# A TOP HAT FOR MOSER'S FOUR MATHEMAGICAL RABBITS

PIETER MOREE

ABSTRACT. If the equation $1^k + 2^k + \ldots + (m-2)^k + (m-1)^k = m^k$ has a solution with $k \geq 2$, then $m > 10^{10^6}$. Leo Moser showed this in 1953 by amazingly elementary methods. His proof rests on four identities he derives separately. It is shown here that Moser's result can be derived from a von Staudt-Clausen type theorem (a simple proof of which is also presented here). In this approach the four Moser identities can be derived uniformly. The mathematical arguments used in the proofs were already available during the lifetime of Lagrange (1736-1813).

## 1. INTRODUCTION

Consider the Diophantine equation

$$1^k + 2^k + \ldots + (m-2)^k + (m-1)^k = m^k, \tag{1}$$

to be solved in integers $(m, k)$ with $m \geq 2$ and $k \geq 1$. Note that in case $k = 1$ the left hand side of (1) equals $m(m-1)/2$ and this leads to the (unique) solution $1 + 2 = 3$. From now on we will assume that $k \geq 2$. Conjecturally solutions with $k \geq 2$ do not exist (this conjecture was formulated around 1950 by Paul Erdős in a letter to Leo Moser). Leo Moser [10] established the following theorem in 1953.

**Theorem 1.** (Leo Moser, 1953). *If $(m, k)$ is a solution of (1) with $k \geq 2$, then $m > 10^{10^6}$.*

His result has since then been improved upon. Butske et al. [2] have shown, by computing rather than estimating certain quantities in Moser's original proof, that $m > 1.485 \cdot 10^{9321155}$. By proceeding along these lines this bound cannot be improved upon substantially. Butske et al. [2, p. 411] expressed the hope that new insights will eventually make it possible to reach the more natural benchmark $10^{10^7}$. This hope was recently fulfilled by Gallot, the author, and Zudilin [4], who showed that $2k/(2m-3)$ must be a convergent of $\log 2$ and made an extensive continued fraction computation of $(\log 2)/2N$, with $N$ an appropriate integer in order to establish Theorem 2. Note that their result goes well beyond establishing the benchmark. Their approach only works for those $N$ for which it can be shown that $N|k$. In [9] it was e.g. shown that $\mathrm{lcm}(1, 2, \ldots, 200)|k$.

**Theorem 2.** *If $(m, k)$ is a solution of (1) with $k \geq 2$, then $m > 10^{10^9}$.*

Moser's proof of Theorem 1 is quite amazing in the sense that he only uses very elementary number theory. His proof is even mathemagical in the sense that he pulls four rabbits out of a hat, namely the Equations (7), (10), (12), and (13), that a solution $(m, k)$ has to satisfy. He derives each of these equations separately in a quite ingenious way. In this note we will see that a reproof of Moser's result can be given, showing that the following result is the top hat the four rabbits were pulled from.

Put $S_r(y) = \sum_{j=1}^{y-1} j^r$.

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

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

Using this result, an easy proof of which will be given in Section 3, a less mathemagical reproof of Moser's result can be given. For a polished version of Moser's orignal proof, we refer the reader to the extended version of this note [8].

The prime harmonic sum diverges (as Euler already knew) and so given $\alpha > 1/2$ there exists a largest prime $p(\alpha)$ such that $\sum_{p \le p(\alpha)} 1/p < \alpha$. Moser needed $p(3.16)$ in his proof, but could only estimate it using certain prime number estimates. His proof is easily adapted to involve $p(3\frac{1}{6})$ and this was first exactly computed by Butske et al. [2] leading to an improvement of Moser's bound, namely

$$m > \left( 3 \prod_{p \le p(3\frac{1}{6})} p \right)^{\frac{1}{4}} > 1.485 \cdot 10^{9321155}. \tag{3}$$

We obtain, using Theorem 3 and a computer algebra package like PARI to compute $p(3\frac{1}{6}) = 85861889$ and the prime product in (3), the following variant of Moser's result.

**Theorem 4.** *Suppose that $(m, k)$ is a solution of (1) with $k \ge 2$, then*
1) $m > 1.485 \cdot 10^{9321155}$;
2) $k$ *is even,* $m \equiv 3 \pmod{8}$, $m \equiv \pm 1 \pmod{3}$;
3) $m - 1$, $(m+1)/2$, $2m - 1$, *and* $2m + 1$ *are all squarefree;*
4) *If $p$ divides at least one of the above four integers, then $p - 1|k$;*
5) *The number $(m^2 - 1)(4m^2 - 1)/12$ is squarefree and has at least* 4990906 *prime factors.*

The proof we give in this note shows that if Lagrange (1736-1813) would have had a present day computer, he could have well proved Theorem 4.

In order to improve on Theorem 2 by Moser's approach one needs to find additional rabbit(s) in the top hat. The interested reader is wished good luck in finding these elusive animals !

## 2. PROOF OF THEOREM 4

*Proof of Theorem* 4. We will apply Theorem 3 with $r = k$.

In case $k$ is odd, we find on combining (2) (putting $y = m$) with (1) and using

the coprimality of $m$ and $m-1$, that $m=2$ or $m=3$, but these cases are easily excluded. (Since $1^k + 2^k < (1+2)^k$ for $k>1$, one sees that $1+2^k = 3^k$ has only the solution $k=1$.) Therefore $k$ must be even.

Take $y=m-1$. Then using (1) the left hand side of (2) simplifies to

$$S_k(m-1) = 1^k + 2^k + \cdots + (m-2)^k = m^k - (m-1)^k \equiv 1 \ (\text{mod } m-1). \quad (4)$$

We get from (2) and (4) that

$$\sum_{p|m-1, \ p-1|k} \frac{m-1}{p} + 1 \equiv 0 \ (\text{mod } m-1). \quad (5)$$

Suppose there exists $p|m-1$ such that $p-1 \nmid k$. Then on reducing both sides modulo $p$ we get $1 \equiv 0 \ (\text{mod } p)$. This contradiction shows that in (5) the condition $p-1|k$ can be dropped and thus we obtain

$$\sum_{p|m-1} \frac{m-1}{p} + 1 \equiv 0 \ (\text{mod } m-1), \quad (6)$$

Suppose there exists a prime $p$ dividing $m-1$ such that also $p^2$ divides $m-1$. Then on reducing both sides modulo $p^2$, we get $1 \equiv 0 \ (\text{mod } p^2)$. This contradiction shows that $m-1$ must be squarefree. On dividing (6) by $m-1$ we obtain

$$\sum_{p|m-1} \frac{1}{p} + \frac{1}{m-1} \equiv 0 \ (\text{mod } 1). \quad (7)$$

Take $y=m$. Then using (1) and $2|k$ we infer from (2) that

$$\sum_{p-1|k, \ p|m} \frac{1}{p} \equiv 0 \ (\text{mod } 1). \quad (8)$$

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

Take $y=m+1$. Then using (1) and the fact that $k$ is even, the left hand side of (2) simplifies to

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

We obtain

$$\sum_{p|m+1, \ p-1|k} \frac{m+1}{p} + 2 \equiv 0 \ (\text{mod } m+1), \quad (9)$$

but by reasoning as in the case $y=m-1$, it is seen that $p|m+1$ implies $p-1|k$ and thus

$$\sum_{p|m+1} \frac{1}{p} + \frac{2}{m+1} \equiv 0 \ (\text{mod } 1). \quad (10)$$

From (9) and $m \equiv 3 \pmod 8$, we derive that $(m+1)/2$ is squarefree.

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

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

we infer that

$$\sum_{p|2m-1, \ p-1|k} \frac{2m-1}{p} + 2m^k \equiv 0 \pmod{2m-1}. \tag{11}$$

Since $m$ and $2m-1$ are coprime we infer that if $p|2m-1$, then $p-1|k$, $m^k \equiv 1 \pmod p$ and furthermore that $2m-1$ is squarefree. By the Chinese Remainder Theorem it then follows that $2m^k \equiv 2 \pmod{2m-1}$ and hence from (11) we obtain

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

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

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

and proceeding as in case $y = 2m - 1$ we obtain

$$\sum_{p|2m+1} \frac{1}{p} + \frac{4}{2m+1} \equiv 0 \pmod 1. \tag{13}$$

We further see that $2m + 1$ is squarefree.

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 $M = (m-1)(m+1)(2m-1)(2m+1)/12$ is a squarefree integer. On adding (7), (10), (12), and (13), we deduce that

$$\sum_{p|M} \frac{1}{p} + \frac{1}{m-1} + \frac{2}{m+1} + \frac{2}{2m-1} + \frac{4}{2m+1} \geq 4 - \frac{1}{2} - \frac{1}{3} = 3\frac{1}{6}. \tag{14}$$

One checks that the only solutions of (7) with $m \leq 1000$ are $3, 7$, and $43$. These are easily ruled out by (10). Thus (14) yields (with $\alpha = 3.16$) $\sum_{p|M} \frac{1}{p} > \alpha$. From this it follows that if

$$\sum_{p \leq x} \frac{1}{p} < \alpha, \tag{15}$$

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

$$m > 3^{1/4} e^{\theta(x)/4}, \tag{16}$$

with $\theta(x) = \sum_{p \leq x} \log p$, the Chebyshev $\theta$-function. Since for example (15) is satisfied with $x = 1000$, we find that $m > 10^{103}$ and infer from (14) that we can take $\alpha = 3\frac{1}{6} - 10^{-100}$ in (15). Next one computes (using a computer algebra package) the

largest prime $p_k$ such that $\sum_{p_j \leq p_k} \frac{1}{p_j} < 3\frac{1}{6}$, with $p_1, p_2, \ldots$ the consecutive primes (note that $p_k = p(3\frac{1}{6})$). Here one finds that $k = 4990906$ and

$$\sum_{i=1}^{4990906} \frac{1}{p_i} = 3.1666666588101728584 < 3\frac{1}{6} - 10^{-9}.$$

By direct computation one finds that $\theta(p_k) = 8.58510010694053 \cdots * 10^7$. Using this we infer from (16), the inequality (3) and hence part 1 of the theorem is proved.

Notice that along our way towards proving part 1, the remaining parts of the theorem have also been proved. $\qquad\square$

## 3. Proof of the Carlitz-von Staudt theorem

Carlitz [3] gave a proof of Theorem 3 using finite differences and states that the result is due to von Staudt. In the case $r$ 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 [6] gave a reproof using the theory of primitive roots and Kellner [5] a reproof (in case $r$ even only) using Stirling numbers of the second kind. Here a reproof will be given that is easier than all the above. It only uses the following result of Lagrange.

**Theorem 5.** *If $f$ is a one-variable polynomial of degree $n$ over $\mathbb{Z}/p\mathbb{Z}$, then it cannot have more than $n$ roots unless it is identically zero.*

*Proof.* See, e.g. the book of Rose [11, Theorem 2.2]. $\qquad\square$

**Lemma 1.** *Suppose that $p-1 \nmid r$. Then the equation $x^r \not\equiv 1 \pmod{p}$ has a solution.*

*Proof.* Let $r_1 \geq 0$ be the smallest positive integer such that $r_1 \equiv r \pmod{p-1}$. Then $r_1 < p - 1$. Suppose that $x^r \equiv 1 \pmod{p}$ for every $x \in \{1, 2, \ldots, p - 1\}$, then by Fermat's Little Theorem we also have $x^{r_1} \equiv 1 \pmod{p}$ for every $x \in \{1, 2, \ldots, p-1\}$, contradicting Lagrange's Theorem. $\qquad\square$

**Lemma 2.** *Let $p$ be a prime. We have*
$$S_r(p) \equiv \epsilon_r(p) \pmod{p},$$
*where*
$$\epsilon_r(p) = \begin{cases} -1 & \text{if } p - 1 | r; \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* If $p - 1$ divides $r$ the result follows by Fermat's Little Theorem. If $p - 1 \nmid r$, assume that $S_r(p) \not\equiv 0 \pmod{p}$. Let $a$ be an integer not divisible by $p$. Multiplication by $a$ permutes the elements of $\mathbb{Z}/p\mathbb{Z}$ and hence $S_r(p) \equiv a^r S_r(p) \pmod{p}$, from which we infer that $a^r \equiv 1 \pmod{p}$. We infer that $a^r \equiv 1 \pmod{p}$ for $a = 1, 2, \ldots, p - 1$. Invoking Lemma 1 gives a contradiction and hence our assumption that $S_r(p) \not\equiv 0 \pmod{p}$ must have been false. $\qquad\square$

The usual proof of this result makes use of the existence of a primitive root modulo $p$, which provides a solution to $x^r \not\equiv 1 \pmod{p}$ in case $p - 1 \nmid r$. The proof given here only makes use of the more elementary theorem of Lagrange, Theorem 5.

**Lemma 3.** *In case $p$ is odd and in case $p = 2$ and $r$ is even, we have $S_r(p^{\lambda+1}) \equiv pS_r(p^\lambda) \pmod{p^{\lambda+1}}$.*

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

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

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. $\qquad\square$

*Proof of Theorem 3.* First let us consider the case where $r$ is odd. Assume $S_r(m)$ is a multiple of $m(m-1)/2$. We need to show that $S_r(m+1) = S_r(m) + m^r$ is a multiple of $m(m+1)/2$.

If $m$ is even, we have that $m/2$ divides $S_r(m)$. But

$$S_r(m+1) = (1^r + m^r) + (2^r + (m-1)^r) + \ldots + ((\frac{m}{2})^r + (\frac{m}{2}+1)^r),$$

which is a multiple of $m+1$ as each pair above is. Thus, $S_r(m+1)$ is a multiple of $m/2$ as well as of $m+1$, which are coprime and hence a multiple of $m(m+1)/2$.

If $m$ is odd, then $m|S_r(m)$. But

$$S_r(m+1) = (1^r + m^r) + (2^r + (m-1)^r) + \ldots + (\frac{m+1}{2})^r,$$

which is a multiple of $(m+1)/2$ as each pair as well as the last remaining term is. Thus $S_r(m+1)$ is a multiple of $m$ as well as $(m+1)/2$ which are coprime and hence a multiple of $m(m+1)/2$.

Next we consider the case where $r$ is even. Suppose that $p^f|y$, then

$$S_r(y) = \sum_{\alpha=0}^{\frac{y}{p^f}-1}\sum_{\beta=0}^{p^f-1}(\alpha p^f + \beta)^r \equiv \frac{y}{p^f}S_r(p^f) \pmod{p^f}. \qquad (17)$$

By the Chinese Remainder Theorem it is enough to show that

$$S_r(y) \equiv \frac{y}{p}\epsilon_r(p) \pmod{p^{e_p}},$$

where $y = \prod_p p^{e_p}$ is a factorization of $y$ into prime powers $p^{e_p}$. By (17), Lemma 3, and Lemma 2, we then infer that

$$S_r(y) \equiv \frac{y}{p^{e_p}}S_r(p^{e_p}) \equiv \frac{y}{p}S_r(p) \equiv \frac{y}{p}\epsilon_r(p) \pmod{p^{e_p}},$$

thus concluding the proof. $\qquad\square$

## 4. Concluding remarks

A further application of the Carlitz-von Staudt theorem is to show that Giuga's conjecture (1950) and Agoh's conjecture (1990) are equivalent, see Kellner [5]. Giuga's conjecture states that if $n \geq 2$, then $S_{n-1}(n) \equiv -1 \pmod{n}$ iff $n$ is prime. Agoh's conjecture states that if $n \geq 2$, then $nB_{n-1} \equiv -1 \pmod{n}$ iff $n$ is a prime, where $B_r$ denotes the $r$-th Bernoulli number.

The author has generalized the Carlitz-von Staudt theorem to deal with consecutive $r$-th powers in arithmetic progression, see [6]. The method of proof in case $r$ odd given in Section 3 does not generalize then.

That Theorem 3 can be used to reprove Moser's result is a discovery due to the author and first used in [7], where it was used to study the equation $1^k + 2^k + \cdots + (m-1)^k = am^k$. The presentation given here also draws on computer improvements since 1996 and [2]. The proof of Theorem 3 given here is clearly easier than those given in [3, 5, 6], and is the main new contribution in this paper.

Some variants of the Erdős-Moser problem require computing $p(\alpha)$ for $\alpha > 3\frac{1}{6}$, see e.g. [7]. The largest value for which $p(\alpha)$ has been computed is $\alpha = 4$. Bach et al. [1] found that $p(4) = 1801241230056600467$, but whereas the computation of $p(3\frac{1}{6})$ is straightforward with a computer algebra package, computation of $p(4)$ is rather more involved (using the Meissel-Lehmer algorithm). For $\alpha > 4$ one presently has to resort to deriving a sharp lower bound for $p(\alpha)$ and here, as was Moser, one is forced to use prime number estimates, cf. [7].

## References

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

[2] 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.

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

[4] 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, arXiv:0907.1356, 2009, to appear in Mathematics of Computation.

[5] B.C. Kellner, The equivalence of Giuga's and Agoh's conjectures, Preprint: arXiv:math.NT/0409259, 2004.

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

[7] P. Moree, Diophantine equations of Erdős-Moser type, *Bull. Austral. Math. Soc.* **53** (1996), 281–292.

[8] P. Moree, Moser's mathemagical work on the equation $1^k + 2^k + \cdots + (m-1)^k = m^k$, MPIM-preprint 2009-92, http://www.mpim-bonn.mpg.de/preprints/retrieve, 2009.

[9] 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.

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

[11] H.E. Rose, *A course in number theory*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1988.

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