ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 43, Number 5, 2013

MOSER'S MATHEMAGICAL WORK ON THE EQUATION $1^k + 2^k + \cdots + (m-1)^k = m^k$

PIETER MOREE

In memory of Alf van der Poorten (1942–2010)

ABSTRACT. If the equation of the title has an integer solution with $k \geq 2$, then $m > 10^{10^6}$. Leo Moser showed this in 1953 by amazingly elementary methods. With the hindsight of more than 50 years, his proof can be somewhat simplified. We give a further proof showing that Moser's result can be derived from a von Staudt-Clausen type theorem. Based on more recent developments concerning this equation, we derive a new result using the divisibility properties of numbers in the sequence $\{2^{2e+1}+1\}_{e=0}^{\infty}$. In the final section we show that certain Erdős-Moser type equations arising in a recent paper of Kellner can be solved completely.

1. Introduction. In this paper we are interested in non-trivial solutions, that is, solutions with $k \geq 2$, of the equation

$$1^k + 2^k + \cdots + (m-2)^k + (m-1)^k = m^k. \tag{1}$$

The conjecture that such solutions do not exist was formulated around 1950 by Paul Erdős in a letter to Leo Moser. For $k = 1$, one has the solution $1 + 2 = 3$ (and no further solutions). From now on, we will assume that $k \geq 2$. Moser [**29**] established the following theorem in 1953.

Theorem 1 [**29**]. *If (m, k) is a solution of* (1), *then $m > 10^{10^6}$.*

His result has since been improved. 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 substantially improved. Butske et al. [**6**, page 411]

2010 AMS *Mathematics subject classification.* Primary 11D61, 11A07.
Received by the editors on November 18, 2010, and in revised form on March 21, 2011.
DOI:10.1216/RMJ-2013-43-5-1707

expressed the hope that new insights will eventually make it possible to reach the benchmark 10^{10^7}.

The main purpose of this paper is to make Moser's remarkable proof of Theorem 1 better known, and, with the hindsight and technological developments of more than 50 years, to give an even cleaner version of Moser's proof. This is contained in Section 2.[1] Moreover, we obtain the following refinement of Moser's result.

Theorem 2. *Suppose that* (m, k) *is a solution of* (1) *with* $k \geq 2$. *Then*:

i) $m > 1.485 \cdot 10^{9\,321\,155}$.

ii) k *is even,* $m \equiv 3 \pmod 8$, $m \equiv \pm 1 \pmod 3$;

iii) $m-1$, $(m+1)/2$, $2m-1$ *and* $2m+1$ *are all square-free.*

iv) *If* p *divides at least one of the integers in* (iii), *then* $p-1 \mid k$.

v) *The number* $(m^2-1)(4m^2-1)/12$ *is square-free and has at least* 4 990 906 *prime factors.*

In fact, Moser proved (iii) and (iv) of Theorem 2 and weaker versions of parts (ii) and (v). Readers interested in the shortest (currently known) proof of Theorem 2 are referred to Moree [**25**]. The deepest result used to prove Theorem 2 is Lemma 1. Using a binomial identity due to Pascal (1654) a reproof of Lemma 1 was given recently by MacMillan and Sondow [**18**]. To wit, had Blaise Pascal's computing machine from 1642, the Pascaline[2], worked like a modern computer, then Theorem 2 could have already been proved in 1654.

In Section 3 we compare our alternative proof with Moser's original proof.

In Section 4 we give a more systematic proof of Moser's result, which uses a variant of the von Staudt-Clausen theorem.[3] The relevance of this result for the study of the Erdős-Moser equation was first pointed out in 1996 by Moree [**21**] who used the result to show that the Moser approach can also be used to study the equation $1^k+2^k+\cdots+(m-1)^k = am^k$ and $a \geq 1$, an integer. An improvement of the main result of [**21**] will be presented in Section 8.

The reader might wonder which other techniques have been brought to bear for the study of (1). Such techniques include Bernoulli numbers, considering the equation modulo prime powers, analysis (taking k to be a real, rather than an integer) and continued fraction methods. There is an extensive literature on the more general equation

$$1^k + \cdots + (m-1)^k = y^n, \ n \geq 2,$$

see, e.g., Bennett et al. [**3**]. That work incorporates several further techniques. However, those results do not appear to have any implications for the study of (1). In Section 5, we give a taste of what can be done using Bernoulli numbers and considering (1) modulo prime powers. The main result here is Theorem 1 of [**27**]. We give a weakened (far less technical) version of this, namely, Lemma 4. Using that result and an heuristic assumption on the behavior of $S_r(a)$, an heuristic argument validating the Erdős-Moser conjecture can be given [**26**, Section 6].

In Section 6, we consider implications for (1) based on analytic methods, and in particular, the recent work of Gallot, Moree and Zudilin [**11**] who obtained the benchmark 10^{10^7} and further improved this to 10^{10^9} by computing $3 \cdot 10^9$ digits of $\log 2$.

Section 7 is the most original part of the paper. Results on divisors of numbers of the form $2^{2e+1}+1$ are used to show that if (m, k) is a solution of (1) such that $m+2$ is only composed of primes p satisfying $p \equiv 5, 7 \pmod 8$, then $m \geq 10^{10^{16}}$.

In the final two sections we consider the Erdős-Moser variants

$$1^k + 2^k + \cdots + (m-1)^k = am^k,$$

respectively,

$$a(1^k + 2^k + \cdots + (m-1)^k) = m^k$$

(with $a \geq 1$ a fixed integer) and show that the latter equation (arising in a recent paper by Kellner [**16**]) can be solved completely for infinitely many integers a.

This paper is in part scholarly and in part research. Moser (1921-1970) was a mathematician of the problem solver type. For bibliographic information, the reader is referred to the MacTutor history of mathematics archive [**30**] or Wyman [**40**].

2. Moser's proof revisited. Let $S_r(n) = \sum_{j=0}^{n-1} j^r$. In what follows, we assume that

$$S_k(m) = m^k, \quad k \geq 2, \tag{2}$$

which corresponds to a non-trivial solution of (1). Throughout this note, p will be used to indicate primes.

Lemma 1. *Let p be a prime. We have*

$$S_r(p) \equiv \varepsilon_r(p) \pmod{p},$$

where

$$\varepsilon_r(p) = \begin{cases} -1 & \text{if } p-1 \mid r; \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let g be a primitive root modulo p. In case $p-1 \nmid r$, we have

$$S_r(p) \equiv \sum_{j=0}^{p-2} (g^j)^r \equiv \frac{g^{r(p-1)}-1}{g^r-1} \pmod{p},$$

and the numerator is divisible by p. In case $p-1 \mid r$, we find by Fermat's little theorem that $S_r(p) \equiv p-1 \equiv -1 \pmod{p}$, as desired. $\square$

Another proof using only Lagrange's theorem on roots of polynomials over $\mathbf{Z}/p\mathbf{Z}$ can be given, see Moree [**25**]. The most elementary proof presently known is due to MacMillan and Sondow [**18**] and is based on Pascal's identity (1654), valid for $n \geq 0$ and $a \geq 2$:

$$\sum_{k=0}^{n} \binom{n+1}{k} S_k(a) = a^{n+1} - 1.$$

A further proof can be given using the polynomial identity

$$X^{p-1} - 1 \equiv \prod_{j=1}^{p-1} (X - j) \pmod{p}$$

and Newton's identities expressing power sums in elementary symmetric polynomials.

Lemma 2. *In the case where p is an odd prime or in the case where $p = 2$ and r is even, we have $S_r(p^{\lambda+1}) \equiv pS_r(p^\lambda) \pmod{p^{\lambda+1}}$.*

Proof. Every $0 \leq j < p^{\lambda+1}$ can be uniquely written as $j = \alpha p^\lambda + \beta$ with $0 \leq \alpha < p$ and $0 \leq \beta < p^\lambda$. Hence, we obtain by invoking the binomial theorem,

$$\begin{aligned} S_r(p^{\lambda+1}) &= \sum_{\alpha=0}^{p-1} \sum_{\beta=0}^{p^\lambda-1} (\alpha p^\lambda + \beta)^r \\ &\equiv p \sum_{\beta=0}^{p^\lambda-1} \beta^r + rp^\lambda \sum_{\alpha=0}^{p-1} \alpha \sum_{\beta=0}^{p^\lambda-1} \beta^{r-1} \pmod{p^{2\lambda}}. \end{aligned}$$

Since the first sum equals $S_r(p^\lambda)$ and $2\sum_{\alpha=0}^{p-1} \alpha = p(p-1) \equiv 0 \pmod{p}$, the result follows. □

Proof of Theorem 2. Suppose that $p \mid m - 1$; then using Lemma 1, we infer that

$$\begin{aligned} S_k(m) &= \sum_{i=0}^{(m-1)/p-1} \sum_{j=1}^{p} (j + ip)^k \\ &\equiv \frac{m-1}{p} S_k(p) \equiv \frac{m-1}{p} \varepsilon_k(p) \pmod{p}. \end{aligned} \tag{3}$$

On the other hand, $m \equiv 1 \pmod{p}$, so that by (2) we must have

$$\frac{m-1}{p} \cdot \varepsilon_k(p) \equiv 1 \pmod{p}. \tag{4}$$

Hence $\varepsilon_k(p) \not\equiv 0 \pmod{p}$, so that from the definition of $\varepsilon_k(p)$, it follows that $\varepsilon_k(p) = -1$, and

$$p \mid m - 1 \text{ implies } p - 1 \mid k. \tag{5}$$

Thus, (4) can be put in the form

$$\frac{m-1}{p}+1 \equiv 0 \pmod{p}, \tag{6}$$

or

$$m-1 \equiv -p \pmod{p}^2. \tag{7}$$

We claim that $m-1$ must have an odd prime divisor p and that, hence by (5), k must be even. It is easy to see that $m-1>2$. If $m-1$ does not have an odd prime divisor, then $m-1=2^e$ for some $e \geq 2$. However, by (7), we see that $m-1$ is square-free. This contradiction shows that $m-1$ has indeed an odd prime factor p.

We now multiply together all congruences of the type (6), that is, one for each prime p dividing $m-1$. Since $m-1$ is square-free, the resulting modulus is $m-1$. Furthermore, products containing two or more distinct prime factors of the form $(m-1)/p$ will be divisible by $m-1$. Thus, we obtain

$$(m-1)\sum_{p \mid m-1}\frac{1}{p}+1 \equiv 0 \pmod{m-1}, \tag{8}$$

or

$$\sum_{p \mid m-1}\frac{1}{p}+\frac{1}{m-1} \equiv 0 \pmod{1}. \tag{9}$$

We proceed to develop three more congruences, similar to (9) which, when combined with (9), lead to the proof of part 1. Equation (2) can be written in the form

$$S_k(m+2)=2m^k+(m+1)^k. \tag{10}$$

Using Lemma 1 and the fact that k is even, we obtain, as before,

$$p \mid m+1 \text{ implies } p-1 \mid k, \tag{11}$$

and

$$\frac{m+1}{p}+2 \equiv 0 \pmod{p}. \tag{12}$$

From (12), it follows that no odd prime appears with exponent greater than one in the prime factorization of $m+1$. The prime 2 (according to H. Zassenhaus 'the oddest of primes'), requires special attention. If we inspect (1) with modulus 4 and use the fact that k is even, then we find that $m+1 \equiv 1$ or $4 \pmod 8$. Now let us assume that we are in the first case, and we let $2^f \parallel m$ (that is, $2^f \mid m$ and $2^{f+1} \nmid m$). Note that $f \geq 3$. By an argument similar to that given in (3) we infer that $S_k(m+1) \equiv (m/2^f)S_k(2^f) \pmod{2^f}$. Using Lemma 2, we see that $S_k(m+1) \equiv (m/2^f)S_k(2^f) \equiv 2^{f-1} \pmod{2^f}$, contradicting $S_k(m+1) = 2m^k \equiv 0 \pmod{2^f}$. Thus, $m+1$ contains 2 exactly to the second power, and hence (12) can be put in the form

$$\frac{m+1}{2p} + 1 \equiv 0 \pmod{p}. \tag{13}$$

We multiply together all congruences of type (13). The modulus then becomes $(m+1)/2$. Further, any term involving two or more distinct factors $(m+1)/(2p)$ will be divisible by $(m+1)/2$, so that on simplification we obtain

$$\sum_{p \mid m+1} \frac{1}{p} + \frac{2}{m+1} \equiv 0 \pmod{1}. \tag{14}$$

We proceed to find two similar equations to (14). Suppose that $p \mid 2m-1$, and let $t = ((2m-1)/p - 1)/2$. Clearly t is an integer, and $m - 1 = tp + (p-1)/2$. We have $a^k = (-a)^k$ since k is even so that $2S_k((p+1)/2) \equiv S_k(p) \pmod p$ and, hence, by Lemma 1,

$$S_k\left(\frac{p+1}{2}\right) \equiv \frac{\varepsilon_k(p)}{2} \pmod{p}.$$

It follows that

$$S_k(m) \equiv \sum_{i=0}^{t-1}\sum_{j=1}^{p-1}(j+ip)^k + \sum_{i=1}^{(p-1)/2} i^k \equiv \left(t + \frac{1}{2}\right)\varepsilon_k(p) \pmod{p}. \tag{15}$$

On the other hand, $1 \equiv (2m-1+1)^k \equiv (2m)^k \pmod p$; hence, $m^k \not\equiv 0 \pmod p$, so that (2) and (15) imply $\varepsilon_k(p) \neq 0$. Hence, $p-1 \mid k$ and, by Fermat's little theorem, $m^k \equiv 1 \pmod p$. Thus, (2) and

(15) yield $-(t+(1/2)) \equiv 1 \pmod p$. Replacing t by its value and simplifying, we obtain

$$\frac{2m-1}{p} + 2 \equiv 0 \pmod p. \tag{16}$$

Since $2m-1$ is odd, (16) implies that $2m-1$ is square-free. Multiplying congruences of the type (16), one for each of the r prime divisors of $2m-1$, yields

$$2^{r-1}\left((2m-1)\sum_{p\mid 2m-1}\frac{1}{p} + 2\right) \equiv 0 \pmod{2m-1}.$$

Since the modulus $2m-1$ is odd, this gives

$$\sum_{p\mid 2m-1}\frac{1}{p} + \frac{2}{2m-1} \equiv 0 \pmod 1. \tag{17}$$

Finally, we obtain a corresponding congruence for primes p dividing $2m+1$, namely, (19) below. For this purpose, we write (2) in the form

$$S_k(m+1) = 2m^k. \tag{18}$$

Suppose $p \mid 2m+1$. Set $v = ((2m+1)/p-1)/2$. Clearly v is an integer. We have $m = pv + (p-1)/2$ and find $S_k(m+1) \equiv (v+(1/2))\varepsilon_k(p) \pmod p$. From this and (18), it is easy to infer that $\varepsilon_k(p) = -1$, and so $v + (1/2) \equiv -2 \pmod p$. We conclude that

$$p \mid 2m+1 \text{ implies } p-1 \mid k.$$

Replacing v by its value and simplifying, we obtain

$$\frac{2m+1}{p} + 4 \equiv 0 \pmod p.$$

Note that this implies that $2m+1$ is square-free. Reasoning as before, we obtain

$$\sum_{p\mid 2m+1}\frac{1}{p} + \frac{4}{2m+1} \equiv 0 \pmod 1. \tag{19}$$

If we now add the left hand sides of (9), (14), (17) and (19), we get an integer, at least 4. By an argument similar to that showing $2 \nmid m$, we show that $3 \nmid m$ (but in this case we use Lemma 2 with $p = 3$ and $3^\lambda \parallel m$ and the fact that k must be even). No prime $p > 3$ can divide more than one of the integers $m-1$, $m+1$, $2m-1$ and $2m+1$. Further, since $m \equiv 3 \pmod 8$ and $3 \nmid m$, 2 and 3 divide precisely two of these integers. We infer that $M_1 = (m-1)(m+1)(2m-1)(2m+1)/12$ is a square-free integer. We deduce that

$$\sum_{p \mid M_1} \frac{1}{p} + \frac{1}{m-1} + \frac{2}{m+1} + \frac{2}{2m-1} + \frac{4}{2m+1} \geq 4 - \frac{1}{2} - \frac{1}{3} = 3\frac{1}{6}. \tag{20}$$

One checks that (17) has no solutions with $m \leq 1\,000$. Thus, (20) yields (with $\alpha = 3.16$) $\sum_{p \mid M_1} (1/p) > \alpha$. From this, it follows that, if

$$\sum_{p \leq x} \frac{1}{p} < \alpha, \tag{21}$$

then $m^4/3 > M_1 > \prod_{p \leq x} p$, and hence

$$m > 3^{1/4} e^{\theta(x)/4}, \tag{22}$$

with $\theta(x) = \sum_{p \leq x} \log p$, the Chebyshev θ-function. Since, for example, (21) is satisfied with $x = 1\,000$, we find that $m > 10^{103}$ and infer from (20) that we can take $\alpha = 19/6 - 10^{-100}$ in (21). Next one computes (using a computer algebra package, say PARI) the largest prime p_k such that $\sum_{p \leq p_k} (1/p) < 19/6$, with $p_1, p_2, \ldots$ the consecutive primes. Here one finds that $k = 4\,990\,906$ and

$$\sum_{i=1}^{4\,990\,906} \frac{1}{p_i} = 3.166\,666\,658\,810\,172\,858\,4 < 3\frac{1}{6} - 10^{-9}.$$

This completes the proof of part 1 of the theorem; the remaining parts of the theorem have been proved along the way. $\square$

Remark 1. Since, for a solution of (1), $(m^2-1)(4m^2-1)/12$ has at least $4\,990\,906$ distinct prime factors, it is perhaps reasonable to expect that each of the factors $m-1$, $m+1$, $2m-1$ and $2m+1$ must have many distinct prime factors. Brenton and Vasiliu [**5**], using the bound given in part 1 of Theorem 2, showed that $m-1$ has at least 26 prime factors. Gallot et al. [**11**] increased this, using Theorem 5, to 33.

Remark 2. Moser [**1**] considered modulo $m-1$, $m+1$, $2m-1$ and $2m+1$. Sondow and MacMillan [**38**] considered the equation also modulo $(m-1)^2$ and obtained some further information (this involves the Fermat quotient).

3. Comparison of the proof with Moser's. In this section we compare and contrast the proof of Theorem 2 with Moser's proof of Theorem 1.

Moser used only Lemma 1, not Lemma 2. Consequently, he concluded that either $m \equiv 3 \pmod 8$ or $m \equiv 0 \pmod 8$. In the first case we followed his proof, but in the second case one has to note that we cannot use (14). Letting $M_2 = (m-1)(2m-1)(2m+1)$, we get from (9), (17), (19),

$$\sum_{p \mid M_2} \frac{1}{p} + \frac{1}{m-1} + \frac{2}{2m-1} + \frac{4}{2m+1} > 3 - \frac{1}{3}. \tag{23}$$

However, since $2 \nmid M_2$, (23) is actually a stronger condition on m than is (20).

The idea to use $3 \nmid m$, leading to a slight improvement for the bound on m, is taken from Butske et al. [**6**] and not present in Moser's proof. (Actually, they consider the cases $3 \nmid m$ and $3 \mid m$ separately. We show that only $3 \nmid m$ can occur.)

By using some prime number estimates from Rosser, Moser deduces that (21) holds with $x = 10^7$ and $\alpha = 3.16$. In his argument, he claims that, by direct computation, one sees that (21) holds with $x = 1\,000$ and $\alpha = 2.18$. This is not true (as pointed out to me by Buciumas and Havarneanu). However, replacing 2.18 by 2.2 in Moser's equation (21), one sees that his proof still remains valid. The present day possibilities of computers allow us to proceed by direct computation, rather than resorting to prime number estimates as Moser was forced to do.

The advantage of the proof given in Section 2 is that it shows, in contrast to Moser's proof and Butske et al.'s variation thereof, that *every* non-trivial solution satisfies the crucial inequality (20).

4. A second proof using a von Staudt-Clausen type theorem. In this section we show that Moser's four formulas (9), (14), (17) and (19) can be easily derived from the following theorem. Indeed, using it, a fifth formula can be derived, namely (26), below.

Theorem 3 (Carlitz-von Staudt, 1961). *Let r, y be positive integers. Then*:

$$S_r(y) = \sum_{j=1}^{y-1} j^r = \begin{cases} 0 \pmod{y(y-1)/2} & \text{if } r \textit{ is odd;} \\ -\sum_{p-1|r,\ p|y}(y/p) \pmod{y} & \textit{otherwise.} \end{cases} \tag{24}$$

Carlitz [**7**] gave a proof of Theorem 3 using finite differences and stated that the result is due to von Staudt. In the case where r is odd, he claims that $S_r(y)/y$ is an integer, which is not always true (it is true though that $2S_r(y)/y$ is always an integer). The author [**20**] gave a proof of a generalization to sums of powers in arithmetic progression using the theory of primitive roots. Kellner [**15**] gave a reproof for even r only using Stirling numbers of the second kind. For the easiest proof known and some further applications of the Carlitz-von Staudt theorem, we refer the reader to Moree [**25**].

Second proof of Theorem 2. We will apply Theorem 3 with $r = k$.

In case k is odd, we find by combining (24) (with $y = m$) with (1) and using the coprimality of m and $m-1$, that $m = 2$ or $m = 3$, but these cases are easily excluded. Therefore, k must be even.

Take $y = m-1$. Then, using (1), the left hand side of (24) simplifies to

$$\begin{aligned} S_k(m-1) &= 1^k + 2^k + \cdots + (m-2)^k \\ &= m^k - (m-1)^k \\ &\equiv 1 \pmod{m-1}. \end{aligned}$$

We get from (24) that

$$\sum_{\substack{p|m-1 \\ p-1|k}} \frac{(m-1)}{p} + 1 \equiv 0 \pmod{m-1}. \tag{25}$$

Suppose there exists a $p \mid m-1$ such that $p-1 \nmid k$. Reducing both sides modulo p, we get $1 \equiv 0 \pmod{p}$. This contradiction shows that, in (25), the condition $p-1 \mid k$ can be dropped, and thus we obtain (8). From (8), we see that $m-1$ must be square-free and also we obtain (9).

Take $y = m$. Then, using (1) and $2 \mid k$, we infer from (24) that

$$\sum_{\substack{p-1 \mid k \\ p \mid m}} \frac{1}{p} \equiv 0 \pmod{1}. \tag{26}$$

Since a sum of reciprocals of distinct primes can never be a positive integer, we infer that the sum in (26) equals zero and hence conclude that, if $p-1 \mid k$, then $p \nmid m$. We conclude, for example, that $(6, m) = 1$. Now, on considering (1) with modulus 4, we see that $m \equiv 3 \pmod{8}$.

Take $y = m+1$. Then, using (1) and the fact that k is even, the left hand side of (24) simplifies to

$$S_k(m+1) = S_k(m) + m^k = 2m^k \equiv 2 \pmod{m+1}.$$

We obtain

$$\sum_{\substack{p \mid m+1 \\ p-1 \mid k}} \frac{(m+1)}{p} + 2 \equiv 0 \pmod{m+1},$$

and by reasoning as in the case $y = m-1$, it is seen that $p \mid m+1$ implies $p-1 \mid k$, and thus (14) is obtained. From (14) and $m \equiv 3 \pmod{8}$, we derive that $(m+1)/2$ is square-free.

Take $y = 2m-1$. On noting that

$$\begin{aligned} S_k(2m-1) &= \sum_{j=1}^{m-1} (j^k + (2m-1-j)^k) \equiv 2S_k(m) \\ &\equiv 2m^k \pmod{2m-1}, \end{aligned}$$

we find that

$$\sum_{\substack{p \mid 2m-1 \\ p-1 \mid k}} \frac{(2m-1)}{p} + 2m^k \equiv 0 \pmod{2m-1}. \tag{27}$$

Since m and $2m-1$ are coprime, we infer that, if $p \mid 2m-1$, then $p-1 \mid k$, $m^k \equiv 1 \pmod{p}$, and furthermore that $2m-1$ is square-free. It follows from the Chinese remainder theorem that $2m^k \equiv 2 \pmod{2m-1}$, and hence from (27) we obtain (17).

Take $y = 2m+1$. Noting that

$$\begin{aligned} S_k(2m+1) &= \sum_{j=1}^{m} (j^k + (2m+1-j)^k) \equiv 2S_k(m+1) \\ &= 4m^k \pmod{2m+1} \end{aligned}$$

and proceeding as in the case $y = 2m-1$, we obtain (19) and the square-freeness of $2m+1$. To finish the proof, we proceed as in Section 2 just below (19). $\square$

With some of the magic behind the four Moser identities revealed, the reader might be well tempted to derive further identities. A typical example would start from

$$4^k - 1^k - 2^k - 3^k \equiv -\sum_{\substack{p-1 \mid k \\ p \mid m-4}} \frac{(m-4)}{p} \pmod{m-4}. \tag{28}$$

For simplicity, let us assume that $m \equiv 2 \pmod{3}$. We have $(6, m-4) = 1$. For this to lead to a further equation, we need the left hand side to be a constant modulo $m-4$. If we could infer that $p \mid m-4$ implies $p-1 \mid k$, then the left hand side would equal $-2 \pmod{m-4}$, and we would be in business. (For the reader familiar with the Carmichael function λ, this can be more compactly formulated as $\lambda(m-4) \mid k$.) Unfortunately, a problem is caused by the fact that the left hand side could be divisible by p. Thus, all we seem to obtain is that if $m \equiv 2 \pmod{3}$, and $\lambda(m-4) \mid k$ or $4^k - 1^k - 2^k - 3^k$ and $m-4$ are coprime, then

$$\sum_{p \mid m-4} \frac{1}{p} - \frac{2}{m-4} \equiv 0 \pmod{1}.$$

In Section 7, we will see that, if we replace $m-4$ by $m+2$, we can do a little better, the reason being that, in this case, $2^{k+1}+1$ appears on the left hand side, and numbers of this form have only a restricted set of possible prime factors.

5. Bernoulli numbers and a cascade process. Recall that the Bernoulli numbers B_k are defined by the power series

$$\frac{t}{e^t-1}=\sum_{k=0}^{\infty}\frac{B_k t^k}{k!}.$$

They are rational numbers and can be written as $B_k=U_k/V_k$, with $(U_k,V_k)=1$. One has $B_0=1$, $B_1=-1/2$ and $B_{2j+1}=0$ for $j\geq 1$. By the von Staudt-Clausen theorem we can take for $k\geq 2$ even $V_k=\prod_{p-1|k}p$. The Kummer congruences state that if k and r are even and $k\equiv r\not\equiv 0 \pmod{p-1}$, then $B_k/k\equiv B_r/r \pmod p$. A prime p will be called *regular* if it does not divide any of the numerators U_k with k even and $\leq p-3$. Otherwise, it is said to be *irregular*. The first few irregular primes are $37,\ 59,\ 67,\ 101,\dots$.

The power sum $S_r(n)$ can be expressed using Bernoulli numbers. One has, see e.g., [**33**, (2.1)],

$$S_r(n)=\sum_{j=0}^{r}\binom{r}{j}B_{r-j}\frac{n^{j+1}}{j+1}.$$

Voronoi in 1889, see, e.g., [**33**, Theorem 2.8], proved that if k is even and ≥ 2, then $V_rS_r(n)\equiv U_kn \pmod{n^2}$. From this result, we infer that, for a solution (k,m) of [**1**], we must have $m\mid U_k$, and thus in particular, $\nu_p(U_k)\geq\nu_p(m)$, where we put $\nu_p(m)=f$ if $p^f\parallel m$. By a more elaborate analysis, Moree et al. [**27**] improved this to $\nu_p(B_k/k)\geq 2\nu_p(m)$. It shows (by the von Staudt-Clausen theorem) that if $p\mid m$, then $p-1\nmid k$ (a conclusion we already reached using identity (26)). Invoking the Kummer congruences, we then obtain the following result.

Lemma 3. *Let* (k,m) *be a solution of* (1) *with* $k\geq 2$ *and even. If* $p\mid m$*, then* p *is irregular.*

Let us call a pair (r,p) with p a regular prime and $2\leq r\leq p-3$ even, *helpful* if, for every $a=1,\dots,p-1$, we have $S_r(a)\not\equiv a^r \pmod p$.

Lemma 4. *If* (r,p) *is a helpful pair and* (k,m) *a solution of* (1) *with* k *even, then we have* $k\not\equiv r \pmod{p-1}$.

Proof. Assume that $k \equiv r \pmod{p-1}$. By the previous lemma, we must have $p \nmid m$. Now write $m = m_0 p + b$. Thus, $1 \le b \le p-1$. We have, modulo p, $S_k(m) \equiv S_r(m) \equiv m_0 S_r(p) + S_r(b) \equiv S_r(b)$. Thus, if (1) is satisfied, we must have $S_r(b) \equiv b^r \pmod{p}$. By the definition of a helpful pair, this is impossible. □

Since $2 \mid k$, and $(2,5)$ is a helpful pair, we infer that $4 \mid k$. Since $(2,7)$ and $(4,7)$ are helpful pairs, it follows that $6 \mid k$. From $4 \mid k$ and the fact that $(4,17)$ and $(12,17)$ are helpful pairs, it follows that $8 \mid k$. We thus infer that $24 \mid k$. It turns out that this process can be continued to deduce that more and more small prime factors must divide k; for a detailed account with many tables, see [**26**]. Given an irregular prime p and $2 \le r \le p-3$ even, one would heuristically expect that it is helpful with probability $(1-1/p)^{p-1}$ which tends to $1/e$, assuming that the values $S_r(a)$ are randomly distributed modulo p; this is supported by current numerical data.

Moree et al. [**27**], using good pairs (of which the helpful pairs are a special case), showed that $N_1 := \operatorname{lcm}(1,2,\ldots,200)$ divides k. Kellner [**14**] showed in 2002 that also all primes $200 < p < 1000$ divide k. Actually Moree, et al. [**27**, page 814] proved a slightly stronger result which, combined with Kellner's, shows that $N_2 \mid k$ with

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

An heuristic argument can be given, suggesting that if $L_v := \operatorname{lcm}(1,2,\ldots,v)$ divides k, with tremendously high likelihood we can infer that L_w divides k, where w is the smallest prime power not dividing L_v. It suffices that $v \ge 11$. To deduce that k is divisible by 24 is delicate, but once one has $L_v \mid k$, there is an explosion of further helpful pairs one can use to establish divisibility of k by a larger integer. To add the first prime power w not dividing L_v, one needs a number of helpful pairs that is roughly linear in v, whereas an exponential number (in v) is available. However, the required computation time increases sharply with w. This heuristic argument is the most convincing known to the author in support of the Erdős-Moser conjecture; details may be found in [**26**], the extended version of [**27**].

Given a fixed integer a, one can try the same approach to study the equation $S_k(m) = am^k$. Again, one sees that once one manages to infer, for example, that $120 \mid k$, one can show that there must be larger and larger divisors. For many a, however, this 'cascade process' does not seem to 'take off,' and it remains unknown whether all solutions with $k \geq 2$, if any, satisfy $120 \mid k$, for example.

If $n = \prod_i p_i^{e_i}$ denotes the canonical prime factorization of n, then $\Omega(n) = \sum e_i$ is the total number of prime divisors of n. Urbanowicz [**39**] proved a result which implies that, given an arbitrary t, there exists an integer m_t such that, if (k, m) is a solution of (1) with $k \geq 2$ and $m \geq m_t$, then $\Omega(k) \geq t$.

6. The analytical approach and continued fractions. Comparing $S_k(m)$ with the appropriate integrals, it is easy to see that the ratio k/m must be bounded. A more refined approach gives

$$S_k(m) = \frac{(m-1)^k}{1 - e^{-(k+1)/(m-1)}}\left(1 + O\left(\frac{1}{\sqrt{m}}\right)\right).$$

On equating the left-hand side to m^k and using $(1-1/m)^m = \exp(-1+O(m^{-1}))$, one concludes that, as $m \to \infty$, we have

$$\frac{k}{m} = \log 2 + O\left(\frac{1}{\sqrt{m}}\right).$$

By a rather more delicate analysis, Gallot et al. [**11**] obtain that, for $m > 10^9$, one has

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

As a corollary, this gives that, if (k, m) is a solution of (1) with $k \geq 2$ and even, then $2k/(2m-3)$ is a convergent p_j/q_j of $\log 2$ with j even. This approach was first explored in 1976 by Best and te Riele [**4**] in their attempt to solve a related conjecture by Erdős, see also Guy [**12**, D7]. The main result of [**11**] reads as follows: where given $N \geq 1$, we define

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

Theorem 4. *Let $N \geq 1$ be an arbitrary integer. Let*

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

be the (regular) continued fraction of $(\log 2)/(2N)$, *with* $p_i/q_i = [a_0, a_1, \ldots, a_i]$ *its ith 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 even integer such that:*

a) $a_{j+1} \geq 180N - 2$;

b) $(q_j, 6) = 1$;

c) $\nu_p(q_j) = \nu_p(3^{p-1} - 1) + \nu_p(N) + 1$ *for all primes* $p \in \mathcal{P}(N)$ *dividing* q_j.

Then $m > q_j/2$.

Note that if, for some integer N, we could prove that if all continued fraction digits a_i satisfy $a_i \leq 100N$, say, and $N \mid k$, then (1) would be resolved! However, for a generic number, $\xi \in [0, 1]$, that is not a rational, one can show that the sequence of a_i is not bounded above. The Gauss-Kuz'min statistics make this more precise and assert that the probability that a given term in the continued fraction expansion of a generic ξ is at least b, equals $\log_2(1 + 1/b)$. Thus, for a sufficiently large N, one expects that $j(N)$ is quite large. This, in combination with the exponential growth of q_j then ensures a large lower bound for m. (The numbers $(\log 2)/2N$ are expected to be generic.)

Conditions b) and c) are of lesser importance. It seems that condition b) is satisfied with probability 1/2. In practice, sometimes condition a) is satisfied, but not b) or c), and this leads to a larger lower bound for m. Condition c) is derived using the Moser method, namely, by analyzing the equation

$$\frac{2(3^k - 1)(m - 1)^k}{2m - 3} \equiv - \sum_{\substack{p \mid 2m-3 \\ p-1 \mid k}} \frac{1}{p} \pmod{1}, \tag{29}$$

that a solution (m, k) of (1) must satisfy.

We leave it as a challenge to experts in metric theory of continued fractions to determine the expected value of $q_{j(N)}$ on replacing $(\log 2)/(2N)$ above by a generic number ξ. Gallot et al. [**11**] expect that conditions a) and b) lead to $E(\log q_{j(N)}(\xi)) \sim c_1 N$ and, taking into account also condition c), $E(\log q_{j(N)}(\xi)) \sim c_2 N \log^{\beta} N$ for some positive constants c_1, c_2 and β.

Crucial in applying the result is a very good algorithm to determine $\log 2$ with many decimals of accuracy. Indeed, it is a well-known result of Lochs, that if one knows a generic number up to n decimal digits, then one can accurately compute approximately 0.97n continued fraction digits. For example, knowing 1000 decimal digits of π allows one to compute 968 continued fraction digits.

Applying Theorem 4 with $N = 2^8 \cdot 3^5 \cdot 5^3$ or $N = 2^8 \cdot 3^5 \cdot 5^4$, and using that $N \mid N_2$ and $N_2 \mid k$, Gallot et al. obtained the current world record:

Theorem 5. *If an integer pair* (m,k) *with* $k \geq 2$ *satisfies* (1), *then*

$$m > 2.7139 \cdot 10^{1\,667\,658\,416} > 10^{10^9}.$$

Gallot et al. argue that, assuming one can compute $\log 2$ with arbitrary precision, applying Theorem 4 with $N = N_2$ should give rise to $m > 10^{10^{400}}$.

Interestingly, the results obtained by invoking Bernoulli numbers ('arithmetic') and analysis seem to be completely unrelated ('the arithmetic does not feel the analysis'). This strongly suggests that the Erdős-Moser conjecture ought to be true.

7. A new result. This section focuses on new research; familiarity with the theory of divisors of second order sequences is helpful. The reader is referred to Ballot [**2**] or Moree [**24**] for more introductory accounts.

Let S be an infinite sequence of positive integers. We say that a prime p divides the sequence if it divides at least one of its terms. Here we will be interested in the sequence $S_2 := \{2^{2e+1}+1\}_{e=0}^{\infty}$. It can be shown that $p > 2$ divides S_2 if and only if $\mathrm{ord}_2(p) \equiv 2 \pmod 4$, with $\mathrm{ord}_g(p)$ (with $p \nmid g$) the smallest positive integer t such that $g^t \equiv 1 \pmod p$. The set of these primes is known to have natural density $7/24$ [**22**]. Furthermore, if $\mathrm{ord}_2(p) \equiv 2 \pmod 4$, then

$$p \mid 2^{2e+1} + 1 \quad \text{if and only if } 2e \equiv \frac{\mathrm{ord}_2(p)}{2} - 1 \pmod{\mathrm{ord}_2(p)}. \tag{30}$$

In some coding theoretical work the sequence S_2 and its variants play an important role, as in [**8, 13**] and similarly in the study of the Stufe of cyclotomic fields [**9, 22**] and the study of Fermat varieties [**31, 37**].

If $m+2$ is coprime with S_2, then from (33) and $2 \mid k$ we can infer a fifth identity of Moser type, (32). This then leads to $m > 10^{10^{11}}$ for such m. We now consider the situation in greater detail.

Theorem 6. *Let* $N \equiv 0 \pmod{24}$ *be an arbitrary integer. Suppose that* (m,k) *is a solution of* (1) *with*

$$k \geq 2, \qquad N \mid k \quad \textit{and} \quad m < 10^{10^{11}}.$$

Then $m+2$ *has a prime divisor* $p > 3$ *such that*:

1) $(\mathrm{ord}_2(p), N) = 2$;

2) $k \equiv \mathrm{ord}_2(p)/2 - 1 \pmod{\mathrm{ord}_2(p)}$.

In the case where $m \equiv 2 \pmod 3$, *we can replace* $10^{10^{11}}$ *by* $10^{10^{16}}$. *In the case* $N = N_2$, *we have* $p \geq 2\,099$.

We first prove a corollary.

Corollary 1. *Suppose every prime divisor* $p > 3$ *of* $m+2$ *satisfies* $p \equiv 5, 7 \pmod 8$. *Then*

$$m \geq \begin{cases} 10^{10^{16}} & \text{if } 3 \nmid m+2; \\ 10^{10^{11}} & \text{if } 3 \mid m+2. \end{cases} \tag{31}$$

Proof. Using the supplementary law of quadratic reciprocity, $(2/p) = (-1)^{(p^2-1)/8}$, one sees that, if $p \equiv 5, 7 \pmod 8$, then $\mathrm{ord}_2(p) \not\equiv 2 \pmod 4$. Thus, condition 1 is not satisfied, as for it to be satisfied we must have $\mathrm{ord}_2(p) \equiv 2 \pmod 4$. $\square$

Put

$$P(N) = \{p > 3 : (\mathrm{ord}_2(p), N) = 2\}.$$

We will study the set $P(N)$ in greater detail with the ultimate goal of studying the N-good integers, that is, the odd integers n having no prime divisors in $P(N)$. Note that, in the proof of Corollary 1, we established that integers composed only of primes $p \equiv 5, 7 \pmod 8$ are N-good (with $24 \mid N$).

Corollary 2. *Let $N \equiv 0 \pmod{24}$ be an arbitrary integer. If (m, k) satisfies* (1)*, $N|k$ and $m+2$ is N-good, then m satisfies inequality* (31).

If p is to be in $P(N)$, then $p \equiv 1 \pmod 8$ or $p \equiv 3 \pmod 8$. In the latter case, we have $\mathrm{ord}_2(p) \equiv 2 \pmod 4$. In the former case, it is not necessarily so that $\mathrm{ord}_2(p) \equiv 2 \pmod 4$, and numerically there is a strong preponderance of primes $p \equiv 3 \pmod 8$ in $P(N)$. Indeed, we have the following result.

Lemma 5. *The relative density of primes $p \equiv 1 \pmod 8$ satisfying $\mathrm{ord}_2(p) \equiv 2 \pmod 4$ within the set of primes $p \equiv 1 \pmod 8$ is $1/6$.*

Proof. We have seen that, if $\mathrm{ord}_2(p) \equiv 2 \pmod 4$, then $p \equiv 1, 3 \pmod 8$. If $p \equiv 3 \pmod 8$, then $\mathrm{ord}_2(p) \equiv 2 \pmod 4$. From this, the fact that $\delta(\mathrm{ord}_2(p) \equiv 2 \pmod 4) = 7/24$ and the prime number theorem for primes in arithmetic progression, we infer that the density of primes $p \equiv 1 \pmod 8$ is such that $\mathrm{ord}_2(p) \equiv 2 \pmod 4$ equals $(7/24) - (1/4) = (1/24)$. The sought for relative density is then $(1/24)/(1/4) = 1/6$. $\square$

Thus, if $p \equiv 3 \pmod 8$, then $\mathrm{ord}_2(p) \equiv 2 \pmod 4$ and, if $p \equiv 1 \pmod 8$, then in 1/6th of the cases we have $\mathrm{ord}_2(p) \equiv 2 \pmod 4$.

A further observation concerning the set $P(N)$ is related to Sophie Germain primes. A prime q such that $2q + 1$ is a prime, is called a *Sophie Germain prime*. Let q_M denote the largest prime factor of M.

Lemma 6. *Let $N \equiv 0 \pmod{24}$ be an arbitrary integer. If q is a Sophie Germain prime, $q \equiv 1 \pmod 4$ and q and N are coprime, then $p = 2q + 1 \in P(N)$.*

Proof. The assumptions imply that $(2/p) = -1$ and, since $p > 3$, we infer that $\mathrm{ord}_2(p) = 2q$. Since $(\mathrm{ord}_2(p), N) = (2q, N) = 2$, we are done. $\square$

There are 42 primes p in $P(N_2)$ not exceeding 10 000. Of those 7 primes, p are such that $(p-1)/2$ is not Sophie Germain, the smallest one being 7 699. However, the Sophie Germain primes have natural density zero, whereas as we shall see $P(N)$ has positive natural density.

Given a rational number g such that $g \notin \{-1, 0, 1\}$, the natural density $\delta_g(d)$ of the set of primes p such that the order of $g \pmod p$ is divisible by d is known to exist and can be computed, see e.g., Moree [**23**]. Using inclusion and exclusion, one then finds that the set $P(N)$ has natural density

$$\delta(N) = \sum_{d|N_0} (\delta_2(2d) - \delta_2(4d))\mu(d),$$

where N_0 is the product of the odd prime divisors dividing N and μ denotes the Möbius function. By Moree [**23**, Theorem 2], we then find that, for odd d,

$$\delta_2(2d) - \delta_2(4d) = \frac{7}{24} \prod_{p|d} \frac{p}{p^2 - 1},$$

and hence

$$\delta(N) = \frac{7}{24} \sum_{d|N_0} \mu(d) \prod_{p|d} \frac{p}{p^2 - 1} = \frac{7}{24} \prod_{p|N_0} \left(1 - \frac{p}{p^2 - 1}\right),$$

where we used that a multiplicative function f satisfies

$$\sum_{d|N_0} \mu(d) f(d) = \prod_{p]midN_0} (1 - f(p)).$$

Taking $N = N_2$, one finds that

$$\delta(N_2) = \frac{7}{24} \prod_{2<p\leq 1000} \left(1 - \frac{p}{p^2 - 1}\right) \approx 0.043\,578\,833\ldots.$$

By a result of Wiertelak, quoted as Theorem 1 in [**23**], we have

$$\sum_{\substack{p\leq x \\ p\notin P(N)}} 1 = (1 - \delta(N)) \frac{x}{\log x} + O_N\left(\frac{x}{\log^2 x}\right),$$

where the implicit constant may depend on N. From this result and [**10**, Proposition 4], we then infer that asymptotically the number of integers $n \le x$ that are N-good, $N_G(x)$, satisfies

$$N_G(x) \sim c_N x \log^{-\delta(N)} x,$$

where

$$c_N = \frac{1}{\Gamma(1-\delta(N))} \lim_{x\to\infty} \prod_{p\le x} \left(1-\frac{1}{p}\right)^{1-\delta(N)} \left(1-\frac{\chi_N(p)}{p}\right)^{-1},$$

with $\chi_N(p) = 0$ if $p = 2$ or p is in $P(N)$ and 1 otherwise. (As usual Γ denotes the Gamma-function.) Taking $N = N_2$, a computer calculation suggests that $c_{N_2} \approx 0.54$.

Now, if we have a sequence of random integers n_j growing roughly as $e^{\beta j}$ for some constant $\beta > 0$, the integer n_j is N-good with probability $c_N \log^{-\delta(N)} n_j \approx c_N(\beta j)^{-\delta(N)}$. The expected number of N-good n_j with $j \le x$ is then approximately

$$c_N \sum_{j\le x} (\beta j)^{-\delta(N)} \sim c_N \frac{(\beta x)^{1-\delta(N)}}{(1-\delta(N))\beta}.$$

The result that $2k/(2m-3)$ is a convergent p_j/q_j of $\log 2$ with j even and the result of Lévy [**17**] that, for a generic $\xi \in [0,1]$ that is not a rational

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

leads us to expect that the sequence m_j of potential solutions (k_j, m_j) to (1) coming from this result, is of exponential growth. Thus, of the potential solutions (m_j, k_j), with $j \le x$, one expects about $x^{1-\delta(N_2)}$, that is, roughly $x^{0.96}$, to be N_2-good. For those, (31) holds with $m = m_j$. Thus, if there would be, say, 10^{10} potential solutions with $m \le 10^{10^{11}}$, then one expects roughly $3 \cdot 10^9$ to be N_2-good, and those can be excluded by Corollary 2.

Remark 3. Given positive integers a, b, c, d, the density of primes $p \equiv c \pmod{d}$ such that $p \mid \{a^e + b^e\}_{e=0}^{\infty}$ is known, see Moree and

Sury [**28**]. Since $S_2 = \{2 \cdot 4^e + 1\}_{e=0}^{\infty}$, that result cannot be applied to establish Lemma 5.

Proof of Theorem 6. The idea of the proof is to show that, if, for every prime divisor $p > 3$ of $m+2$, at least one the conditions 1 or 2 is not satisfied, then the identity

$$\sum_{p|m+2} \frac{1}{p} + \frac{3}{m+2} \equiv 0 \pmod{1} \tag{32}$$

holds. Using this, we then show that m is bigger than the bound in the theorem; this is a contradiction. As usual, we make heavy use of the fact that k must be even.

We start with the equation

$$2^{k+1} + 1 \equiv - \sum_{\substack{p-1|k \\ p|m+2}} \frac{(m+2)}{p} \pmod{m+2}, \tag{33}$$

found on noting that $S_k(m+2) = 2m^k + (m+1)^k \equiv 2^{k+1} + 1 \pmod{m+2}$ and on invoking Theorem 3. Suppose that $p \mid m+2$. The idea is to reduce (33) modulo p (except if $p = 3$, then we reduce modulo 9).

If $p = 3$, then using $6 \mid k$, we see that $2^{k+1} + 1 \equiv 3 \pmod{9}$, and we infer that $3^2 \parallel m+2$, that is, we must have $m \equiv 7, 16 \pmod{27}$. Next assume $p > 3$.

First assume that $\mathrm{ord}_2(p) \not\equiv 2 \pmod{4}$. Then p does not divide S_2. Thus, the right hand side of (33) is non-zero modulo p, and this implies that $p-1 \mid k$ and $p^2 \nmid m+2$, and hence $2^{k+1} + 1 \equiv 3 \pmod{p}$.

Next, assume that $\mathrm{ord}_2(p) \equiv 2 \pmod{4}$, and condition 1 is not satisfied. Then $\mathrm{ord}_2(p)$ and N have an odd prime factor in common, and by (30) (with $e = k/2$) we get a contradiction to the assumption $N \mid k$.

Finally, assuming that condition 1 is satisfied but not condition 2, the right hand side of (33) is non-zero modulo p, and the same conclusion as before holds. By the Chinese remainder theorem we then infer that $2^{k+1} + 1 \equiv 3 \pmod{m+2}$, and hence from (33) we see that (32) holds.

Put $M_3 = (m^2-1)(4m^2-1)(m+2)$. By part 2 of Theorem 2 we infer that, amongst the numbers $m-1$, $m+1$, $m+2$, $2m-1$, $2m+1$, no prime $p \geq 7$ occurs more than once as divisor, the prime 2 occurs precisely twice, the prime 3 at most three times and the prime 5 at most two times. Using this, we obtain, on adding Moser's equations (9), (14), (17) and (19) to (32):

$$\sum_{p \mid M_3} \frac{1}{p} + \frac{1}{m-1} + \frac{2}{m+1} + \frac{2}{2m-1} + \frac{4}{2m+1} + \frac{3}{m+2} \geq \frac{109}{30}, \tag{34}$$

where

$$\frac{109}{30} = 5 - \frac{1}{2} - \frac{2}{3} - \frac{1}{5} = 3.633\,333\,333\,333\ldots.$$

Using the estimate

$$\sum_{p \leq x} \frac{1}{p} < \log\log x + 0.2615 + \frac{1}{\log^2 x} \quad \text{for } x > 1,$$

due to Rosser and Schoenfeld [**35**, (3.20)], we find that $\sum_{p \leq \beta} 1/p < 3.633\,32$ with $\beta = 4.33 \cdot 10^{12}$. From another paper by the same authors [**36**] we have

$$|\theta(x) - x| < \frac{x}{40 \log x}, \quad x \geq 678\,407.$$

Hence,

$$\log(4m^5) > \log(N_3) > \log \prod_{p \leq \beta} p = \theta(\beta) > 0.999\beta,$$

from which we infer that $m \geq 10^{10^{11}}$.

In the case where $m \equiv 2 \pmod 3$ there are precisely two of the five terms $m-1$, $m+1$, $2m-1$, $2m+1$ and $m+2$ divisible by 3, and in (34) we can replace $109/30$ by $109/30 + 1/3 = 119/30 = 3.966\,666\ldots$. In that case we can take $\beta = 4.425 \cdot 10^{17}$, and this leads to $m \geq 10^{10^{16}}$.

The smallest two primes in $P(N_2)$ are $2\,027$ and $2\,099$. For $p = 2\,027$, we can actually show that condition 2 is not satisfied. To this end, we must show that $k \not\equiv 1\,012 \pmod{2\,026}$. Computation shows that $(1\,012, 6\,079)$, $(3\,038, 6\,079)$ and $(5\,064, 6\,079)$ are helpful pairs. By

Lemma 4, it then follows that $k \not\equiv 1\,012 \pmod{2\,026}$. The smallest prime that possibly satisfies both condition 1 and 2 is hence 2 099. $\square$

Remark 4. We leave it as an exercise to the reader to show that (31) can be refined to

$$m \geq \begin{cases} 10^{10^{20}} & \text{if } 3 \nmid m+2;\ 5 \nmid m+2 \\ 10^{10^{16}} & \text{if } 3 \nmid m+2;\ 5 \mid m+2 \\ 10^{10^{14}} & \text{if } 3 \mid m+2;\ 5 \nmid m+2 \\ 10^{10^{11}} & \text{if } 3 \mid m+2;\ 5 \mid m+2. \end{cases} \tag{35}$$

In the same vein, one can show that, if (m, k) satisfies (1), $k \geq 2$ and $m \equiv \pm 1 \pmod{15}$ or $m \equiv \pm 1 \pmod{21}$, then $m \geq 10^{10^{20}}$. If, e.g., $m \equiv 1 \pmod{15}$, then the sum in the left hand side of (9) exceeds 1, so it must be at least two. We infer that (20) holds with $3.1666\ldots$ replaced by $4.1666\ldots$. This then leads to $m \geq 10^{10^{20}}$. The remaining cases are similar (they all lead to (20) with $3.1666\ldots$ replaced by $4.1666\ldots$).

Remark 5. Using the methods from Bach et al. [**1**], it should be possible to compute the largest β such that $\sum_{p \leq \beta} 1/p < 109/30$, respectively, 119/30 exactly. They found, e.g., that the prime $p_0 = 180\,124\,123\,005\,660\,046\,7$ is the largest one such that $\sum_{p \leq p_0} 1/p < 4$.

8. The generalized Erdős-Moser conjecture. The Erdős-Moser conjecture has the following generalization.

Conjecture 1. *There are no integer solutions* (m, k, a) *of*

$$1^k + 2^k + \cdots + (m-1)^k = am^k \tag{36}$$

with $k \geq 2$*,* $m \geq 2$ *and* $a \geq 1$*.*

In this direction the author proved in 1996 [**21**] that (36) has no integer solutions (a, m, k) with $k > 1$ and $m < \max(10^{10^6}, a \cdot 10^{22})$. With the hindsight of more than 10 years this can be improved.

Theorem 7. *Equation* (36) *has no integer solutions* (a, m, k) *with*

$$k \geq 2, \qquad m < \max\left(10^{9\cdot 10^6}, a \cdot 10^{28}\right).$$

Proof. (In this proof references to propositions and lemmas are exclusively to those in [**21**].) The Moser method yields that $2 \mid k$ and gives the following four inequalities:

$$\sum_{\substack{p-1|k \\ p|m-1}} \frac{1}{p} + \frac{a}{m-1} \geq 1, \qquad \sum_{\substack{p-1|k \\ p|m+1}} \frac{1}{p} + \frac{a+1}{m+1} \geq 1. \tag{37}$$

$$\sum_{\substack{p-1|k \\ p|2m-1}} \frac{1}{p} + \frac{2a}{2m-1} \geq 1, \qquad \sum_{\substack{p-1|k \\ p|2m+1}} \frac{1}{p} + \frac{2(a+1)}{2m+1} \geq 1. \tag{38}$$

Since $p \mid m$ implies $p - 1 \nmid k$ (Proposition 9), we infer that $(6, m) = 1$. Using this, we see that $M_1 = (m^2 - 1)(4m^2 - 1)/12$ is an even integer. Since no prime > 3 can divide more than one of the numbers $m - 1$, $m+1$, $2m-1$ and $2m+1$, and since 2 and 3 divide two of these numbers, we find on adding the inequalities that

$$\sum_{\substack{p-1|k \\ p|M_1}} \frac{1}{p} + \frac{a}{m-1} + \frac{a+1}{m+1} + \frac{2a}{2m-1} + \frac{2(a+1)}{2m+1} \geq 4 - \frac{1}{2} - \frac{1}{3} = 3\frac{1}{6}.$$

Using that $a(k + 1) < m < (a + 1)(k + 1)$ (Proposition 2), we see that in the latter equation the four terms involving a are bounded above by $6/(k + 1)$. Since $k \geq 10^{22}$ (Lemma 2), we can proceed as in the proof of Theorem 2 and find the same bound for m, namely, $m > 1.485 \cdot 10^{9321155}$.

Earlier it was shown that, if $k > 1$, then $k \geq 10^{22}$. To this end, Proposition 6 with $C = 3.16$, $s = 664\,579 = \pi(10^7)$ and n the 200th highly composite number was applied. Instead, we apply it with $C = 19/6 - 10^{-10}$, $s = 4\,990\,906$ and n the 259th composite number c_{250} (this has the property that the number of divisors of $c_{259} < s$, whereas the number of divisors of c_{260} exceeds s). Since $n = c_{259} > 5.583\,4 \cdot 10^{27}$, it follows that $k \geq 2n > 10^{28}$. Since $m > a(k + 1)$, the proof is completed. $\square$

Remark 6. The above proof shows that if (36) has a solution with $k \geq 2$, $m \geq 2$ and $a \geq 1$, then m must be odd. An easy reproof of this was given by MacMillan and Sondow [**19**].

Remark 7. The reader might wonder whether the method of Gallot et al. can be applied here as well to break the 10^{10^7} barrier. For a fixed integer a, this is possible if one manages to establish that $N \mid k$ with N large enough. Gallot et al. showed that $2k/(2m-2a-1)$ is a convergent with even index of $\log(1+1/a)$ for m large enough. For a given a, this can be made effective, establishing that $N \mid k$ along the lines of Section 5 is not always possible (see the last paragraph of that section).

Challenge. Reach the benchmark 10^{10^7} in Theorem 7.

9. The Kellner-Erdős-Moser conjecture. Kellner [**16**] conjectured that, if k, m are positive integers with $m \geq 3$, the ratio $S_k(m+1)/S_k(m)$ is an integer if and only if $(k, m) \in \{(1,3), (3,3)\}$. Noting that $S_k(m+1) = S_k(m) + m^k$, one easily observes that this conjecture is equivalent with the following one.

Conjecture 2. *We have* $aS_k(m) = m^k$ *if and only if* $(a, k, m) \in \{(1,1,3), (3,3,3)\}$.

If this conjecture holds true, then obviously so does the Erdős-Moser conjecture.

It is easy to deal with the case $m = 3$. Then we must have $a(1+2^k) = 3^k$, and hence $a = 3^e$ for some $e \leq k$. It follows that $1 + 2^k = 3^{k-e}$. This Diophantine equation was already solved by the famous medieval astronomer Levi ben Gerson (1288–1344), alias Leo Hebraeus, who showed that 8 and 9 are the only consecutive integers in the sequence of powers of 2 and 3, see Ribenboim [**34**, pages 124–125]. This leads to the solutions $(e, k) \in \{(0,1), (3,1)\}$, and hence $(a, k, m) \in \{(1,1,3), (3,3,3)\}$. Next assume that $m \geq 4$ and k is odd. Then, by Theorem 3, we find that $m(m-1)/2$ divides m^k, which is impossible. We infer that, to establish Conjecture 2, it is enough to establish Conjecture 3, where

$$\mathcal{A} = \{a \geq 1 : aS_k(m) = m^k \text{ has a solution with } 2 \mid k,\ k \geq 2,\ m \geq 4\}.$$

Conjecture 3. *The set $\mathcal{A}$ is empty.*

The next result shows that, if $a \equiv 2 \pmod 4$ or $a \equiv 3, 6 \pmod 9$, then $a \notin \mathcal{A}$.

Theorem 8. *Let $k \geq 2$ be even. Suppose that $q \mid a$ is a prime such that $q^2 \nmid a$ and $q - 1 \mid k$. Then $aS_k(m) \neq m^k$, and hence $a \notin \mathcal{A}$.*

Proof. Suppose that $aS_k(m) = m^k$. Let $q^e \parallel m$. Note that $e \geq 1$. Using Theorem 3, we find that $S_k(m) \equiv (m/q^e)S_k(q^e) \equiv -(m/q) \pmod{q^e}$. Now we consider the identity $aS_k(m) = m^k$ modulo q^{e+1} and find $-a(m/q) \equiv m^k \equiv 0 \pmod{q^{e+1}}$, contradicting $q^{e+1} \parallel am$. It follows that $aS_k(m) \neq m^k$. $\square$

Note that, if $a \notin \mathcal{A}$, then the equation $aS_k(m) = m^k$ can be solved completely. The author is not aware of earlier 'naturally' occurring Erdős-Moser type equations that can be solved completely. He expects that further values of a can be excluded and might come back to this in a future publication.

Acknowledgments. Part of this article was written whilst I had four interns (Valentin Buciumas, Raluca Havarneanu, Necla Kayaalp and Muriel Lang) studying variants of the Erdős-Moser equation. Raluca and Valentin found a (reparable) mistake in Moser's paper and Muriel showed that 2027 is the smallest prime in $P(N_2)$, but does not satisfy condition 2 of Theorem 6. I thank them all for their questions, comments and cheerful presence. Paul Tegelaar provided some helpful comments on an earlier version. Jonathan Sondow I thank for helpful e-mail correspondence. This note profited a lot from corrections by Julie Rowlett (a native English speaker). Particular thanks are due to the referee for many very detailed and constructive comments.

The academic year 1994/1995 the author spent as a postdoc of Alf van der Poorten at Macquarie. Alf told me various times it would be

so nice if mathematicians could be less serious in their mathematical presentation, e.g., talk about a 'troublesome double sum,' if there are smooth numbers, then also consider hairy numbers, etc. In this spirit, I 'spiced up' my initial submission of [**21**]; the red pencil of the referee was harsh, though, but somehow the words 'mathemagics' and 'rabbits' survived. Rabbits are difficult to suppress, and least of all mathemagical ones. So I am happy they are back full force in the title of [**25**]. Also whilst at Macquarie, thanks to questions by then visitor Patrick Solé, I got into the study of divisors of $a^k + b^k$, not realizing there is a connection with the Erdős-Moser equation (as the present article shows).

Had Alf learned that the present record for solutions for EM is based on a continued fraction expansion (of $\log 2$), I am sure he would have been pleased.

I had a wonderful year in Australia and will be always grateful to Alf for having made that possible.

ENDNOTES

1. A large part of the material in Section 2 is copied verbatim from Moser's paper.

2. The Pascaline was originally developed for tax collecting purposes!

3. The proof given in Section 4 is implicit in Moree's [**21**] with $a = 1$.

REFERENCES

1. E. Bach, D. Klyve and J.P. Sorenson, *Computing prime harmonic sums*, Math. Comp. **78** (2009), 2283–2305.

2. C. Ballot, *Density of prime divisors of linear recurrences*, Mem. Amer. Math. Soc. **115** (1995), viii+102 pp.

3. M. Bennett, K. Győry and A. Pintér, *On the Diophantine equation* $1^k+2^k+\cdots+x^k=y^n$, Compos. Math. **140** (2004), 1417–1431.

4. M.R. Best and H.J.J. te Riele, *On a conjecture of Erdős concerning sums of powers of integers*, Report NW 23/76, Math. Centr., Amsterdam, 1976.

5. L. Brenton and A. Vasiliu, *Znam's problem*, Math. Mag. **75** (2002), 3–11.

6. W. Butske, L.M. Jaje and D.R. Mayernik, *On the equation* $\sum_{p|N} \frac{1}{p} + \frac{1}{N} = 1$, *pseudoperfect numbers, and perfectly weighted graphs*, Math. Comp. **69** (2000), 407–420.

7. L. Carlitz, *The Staudt-Clausen theorem*, Math. Mag. **34** (1960/1961), 131–146.

8. L. Dicuangco, P. Moree and P. Solé, *The lengths of Hermitian self-dual extended duadic codes*, J. Pure Appl. Alg. **209** (2007), 223–237.

9. B. Fein, B. Gordon and J.H. Smith, *On the representation of* -1 *as a sum of two squares in an algebraic number field*, J. Number Theory **3** (1971), 310–315.

10. S. Finch, G. Martin and P. Sebah, *Roots of unity and nullity modulo* n, Proc. Amer. Math. Soc. **138** (2010), 2729–2743.

11. Y. Gallot, P. Moree and W. Zudilin, *The Erdős-Moser equation* $1^k + 2^k + \cdots + (m-1)^k = m^k$ *revisited using continued fractions*, Math. Comp. **80** (2011), 1221–1237.

12. R.K. Guy, *Unsolved problems in number theory*, Third edition, *Problem books in mathematics*, Springer-Verlag, New York, 2004.

13. Y. Jia, S. Ling and C. Xing, *On self-dual cyclic codes over finite fields*, IEEE Trans. Inf. Theor. **57** (2011), 2243–2251.

14. B.C. Kellner, *Über irreguläre Paare höhere Ordnungen*, Diplomarbeit, Mathematisches Institut der Georg, August, University of Gottingen, Germany, 2002. (Also available at `http://www.bernoulli.org/~bk/irrpairord.pdf`).

15. ———, *The equivalence of Giuga's and Agoh's conjectures*, arXiv:math.NT/0409259, preprint.

16. ———, *On stronger conjectures that imply the Erdős-Moser conjecture*, J. Number Theory **131** (2011), 1054–1061.

17. P. Lévy, *Sur le développement en fraction continue d'un nombre choisi au hasard*, Compos. Math. **3** (1936), 286–303.

18. K. MacMillan and J. Sondow, *Proofs of power sum and binomial coefficient congruences via Pascal's identity*, Amer. Math. Month. **118** (2011), 549–551.

19. ———, *Divisibility of power sums and the generalized Erdős-Moser equation*, Elem. Math. **67** (2012), 182–186.

20. P. Moree, *On a theorem of Carlitz-von Staudt*, C.R. Math. Rep. Acad. Sci. Canada **16** (1994), 166–170.

21. ———, *Diophantine equations of Erdős-Moser type*, Bull. Austral. Math. Soc. **53** (1996), 281–292.

22. ———, *On the divisors of* $a^k + b^k$, Acta Arith. **80** (1997), 197–212.

23. ———, *On primes* p *for which* d *divides* $\mathrm{ord}_p(g)$, Funct. Approx. Comment. Math. **33** (2005), 85–95.

24. ———, *Artin's primitive root conjecture–A survey*, Integers **12** (2012), 1305–1416.

25. ———, *A top hat for Moser's four mathemagical rabbits*, Amer. Math. Month. **118** (2011), 364–370.

26. P. Moree, H. te Riele and J. Urbanowicz, *Divisibility properties of integers* x, k *satisfying* $1^k + \cdots + (x-1)^k = x^k$, Report NM-R9215, Centr. Wisk. Inform., Amsterdam, August 1992. (Available from the authors upon request.)

27. ———, *Divisibility properties of integers* x, k *satisfying* $1^k + \cdots + (x-1)^k = x^k$, Math. Comp. **63** (1994), 799–815.

28. P. Moree and B. Sury, *Primes in a prescribed arithmetic progression dividing the sequence* $\{a^k + b^k\}_{k=1}^{\infty}$, Int. J. Number Theory **5** (2009), 641–665.

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.

30. ———, `http://www-history.mcs.st-andrews.ac.uk/Biographies/Moser_Leo.html`.

31. W.-G. Nowak, *On an arithmetic function connected with the distribution of supersingular Fermat varieties*, Unif. Distr. Theor. **2** (2007), 11–21.

32. B. Pascal, *Sommation des puissances numériques*, in *Oeuvres complètes*, vol. III, Jean Mesnard, ed., Desclée-Brouwer, Paris, 1964.

33. M. Ram Murty, *Introduction to p-adic analytic number theory*, AMS/IP Stud. Adv. Math. **27**, American Mathematical Society, Providence, RI, 2002.

34. P. Ribenboim, *Catalan's conjecture. Are* 88 *and* 99 *the only consecutive powers?* Academic Press, Inc., Boston, MA, 1994.

35. J.B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illin. J. Math. **6** (1962), 64–94.

36. ———, *Sharper bounds for the Chebyshev functions* $\theta(x)$ *and* $\psi(x)$, Math. Comp. **29** (1975), 243–269.

37. T. Shioda and T. Katsura, *On Fermat varieties*, Tôhoku Math. J. **31** (1979), 97–115.

38. J. Sondow and K. MacMillan, *Reducing the Erdős-Moser equation* $1^n + 2^n + \ldots + k^n = (k+1)^n$ *modulo* k *and* k^2, Integers **11** (2011), 8 pages.

39. J. Urbanowicz, *Remarks on the equation* $1^k + 2^k + \cdots + (x-1)^k = x^k$, Nederl. Akad. Weten. Indag. Math. **50** (1988), 343–348.

40. M. Wyman, *Biographical sketch–Leo Moser*, Rocky Mountain J. Math. **1** (1971), 255–256 (1 plate).

Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany
Email address: moree@mpim-bonn.mpg.de