5k - 6)}{30} y^5
\]
\[
- ky^6 \sum_{n=0}^{12} (a_n k^2 + b_n k + c_n)
\]
\[
= 1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k - 2)}{8} y^4 + \frac{k(5k - 6)}{30} y^5
\]
\[
- ky^6 \left( \frac{1}{6} k^2 - \frac{17}{24} k + \frac{11}{20} \right),
\]

(15)

from which we deduce the left inequality in (8), and the lemma follows. □

**Proof of Theorem 1.** The original equation (1) is equivalent to

\[
1 = m - 1 \sum_{j=1}^{m-1} \left( 1 - \frac{j}{m} \right)^k .
\]

(16)

Applying to each term on the right-hand side the inequality from (8) we obtain

\[
S_0 - \frac{k}{2m^2} S_2 - \frac{k}{3m^3} S_3 + \frac{k(k - 2)}{8m^4} S_4 + \frac{k(5k - 6)}{30m^5} S_5 - \frac{k^3}{6m^6} S_6
\]
\[
< \sum_{j=1}^{m-1} \left( 1 - \frac{j}{m} \right)^k < S_0 - \frac{k}{2m^2} S_2 - \frac{k}{3m^3} S_3 + \frac{k(k - 2)}{8m^4} S_4 + \frac{k^2}{2m^5} S_5,
\]

(17)

with the notation

\[
S_n = \sum_{j=1}^{m-1} j^n e^{-kj/m} = \sum_{j=1}^{m-1} j^n z^j \bigg|_{z=e^{-k/m}}.
\]

By (2) we have \( e^{-1} < z < e^{-1/2} \), where \( z = e^{-k/m} \), and hence \( 1/(1 - z) < 1/(1 - e^{-1/2}) < 3 \), and in the closed-form expression of the sum

\[
S_0 = \sum_{j=1}^{m-1} z^j = \frac{z}{1 - z} - \frac{z^m}{1 - z},
\]

the second term as well as its \( z \frac{d}{dz} \)-derivatives are bounded:

\[
0 < \left. \frac{z^m}{1 - z} \right|_{z=e^{-k/m}} < 3e^{-k} \quad \text{and}
\]
\[
0 < \left( \left. z \frac{d}{dz} \right|_{z=e^{-k/m}} \right)^n < 3^{n+1} m^n e^{-k}, \quad \text{for } n = 1, 2, \ldots
\]
Therefore, we can write the inequality in (17) as

\[
S'_0 - \frac{k}{2m^2} S'_2 - \frac{k}{3m^3} S'_3 + \frac{k(k-2)}{8m^4} S'_4 + \frac{k(5k-6)}{30m^5} S'_5 - \frac{k^3}{6m^6} S'_6 < \frac{500k^3e^{-k}}{m^3},
\]

implying

\[
S'_0 - \frac{k}{2m^2} S'_2 - \frac{k}{3m^3} S'_3 + \frac{k(k-2)}{8m^4} S'_4 + \frac{k(5k-6)}{30m^5} S'_5 - \frac{k^3}{6m^6} S'_6 < \frac{500k^3e^{-k}}{m^3},
\]

(18)

where

\[
S'_n = \sum_{j=1}^{\infty} j^n z^j = \left. \left( \frac{z d}{dz} \right)^n \frac{z}{1-z} \right|_{z=e^{-k/m}} = (-1)^n \left( \frac{z d}{dz} \right)^n \frac{1}{z-1} \right|_{z=e^{k/m}} \text{ for } n = 0, 1, \ldots;
\]

in particular,

\[
S'_0 = \frac{1}{z-1}, \quad S'_2 = \frac{z + z^2}{(z-1)^2}, \quad S'_3 = \frac{z + 4z^2 + z^3}{(z-1)^3}, \quad S'_4 = \frac{z + 11z^2 + 11z^3 + z^4}{(z-1)^4}, \quad S'_5 = \frac{z + 57z^2 + 302z^3 + 302z^4 + 57z^5 + z^6}{(z-1)^5}, \quad S'_6 = \frac{z + 26z^2 + 66z^3 + 26z^4 + z^5}{(z-1)^6},
\]

with \( z = e^{k/m} \). Since \( 500k^3e^{-k} < (2k)^{-3} < m^{-3} \) for \( k > m/2 > 30 \), using our equation (16) we can write the estimates (18) as

\[
1 - \frac{k(5k-6)}{30m^5} S'_5 - \frac{k^3}{6m^6} S'_6 < \frac{1}{m^3},
\]

(19)

Noting that \( e^{1/2} < z = e^{k/m} < e \), we find

\[
0 < S'_5 < \frac{e + 26e^2 + 66e^3 + 26e^4 + e^5}{(e^{1/2} - 1)^6} < 41438,
\]

\[
0 < S'_6 < \frac{e + 57e^2 + 302e^3 + 302e^4 + 57e^5 + e^6}{(e^{1/2} - 1)^7} < 658544.
\]
We continue (19) as follows:

\[
\left| 1 - \frac{1}{z} + \frac{k}{2m^2} \frac{z + z^2}{(z - 1)^3} + \frac{k}{3m^3} \frac{z + 4z^2 + z^3}{(z - 1)^4} - \frac{k(k - 2)}{8m^4} \frac{z + 11z^2 + 11z^3 + z^4}{(z - 1)^5} \right| < \frac{110000}{m^3},
\]

where \( z = e^{k/m} \).

We already know that \( k/m \) is bounded as \( m \to \infty \); making the ansatz \( k/m = c + O(1/m) \), hence \( z = e^{k/m} = e^c + O(1/m) \), we find from (20) that

\[
1 - \frac{1}{e^c - 1} = O\left( \frac{1}{m} \right) \quad \text{as} \quad m \to \infty,
\]

hence \( e^c = 2 \) and \( c = \log 2 \). Now we take

\[
\frac{k}{m} = \log 2 + \frac{a}{m} + \frac{b}{m^2} + O\left( \frac{1}{m^3} \right) \quad \text{as} \quad m \to \infty,
\]

hence

\[
z = e^{k/m} = 2 + \frac{2a}{m} + \frac{a^2 + 2b}{m^2} + O\left( \frac{1}{m^3} \right) \quad \text{as} \quad m \to \infty.
\]

Substituting these formulas into (20) results in

\[
O\left( \frac{1}{m^3} \right) = 1 - \frac{1}{1 + 2a/m + (a^2 + 2b)/m^2 + O(m^{-3})}
+ \frac{\log 2 + a/m + O(m^{-2})}{2m} \frac{6 + 10a/m + O(m^{-2})}{1 + 6a/m + O(m^{-2})}
+ \frac{\log 2 + O(m^{-1})}{3m^2} \frac{26 + O(m^{-1})}{1 + O(m^{-1})}
- \frac{\log^2 2 + O(m^{-1})}{8m^2} \frac{150 + O(m^{-1})}{1 + O(m^{-1})} + O\left( \frac{1}{m^3} \right)
= \frac{2a + 3 \log 2 - 3a^2 - 3a + 13a \log 2 - 2b + \frac{75}{4} \log^2 2 - \frac{26}{3} \log 2}{m}
+ O\left( \frac{1}{m^3} \right),
\]

hence \( a = -\frac{3}{2} \log 2, \ b = (3 \log 2 - \frac{25}{12}) \log 2 \) and, finally, we get the asymptotic formula (3).

To quantify this asymptotic expansion, we introduce the function

\[
F(x, \lambda) = \left( 1 - \frac{1}{z - 1} + \frac{x\lambda}{2} \frac{z + z^2}{(z - 1)^3} + \frac{x^2\lambda}{3} \frac{z + 4z^2 + z^3}{(z - 1)^4}
- \frac{x^2\lambda(\lambda - 2x)}{8} \frac{z + 11z^2 + 11z^3 + z^4}{(z - 1)^5} \right) \bigg|_{z = e^\lambda}.
\]

Then the inequality (20) takes the form

\[
\left| F\left( \frac{1}{m}, \frac{k}{m} \right) \right| < \frac{110000}{m^3}.
\]
Direct computation shows that
\[ F(x, \log 2(1 - \frac{3}{2}x)) > 0.005x^2 - 100x^3 \quad \text{and} \quad F(x, \log 2(1 - \frac{3}{2}x - 0.004x^2)) < -0.00015x^2 + 100x^3 \]
for \( x \leq 0.01 \); therefore,
\[ F\left(\frac{1}{m}, \log 2\left(1 - \frac{3}{2}m - \frac{0.004}{m^2}\right)\right) < -\frac{110000}{m^3} \quad \text{for} \quad m > 734 \cdot 10^6. \]

Comparing these estimates with (21) we conclude that, for \( k \) and \( m > 10^9 \) satisfying (16), we necessarily have
\[ \log 2\left(1 - \frac{3}{2}m - \frac{0.004}{m^2}\right) < \frac{k}{m} < \log 2\left(1 - \frac{3}{2m}\right), \]
which is the desired estimate (14).

Clearly, the strategy to deduce further terms in the expansion remains the same, but in order to achieve precision \( O(m^{-n}) \) for an integer \( n \geq 2 \), we have to use the Taylor expansion of \( (1 - y)^k e^{ky} \) up to \( y^{n+1} \) (each new term in (3) requires two extra terms in the expansion of \( (1 - y)^k e^{ky} \)). In this way we get
\[
k = cm - \frac{3}{2}c - \left(\frac{25}{12}c - 3c^2\right)m^{-1} + \left(-\frac{73}{8}c + \frac{61}{2}c^2 - 25c^3\right)m^{-2}
+ \left(-\frac{41299}{720}c + \frac{657}{2}c^2 - 598c^3 + \frac{1405}{4}c^4\right)m^{-3} + O(m^{-4})
\approx 0.69314718m - 1.03972077 - 0.00269758m^{-1} + 0.00323260m^{-2}
+ 0.00217182m^{-3} + O(m^{-4}),
\]
where \( c = \log 2 \). However, we do not possess any clear general strategy to quantify such expansions. Already proving a sharp dependence on \( k \) for the remainder of the \( n \)-th truncation of the Taylor expansion of \( (1 - y)^k e^{ky} \) (like we do for \( n = 4 \) in Lemma 1) seems to be a difficult task. We discuss related problems in Section 5.

**Proof of Corollary** 1. Let \((m, k)\) be a non-trivial integer solution of (1). By Moser’s result we know that \( m > 10^9 \). It follows from Theorem 1 that
\[
0 < \log 2 - \frac{2k}{2m - 3} < \frac{0.0111}{(2m - 3)^2}.
\]
By Legendre’s theorem, \(| \log 2 - p/q | < 1/(2q^2) \) implies that \( p/q \) is a convergent of \( \log 2 \), while \( \log 2 > p/q \) insures that the index of the convergent is even. Thus, \( 2k/(2m - 3) \) is a convergent \( p_j/q_j \) of the continued fraction of \( \log 2 \) with \( j \) even. 

### 3. The proof of the main theorem

In this section we prove Theorem 2. The restrictions on the prime factorization of \( q_j \) in that result are established using an argument in the style of Moser given in the proof of the following lemma.
Lemma 2. Let $(m,k)$ be a solution of (1) with $k \geq 2$. Let $p$ be a prime divisor of $2m - 3$. If $p - 1 \mid k$, then
\[ \nu_p(2m - 3) = \nu_p(3^{p-1} - 1) + \nu_p(k) + 1 \geq 2. \]

If 3 is a primitive root modulo $p$, then $p - 1 \mid k$.

Proof. Using that $k$ must be even, we find that
\[
\sum_{j=1}^{2m-4} j^k \equiv \sum_{j=1}^{m-1} j^k + \sum_{j=1}^{m-3} (2m - 3 - j)^k \equiv \sum_{j=1}^{m-1} j^k + \sum_{j=1}^{m-3} j^k \pmod{2m - 3}
\]
\[
\equiv m^k + m^k - (m - 1)^k - (m - 2)^k \equiv 2(3^k - 1)(m - 1)^k \pmod{2m - 3},
\]
where we used that $m^k \equiv (2m - 3 + m)^k \equiv 3^k(m - 1)^k \pmod{2m - 3}$ and $(m - 2)^k \equiv (2m - 3 - m + 1)^k \equiv (m - 1)^k \pmod{2m - 3}.$ On applying (3) with $l = 2m - 3$ and $r = k$, we then obtain that
\[
(24) \quad \frac{2(3^k - 1)(m - 1)^k}{2m - 3} \equiv - \sum_{p \mid 2m - 3 \atop p - 1 \mid k} \frac{1}{p} \pmod{1}.
\]

If $p \mid 2m - 3$ and $p - 1 \mid k$, the $p$-order of the right-hand side is $-1$. The $p$-order of the left-hand side must also be $-1$, that is, we must have
\[ \nu_p(2m - 3) = \nu_p(3^k - 1) + kv_p(m - 1) + 1 = \nu_p(3^{p-1} - 1) + \nu_p(k) + 1, \]
where we used that $m - 1$ and $2m - 3$ are coprime. Now suppose that $p \mid 2m - 3$ and 3 is a primitive root modulo $p$ (thus $p \mid 3^k - 1$ implies $p - 1 \mid k$). If $p - 1 \nmid k$, the $p$-order of the left-hand side is $\leq -1$ and $> -1$ on the right-hand side. Thus, we infer that $p - 1 \mid k$. \hfill \Box

This completes the required ingredients needed in order to prove the main result.

Proof of Theorem 2. Since by assumption $N \mid k$, we can write $k = Nk_1$ and thus rewrite (23) as
\[
(25) \quad 0 < \frac{\log 2}{2N} - \frac{k_1}{2m - 3} < \frac{0.0111}{2N(2m - 3)^2}.
\]

We infer that $k_1/(2m-3) = p_j/q_j$ is a convergent to $(\log 2)/(2N)$ with $j$ even. Since $p \mid m$ implies $p - 1 \nmid k$ (see, e.g., M"orre [26 Proposition 9]), we have $(6, q_j) = 1$. We rewrite (25) as
\[
0 < \frac{\log 2}{2N} - \frac{p_j}{q_j} < \frac{0.0111}{2Nd^2q_j^2},
\]
with $d$ the greatest common divisor of $k_1$ and $2m - 3$. On the other hand,
\[
\frac{\log 2}{2N} - \frac{p_j}{q_j} > \frac{1}{(a_{j+1} + 2)d^2q_j^2},
\]
hence $(a_{j+1} + 2)^{-1} < 0.0111/(2Nd^2)$, from which the result follows by also noting that $2m - 3 \geq q_j$ and invoking Lemma 2 (note that if $\nu_p(q_j) \geq 1$, then $\nu_p(q_j) = \nu_p(2m - 3) - \nu_p(k_1)$). \hfill \Box

To prove that $p \mid m$ implies $p - 1 \nmid k$ one uses that $k$ must be even and takes $l = m$ in (3), showing that $\sum_{p \mid (m, p-1) \mid k} k_p^l$ must be an integer. Since a sum of reciprocals of distinct primes can never be an integer, the result follows.
Table 1. Smallest integers $j$ satisfying conditions (a), (b) and (c) of Theorem 2.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$j = j(N)$</th>
<th>$a_{j+1}$</th>
<th>$q_j$ (rounded down)</th>
<th>$q_j \mod 6$</th>
<th>$p = p(q_j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>642</td>
<td>764</td>
<td>2.383153 · 10^{-130}</td>
<td>−1</td>
<td>149</td>
</tr>
<tr>
<td>2</td>
<td>664</td>
<td>1529</td>
<td>2.383153 · 10^{-130}</td>
<td>−1</td>
<td>149</td>
</tr>
<tr>
<td>2^2</td>
<td>1254</td>
<td>21966</td>
<td>1.132014 · 10^{-638}</td>
<td>+1</td>
<td>5</td>
</tr>
<tr>
<td>2^3</td>
<td>1264</td>
<td>43933</td>
<td>1.132014 · 10^{-638}</td>
<td>+1</td>
<td>5</td>
</tr>
<tr>
<td>2^4</td>
<td>1280</td>
<td>87866</td>
<td>1.132014 · 10^{-638}</td>
<td>+1</td>
<td>5</td>
</tr>
<tr>
<td>2^5</td>
<td>1294</td>
<td>175733</td>
<td>1.132014 · 10^{-638}</td>
<td>+1</td>
<td>5</td>
</tr>
<tr>
<td>2^6</td>
<td>8950</td>
<td>26416</td>
<td>3.458446 · 10^{-589}</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>2^7</td>
<td>8926</td>
<td>52834</td>
<td>3.458446 · 10^{-589}</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>2^8</td>
<td>119476</td>
<td>122799</td>
<td>1.374540 · 10^{-511317}</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3</td>
<td>119008</td>
<td>368398</td>
<td>1.374540 · 10^{-511317}</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3^2</td>
<td>139532</td>
<td>782152</td>
<td>9.351282 · 10^{-71882}</td>
<td>+1</td>
<td>56131</td>
</tr>
<tr>
<td>2^9 · 3^3</td>
<td>6168634</td>
<td>1540283</td>
<td>8.220719 · 10^{-3177670}</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3^4</td>
<td>22383618</td>
<td>5167079</td>
<td>5.128265 · 10^{-11538265}</td>
<td>+1</td>
<td>17</td>
</tr>
<tr>
<td>2^9 · 3^5</td>
<td>155830946</td>
<td>31664035</td>
<td>2.257099 · 10^{-60303211}</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3^5 · 5</td>
<td>351661538</td>
<td>85898211</td>
<td>9.729739 · 10^{-581214202}</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3^5 · 5^2</td>
<td>1738154976</td>
<td>143700727</td>
<td>1.594940 · 10^{-965721905}</td>
<td>+1</td>
<td>5</td>
</tr>
<tr>
<td>2^9 · 3^5 · 5^3</td>
<td>1977626256</td>
<td>853324651</td>
<td>1.196828 · 10^{-1019133881}</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3^5 · 5^4</td>
<td>2015279170</td>
<td>4388327617</td>
<td>5.565196 · 10^{-6038523018}</td>
<td>−1</td>
<td>19</td>
</tr>
<tr>
<td>2^9 · 3^5 · 5^5</td>
<td>3236170820</td>
<td>2307115390</td>
<td>5.427815 · 10^{-1667658416}</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td>2^9 · 3^5 · 5^6</td>
<td>2015385392</td>
<td>21941638090</td>
<td>5.565196 · 10^{-6038523018}</td>
<td>−1</td>
<td>19</td>
</tr>
<tr>
<td>2^9 · 3^5 · 5^7</td>
<td>3236257942</td>
<td>11535576954</td>
<td>5.427815 · 10^{-1667658416}</td>
<td>+1</td>
<td></td>
</tr>
</tbody>
</table>

4. Computation of the continued fractions

We make use of conditions (a), (b), (c) of Theorem 2. We recall that we expect $E(\log q_{(N)}(\xi)) \sim c_1 N$ for a generic $\xi \in [0, 1]$ satisfying these conditions. Indeed, on the basis of theoretical results, heuristics and numerical experiments, we conjecture that $c_1 = 60\pi^2$.

The computation of $(\log 2)/(2N)$ is done in two steps. First, we generate $d$ digits of log 2. For this we use the γ-cruncher [32]. With this program, A. Yee and R. Chan computed 31 billion decimal digits of log 2 in about 24 hours. Second, we set a rational approximation of $(\log 2)/(2N)$ with a relative error bounded by $10^{-d}$. Then partial quotients of the continued fraction of $(\log 2)/(2N)$ are computed: about 0.97$d$ of them can be evaluated, with safe error control [4] (cf. the result of Lochs mentioned in Section 1). We maintain a floating point approximation of numbers $q_j$ (rounded down) and residues of $q_j \mod 6$ by the formula $q_{i+1} = a_{i+1}q_i + q_{i-1}$ for $i \geq 0$, where $q_0 = 1$ and $q_{-1} = 0$.

Table 1 was created with the “basic method” of [4] for $N \leq 2^9 \cdot 3^4$. It was fast enough to reach the benchmark $m > 10^{10^7}$ in four days with $50 \cdot 10^6$ digits of log 2. Bit-complexity of this algorithm (or of the indirect or direct methods [4]) is quadratic and reaching the $m > 10^{10^{10}}$ milestone would take centuries.

Some subquadratic GCD algorithms were discovered that have asymptotic running time $O(n(\log n)^2 \log \log n)$ [24]. A faster version of the program was written: this time a recursive HGCD method is applied. It is adapted for computing a
continued fraction by using Lemma 3 of [4] (which is similar to Algorithm 1.3.13 of [8]) for error control; with it the program leaps over $10^{10^8}$ in just about one hour.

Finally, the new benchmark $m > 10^{10^8}$ is established in no more than 10 hours with $3 \cdot 10^9$ digits of log 2, $N = 1555200$ and condition [d] the first found solution fits conditions [a]–[c] but not [d]. With $N = 7776000$, $m > 10^{10^9}$ is achieved for the smallest $j$. See Table 1 in the last column, $p$ is a prime such that $p \in P(N)$ and $\nu_p(q_j) = 1$, that is, such that condition [d] of Theorem 2 is violated.

5. Miscellaneous

5.1. The number of distinct prime factors of $m - 1$. There is a different application of Theorem 3 suggested by the work of Brenton and Vasiliu [5], to factorization properties of the number $m - 1$ coming from a non-trivial solution $(m,k)$ of (1). A result of Moser [29] (which can also be deduced from the key identity (5), cf. the proof of Lemma 2 above) asserts that

$$\sum_{p \mid m-1} \frac{1}{p} + \frac{1}{m-1} \in \mathbb{Z};$$

in particular, the number $m - 1$ is square-free. Since the sum of reciprocals of the first 58 primes is less than 2, we conclude that either $\omega(m-1) \geq 58$ or the integer in (26) is equal to 1. In the latter case, we can apply Curtiss’s bound [9] for positive integer solutions of Kellogg’s equation

$$\sum_{i=1}^{n} \frac{1}{x_i} = 1,$$

namely, $\max_i \{x_i\} \leq A_n - 1$, where the Sylvester sequence $\{A_n\}_{n \geq 1} = \{2, 3, 7, 43, \ldots \}$ is defined by the recurrence $A_n = 1 + \prod_{i=1}^{n-1} A_i$ (for some further details, see e.g. Odoni [30]). From this result and the estimate $A_n < (1.066 \cdot 10^{13})^{2^{n-7}}$, we infer

$$m < (1.066 \cdot 10^{13})^{2^{\omega(m-1)-6}},$$

which together with the lower bound on $m$ from Theorem 3 yields $\omega(m-1) \geq 33$. A similar estimate on the basis of another (26)-like identity of Moser implies that $\omega(m+1) \geq 32$.

5.2. Generalized EM equation. The method we use in Section 2 for deriving the asymptotics of $k$ in terms of $m$ works for the more general equation

$$(27) \quad 1^k + 2^k + \cdots + (m-1)^k = tm^k,$$

with $t \in \mathbb{N}$ fixed, as well. Indeed, the coefficients in the Taylor series expansion

$$(28) \quad (1 - y)^k e^{k y} = 1 - \frac{k}{2} y^2 - \frac{k}{3} y^3 + \frac{k(k - 2)}{8} y^4 + \cdots = \sum_{n=0}^{\infty} g_n(k) y^n$$

are polynomials satisfying

$$(29) \quad g_0(k) = 1, \quad g_1(k) = 0, \quad \text{and} \quad \deg_k g_n(k) = \left\lfloor \frac{n}{2} \right\rfloor, \quad g_n(0) = 0 \quad \text{for} \ n \geq 2;$$
the latter follows from raising the series \((1 - y)e^y = 1 - y^2/2 - y^3/3 - \cdots\) to the power \(k\). In these settings, equation (27) becomes

\[
t = \sum_{j=1}^{m-1} \left(1 - \frac{j}{m}\right)^k = \sum_{j=1}^{m-1} e^{-kj/m} \sum_{n=0}^{\infty} g_n(k) \left(\frac{j}{m}\right)^n
\]

\[
= \sum_{n=0}^{\infty} g_n(k) \frac{1}{m^n} \sum_{j=1}^{m-1} j^n e^{-jk/m}
\]  

(since \(\sum_{j=m}^{\infty} j^n e^{-jk/m} = O(m^n e^{-k})\))

\[
\sim \sum_{n=0}^{\infty} \frac{g_n(k)}{m^n} \sum_{j=1}^{\infty} j^n e^{-jk/m} = \sum_{n=0}^{\infty} \frac{g_n(k)}{m^n} \left(\frac{z}{z-1}\right)^n \left.\left(\frac{z}{z-1}\right)^n \frac{z}{1-z}\right|_{z=e^{-k/m}}
\]

hence in the notation \(\lambda = k/m\) and \(x = 1/m\) we have

\[
t = \sum_{n=0}^{\infty} g_n \left(\frac{\lambda}{x}\right) (-x)^n \left(\frac{z}{z-1}\right)^n \left.\left(\frac{z}{z-1}\right)^n \frac{z}{1-z}\right|_{z=e^{x}}.
\]

Searching \(\lambda\) in the form \(\lambda = c_0 + c_1 x + c_2 x^2 + \cdots\), we find successively,

\[
c_0 = c(t) = \log \left(1 + \frac{1}{t}\right) = \log \frac{t+1}{t}, \quad c_1 = \left(t + \frac{1}{2}\right) c,
\]

\[
c_2 = \left(t + \frac{1}{2}\right)^3 c^2 - \left(t + \frac{1}{2}\right)^2 c - \frac{1}{4} \left(t + \frac{1}{2}\right) c^2 + \frac{c}{6},
\]

and so on. Note that \(c_n (-(-t+1)) = (-1)^{n+1} c_n(t)\) for \(n = 0, 1, 2, \ldots\); this reflects the equivalence of equation (27) and

\[
1^k + 2^k + \cdots + (m-1)^k + m^k = (t+1)m^k.
\]

From this asymptotics we infer that

\[
\frac{2k}{2m-t_1} = c + \frac{t_1^3 c^2 - 2t_1^2 c - t_1 c^2 + 4c/3}{2(2m-t_1)^2} + O\left(\frac{1}{(2m-t_1)^3}\right),
\]

where \(t_1 = 2t + 1\) and \(c = \log(1+1/t)\). It can be checked that for all positive integers \(t\) we have the inequality

\[-0.22 < t_1^3 c^2 - 2t_1^2 c - t_1 c^2 + \frac{4c}{3} < 0,
\]

and hence \(2k/(2m-2t-1)\) is a convergent (with even index) of this logarithm \(c = \log(1+1/t)\) for \(m\) large enough.

5.3. Saddle-point method. A different approach to treat the asymptotic behaviour of \(k\) in terms of \(m\) for \(k\) and \(m\) satisfying \((11)\) (or, more generally, (27)) is based on the integral representation

\[
1^k + 2^k + \cdots + (m-1)^k = \frac{\Gamma(k)}{2\pi i} \int_{C-i\infty}^{C+i\infty} \frac{e^{mz}}{(e^z - 1)z^{k+1}} dz,
\]
where \( C \) is an arbitrary positive real number (cf. [10, p. 273]). On noting that
\[
\frac{e^{mz}}{e^z - 1} = \frac{e^{(m-1)z}}{1 - e^{-C}} \left( 1 + \frac{1 - e^{-z+C}}{e^z - 1} \right)
\]
one obtains, on taking \( C = (k+1)/(m-1) \) and after invoking some rather trivial estimates, that
\[
1^k + 2^k + \cdots + (m-1)^k = \frac{(m-1)^k}{1 - e^{-(k+1)/(m-1)}} \left( 1 + \rho_k(m) \right),
\]
with
\[
|\rho_k(m)| < \frac{\sqrt{2(k+1)}C}{\sqrt{\pi(k-1)(e^C - 1)}}.
\]
(This part of the argument is due to Delange; for more details see [10, pp. 273–274].) By (2), \( C \) is bounded and we infer that \( |\rho_k(m)| = O(k^{-1/2}) = O(m^{-1/2}) \). On putting \( m^k \) on the left-hand side of (33) and using \( (1-1/m)^m = \exp(-1+O(m^{-1})) \), we immediately conclude that, as \( m \to \infty \),
\[
\frac{k}{m} = \log 2 + O\left( \frac{1}{\sqrt{m}} \right),
\]
where the implied constant is absolute. A more elaborate analysis, using the saddle-point method, will very likely allow one to obtain as many terms in the latter expansion as required.

5.4. Experimental asymptotics. It is worth mentioning a fast experimental approach of doing asymptotics like (22). Given numerically a few hundred terms of a sequence \( s_n \) that one believes has an asymptotic expansion in inverse powers of \( n \), one can try to apply the \texttt{asympt} trick, a simple but often powerful method to numerically determine the coefficients in the ansatz
\[
s_n \sim c_0 + \frac{c_1}{n} + \frac{c_2}{n^2} + \cdots.
\]
As a second step one tries to identify the so-found coefficients with (linear combinations of) known constants. Thus, one arrives at a conjecture that hopefully can be turned into a proof. For more details and some “victories” achieved by the \texttt{asympt} method, see Grünberg and Moree [13].

D. Zagier has applied this trick to the sequence of \( k = k(m) \) obtained from (1) by letting \( m \) run through the first thousand values. Excellent agreement with our theoretical results was obtained in this way.

Acknowledgements

The second author is very indebted to Jerzy Urbanowicz for involving him in the early 1990s in his EM research (with [27] as visible outcome). The second and third author would like to thank D. Zagier for verifying some of our results using the \texttt{asympt} trick and for some informative discussions regarding the saddle-point method (reflected in the final section). H. te Riele provided us with the unpublished report [3], which became the “initial spark” for the current project. C. Baxa pointed out the relevance of [15] to us. Further thanks are due to T. Agoh and I. Shparlinski.
References


[27] P. Moree, H. te Riele, and J. Urbanowicz, Divisibility properties of integers \( x, k \) satisfying \( 1^k + \cdots + (x-1)^k = x^k \), *Math. Comp.* **63** (1994), 799–815. MR1257577 (94m:11041)

[29] L. Moser, On the diophantine equation $1^n + 2^n + 3^n + \cdots + (m-1)^n = m^n$, *Scripta Math.* **19** (1953), 84–88. MR0054627 (14:950g)


12 bis rue Perrey, 31400 Toulouse, France

E-mail address: galloty@orange.fr

Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany

E-mail address: moree@mpim-bonn.mpg.de

School of Mathematical and Physical Sciences, University of Newcastle, Callaghan NSW 2308, Australia

E-mail address: wadim.zudilin@newcastle.edu.au