Periodica Mathematica Hungarica
https://doi.org/10.1007/s10998-023-00551-4

Asymptotics on a class of $\mathcal{S}$-unit integers

Florian Luca[1,2] · **Pantelimon Stănică**[3]

Accepted: 3 April 2023

Abstract
In this paper we consider a (non)congruence generalizing the so-called good/bad numbers introduced by Moree (Acta Arith LXXX 3:197–212, 1997) and give asymptotics for their counting functions. In addition, we give heuristics for some conjectured bounds on primes belonging to such a class.

Keywords Congruences · Quadratic characters · $\mathcal{S}$-units · Sieve methods · Mersenne primes

Mathematics Subject Classification 11A07 · 11N13 · 11N36 · 11N69

1 Introduction

Carlitz [1] showed that all permutation polynomials over $\mathbb{F}_q$, where $q > 2$ is a power of a prime, are generated by the inverse permutation polynomials x^{q-2} and some affine functions $ax + b$ ($0 \neq a, b \in \mathbb{F}_q$). Building upon that result, in a recent work [2], concentrating on the inverse function, $x \mapsto x^{2^n-2}$ in $\mathbb{F}_{2^n}$ (the finite field of dimension n over the two-element prime field $\mathbb{F}_2$), the congruence $\frac{n-1}{2^{\nu_2(n-1)}} \equiv 2^s 3^k \pmod{2^n - 1}$ was used to show that for such an n, the inverse power permutation in $\mathbb{F}_{2^n}$ has a decomposition into affine, quadratic and cubic power permutations of length $k + s + \nu_2(n-1)$. Here $\nu_2(m)$ is the exponent of 2 in the factorization of a positive integer m. Decomposing permutation polynomials into power permutations of small weight is quite useful in reducing hardware needs in the implementation of symmetric cryptographic algorithms, with or without countermeasures to side channel attacks.

✉ Pantelimon Stănică
pstanica@nps.edu

Florian Luca
Florian.Luca@wits.ac.za

1 School of Mathematics, University of the Witwatersrand, Private Bag X3, Wits 2050, Johannesburg, South Africa

2 Centro de Ciencias Matemáticas, UNAM, Morelia, Mexico

3 Applied Mathematics Department, Naval Postgraduate School, Monterey, CA 93943, USA

If $k = 0$, above, the integers n are called *good*, and if the congruence does not happen, they are called *bad* (see [3] for estimates). Here, we let $\mathcal{S} = \{2, 3\}$ and define a positive integer n as $\mathcal{S}$-*bad* if

$$\frac{n-1}{2^{\nu_2(n-1)}} \not\equiv 2^s 3^k \pmod{2^n - 1}, \tag{1}$$

and if the congruence holds, we call them $\mathcal{S}$-*good*. We show that the $\mathcal{S}$-bad numbers form a set of asymptotic density 1. Thus, most positive integers are $\mathcal{S}$-bad. Surely, the concept and results can be further extended to any finite set of primes $\mathcal{S}$.

2 Our result

Let $\mathcal{A}$ be the set of positive integers $n \geq 1$ such that the congruence

$$\frac{n-1}{2^{\nu_2(n-1)}} \equiv 2^s 3^k \pmod{2^n - 1} \tag{2}$$

holds with some integers k and s. Let $\mathcal{A}(x) = \mathcal{A} \cap [1, x]$. For a real number x and a positive integer k we write $\log_k x$ for the kth iterate of the logarithm $\log x := \max\{\ln x, 1\}$, where $\ln x$ is the natural logarithm. We use the notations $O, o, \ll, \gg$ with their customary meanings. We let $\left(\frac{a}{p}\right)$ (or, $(a|p)$) denote the Legendre symbol of an integer a with respect to the odd prime p, which equals $1, -1$, if a is a nonzero quadratic residue, respectively, non-residue, modulo p, and it is 0 if p divides a. We will show the following estimate.

Theorem 1 *The estimate*

$$\#\mathcal{A}(x) \ll \frac{x}{(\log_2 x)^{1+o(1)}}$$

holds as $x \to \infty$.

We will point out later that there are, surely, infinitely many good integers. However, we conjecture that the sum of reciprocals of the good integers is convergent, but our estimate is not strong enough to infer such a result.

3 The proof of theorem 1

Let x be a large positive real number and $n \leq x$. Write $n = 2^a m + 1$, where $a \geq 0$ and m is odd. Let $y := \lfloor 2 \log_3 x \rfloor$. The set $\mathcal{A}_1(x)$ of positive integers $n \leq x$ such that $a \geq y$ is the set of positive integers $n \leq x$ such that $n \equiv 1 \pmod{2^y}$. The number of such n is

$$\#\mathcal{A}_1(x) \ll \left\lfloor \frac{x}{2^y} \right\rfloor + 1 \leq \frac{2x}{2^y} = O\left(\frac{x}{(\log_2 x)^{2\ln 2}}\right) = o\left(\frac{x}{(\log_2 x)^{1.1}}\right)$$

as $x \to \infty$. From now on we assume that $n \notin \mathcal{A}_1(x)$. We let $I = [2, t]$, where

$$t := \frac{(\log x)^{1/2}}{(\log_2 x)^{1/4}}.$$

Let $\varepsilon > 0$ and let $\mathcal{A}_2(x)$ be the set of $n \leq x$ having less than $L := \lfloor \varepsilon \log_3 x \rfloor$ prime factors $q \leq t$. We estimate the cardinality of $\mathcal{A}_2(x)$. Let $L' < L$ and $q_1 < \cdots < q_{L'}$ be L' prime numbers in $[2, t]$. Let d be a positive integer built up only of $q_1, \ldots, q_{L'}$. We count the

positive integers $n \le x$ such that d is the part of n consisting of primes $q \le t$. If d has a prime power divisor of exponent at least y, then there is a prime $q \le t$ such that $q^y \mid n$. As in the case of $\mathcal{A}_1(x)$, for fixed q the number of such $n \le x$ is

$$\ll \frac{x}{q^y} \ll \frac{x}{(\log_2 x)^{1.1}}.$$

This we multiply with a factor of L to account for all the possible q's and get

$$\#\mathcal{A}_{2,1}(x) \ll \frac{xL}{(\log_2 x)^{1.1}} = o\left(\frac{x}{\log_2 x}\right).$$

Here, $\mathcal{A}_{2,1}(x)$ stands for the subset of $\mathcal{A}_2(x)$ of n's such that d is divisible by a prime power of exponent at least y. Assume now that the exponent of each q in d is at most y and let $\mathcal{A}_{2,2}(x)$ be the set of such n's. Then

$$d \le t^{yL} = (\log x)^{O((\log_3 x)^2)} = x^{o(1)}$$

as $x \to \infty$. Writing $n = d\ell$, we need to count $\ell \le x/d$ free of primes $q \le t$. The number of such integers ℓ is

$$\frac{x}{d} \prod_{q \le t} \left(1 - \frac{1}{q}\right) \ll \frac{x}{d \log t} \ll \frac{x}{d \log_2 x}.$$

We now sum up over d's getting a sum of

$$\begin{aligned}
\#\mathcal{A}_{2,2}(x) &\ll \frac{x}{\log_2 x} \sum_{\substack{q \mid d \Rightarrow q \le t \\ \omega(d) \le L}} \frac{1}{d} \\
&\ll \frac{x}{\log_2 x} \sum_{k \le L} \frac{1}{k!} \left(\sum_{q \le t} \sum_{u \ge 1} \frac{1}{q^u} \right)^k \\
&\ll_\varepsilon \frac{x}{\log_2 x} \frac{1}{L!} \left(\log_3 x + O(1)\right)^L \\
&\ll_\varepsilon \frac{x}{\log_2 x} \left(\frac{e \log_3 x + O(1)}{\varepsilon \log_3 x} \right)^{\varepsilon \log_3 x} \\
&\ll \frac{x}{(\log_2 x)^{1 - \varepsilon \log(e/\varepsilon) + o(1)}}
\end{aligned}$$

as $x \to \infty$. Assume $n \notin \mathcal{A}_1(x) \cup \mathcal{A}_2(x)$. Then n has at least $\lfloor \varepsilon \log_3 x \rfloor$ prime factors $q \le t$. Let $\mathcal{A}_3(x)$ be the set of such $n \le x$ such that at least $L_1 := \lfloor 0.5\varepsilon \log_3 x \rfloor$ of such primes are in fact smaller than or equal to $\log_2 x$. Let $q_1 < \cdots < q_{L_1} \le \log_2 x$ be L_1 such prime factors of n. The number of $n \le x$ divisible by $q_1 \dots q_{L_1}$ is at most

$$\frac{x}{q_1 \cdots q_{L_1}}.$$

Summing up over all possibilities of $q_1 < \cdots < q_{L_1}$ we get a count of

$$\begin{aligned}
\#\mathcal{A}_3(x) &\ll x \sum_{q_1 < \cdots < q_{L_1} \le \log_2 x} \frac{1}{q_1 \cdots q_{L_1}} \\
&\ll \frac{x}{L_1!} \left(\sum_{q \le \log_2 x} \frac{1}{q} \right)^{L_1}
\end{aligned}$$

$$\begin{aligned}
&\ll \frac{x}{L_1!}\left(\log_4 x + O(1)\right)^{L_1}\\
&\ll x\left(\frac{e\log_4 x + O(1)}{0.5\varepsilon \log_3 x}\right)^{0.5\varepsilon \log_3 x}\\
&= \frac{x}{(\log_2 x)^{O(\varepsilon \log_4 x)}}\\
&= o\left(\frac{x}{\log_2 x}\right).
\end{aligned}$$

From now on, $n \notin \mathcal{A}_1(x) \cup \mathcal{A}_2(x) \cup \mathcal{A}_3(x)$. Such $n \le x$ is divisible by $K = \lfloor (\log_3 x)^{2/3} \rfloor$ primes in $[\log_2 x, t]$. Let $q_1 < \cdots < q_K$ be primes, all in $[\log_2 x, t]$, and assume $q_1, \ldots, q_K$ are all divisors of n. For $1 \le i < j \le K$ let $p_{i,j}$ be a primitive prime factor of $2^{q_i q_j} - 1$. As $i < j$ vary, $p_{i,j}$ vary as well so they are all distinct. So are the q_is. Let $\mathcal{A}_4(x)$ be the set of $n \le x$ for which $q_1, \ldots, q_K, p_{1,2}, \ldots, p_{K-1,K}$ are not all distinct. Then there is q_k which is a primitive prime factor of $2^{q_i q_j} - 1$ for some $i < j$ both in $\{1, \ldots, K\}$. Thus, $q_k \equiv 1 \pmod{q_i q_j}$. The number of such $n \le x$ divisible by $q_i q_j q_k$ is

$$\ll \frac{x}{q_i q_j q_k}.$$

Summing up over all $q_k \le t$ which are congruent to $1 \pmod{q_i q_j}$ we get a count of

$$\frac{x}{q_i q_j} \sum_{\substack{q_k \le t \\ q_k \equiv 1 \pmod{q_i q_j}}} \frac{1}{q_k} \ll \frac{x \log_3 x}{(q_i q_j)^2}.$$

Summing up the above over all $q_i \in [\log_2 x, t]$ and $q_j \in [\log_2 x, t]$, we get a count of

$$\#\mathcal{A}_4(x) \ll x \log_3 x \left(\sum_{q > \log_2 x} \frac{1}{q^2}\right)^2 \ll \frac{x \log_3 x}{(\log_2 x)^2 (\log_3 x)^2} = o\left(\frac{x}{\log_2 x}\right)$$

as $x \to \infty$. Assume now that $n \notin \mathcal{A}_1(x) \cup \mathcal{A}_2(x) \cup \mathcal{A}_3(x) \cup \mathcal{A}_4(x)$. Assume that n is not $\mathcal{S}$-bad. Write $n = 2^a m + 1$ and fix $a \le y$. Then

$$m \equiv 2^s 3^k \pmod{2^n - 1}.$$

In particular,

$$m \equiv 2^s \cdot 3^k \pmod{2^d - 1}$$

for any $d \mid n$. In particular this is so when $d = q_i q_j$ for $i < j$ both in $\{1, \ldots, K\}$. Then $2^{q_i q_j} \equiv 1 \pmod{p_{i,j}}$, therefore $(2|p_{i,j}) = 1$ holds for all $1 \le i < j \le K$. Suppose first that k is even. Then $(m|p_{i,j}) = 1$. This shows that $m \equiv -2^{-a} \pmod{q_i}$ and in addition $(m|p_{i,j}) = 1$ for all $1 \le i < j \le K$. This puts m into $(p-1)/2$ of the possible p progressions modulo p for any of the $p = p_{i,j}$ and all $1 \le i < j \le K$. Varying i, j, we get that $m \equiv -2^{-a} \pmod{q_1 \cdots q_K}$ and m is in

$$\frac{1}{2^{\binom{K}{2}}} \prod_{1 \le i < j \le K} (p_{i,j} - 1)$$

progressions modulo $\prod_{1 \le i < j \le K} p_{i,j}$. The fact that all these primes are different is because $n \notin \mathcal{A}_4(x)$ and by the Chinese Remainder Theorem. The common modulus of these progres-

sions is

$$\begin{aligned}
&\leq q_1 \cdots q_K \prod_{1 \leq i < j \leq K} (2^{q_i q_j} - 1) \\
&< (\log x)^K (2^{t^2})^{K^2} = \exp\left(O\left(\frac{\log x (\log_3 x)^{4/3}}{(\log_2 x)^{1/2}} \right) \right) \\
&= x^{o(1)}
\end{aligned}$$

as $x \to \infty$. The number of such $n \leq x$ is at most

$$\frac{x}{2^{K(K-1)/2} q_1 \cdots q_k} + \frac{(p_{1,2} - 1) \cdots (p_{K-1,K} - 1)}{2^{K(K-1)/2}} \ll \frac{x}{2^{K(K-1)/2} q_1 \cdots q_K}.$$

We now sum up over $q_1, \ldots, q_K$ getting

$$\begin{aligned}
\frac{x}{2^{K(K-1)/2}} \sum_{q_1 < \cdots < q_K \leq t} \frac{1}{q_1 \cdots q_K} &\ll \frac{x}{2^{K(K-1)/2}} \frac{1}{K!} \left(\log_3 x + O(1) \right)^K \\
&\ll \frac{x}{2^{K(K-1)/2}} \left(\frac{e \log_3 x + O(1)}{(\log_3 x)^{2/3}} \right)^K \\
&\ll \frac{x}{2^{K^2/2 + O(K \log K)}}.
\end{aligned}$$

We now sum up over $a \leq y$, getting a count of

$$\#\mathcal{A}_{5,1}(x) \ll \frac{xy}{2^{K^2/2 + O(K \log K)}} = O\left(\frac{x}{2^{K^2/2 + O(K \log K)}} \right) = o\left(\frac{x}{\log_2 x} \right)$$

as $x \to \infty$, where we wrote $\mathcal{A}_{5,1}(x)$ for the set of $n \leq x$ not in $\cup_{i=1}^4 \mathcal{A}_i(x)$ for which congruence (1) holds with some even k. The case of k being odd is similar and leads to a comparable estimate. For this, we just need to distinguish 4 cases according to the classes of $p_{i,j}$ modulo 8 and modulo 3. For example, if $p_{i,j} \equiv 1 \pmod 8$ and $p_{i,j} \equiv 1 \pmod 3$, then

$$\left(\frac{3}{p_{i,j}} \right) = \left(\frac{p_{i,j}}{3} \right) = 1$$

so again $(m | p_{i,j}) = 1$. If $p_{i,j} \equiv 1 \pmod 8$ but $p_{i,j} \equiv 2 \pmod 3$, then

$$\left(\frac{3}{p_{i,j}} \right) = \left(\frac{p_{i,j}}{3} \right) = -1,$$

from which, together with the fact that k is odd and $(2 | p_{i,j}) = 1$, we get that $(m | p_{i,j}) = -1$. Similar arguments apply to the remaining cases when $p_{i,j} \equiv 7 \pmod 8$ and $p_{i,j} \equiv 1, 2 \pmod 3$. The point is that knowledge of $p_{i,j}$ modulo 24 together with the existence of the congruence $m \equiv 2^s 3^k \pmod{2^n - 1}$ determines the quadratic character of m modulo each of $p_{i,j}$ which saves a proportion of $1/2$ for each $p_{i,j}$. The theorem follows since $\varepsilon > 0$ is arbitrary with the worst (largest) upper bound being the one of $\mathcal{A}_{2,2}(x)$.

4 Generalizations

Surely, one can ask whether such an estimate would be valid if one replaces $2, 3$ in the congruence above by other primes, or even by primes from a fixed finite set. In reality, a similar argument shows the following.

Theorem 2 *Let $p_1, \ldots, p_k$ be primes, c be an integer and $\mathbf{u} = (u_m)_{m\geq 0}$ be a Lucas sequence. Let s denote an $\mathcal{S} = \{p_1, \ldots, p_k\}$-unit, that is an integer whose only prime factors are among $\{p_1, \ldots, p_k\}$. Then the set of positive integers n such that $n \equiv c + s \pmod{u_n}$ holds for some $\mathcal{S}$-unit s is of density 0. In fact, the counting function of such positive integers $n \leq x$ is at most*

$$\frac{x}{(\log_2 x)^{1+o(1)}}$$

as $x \to \infty$.

5 Heuristics for the case when *n* is prime

Our theorem (and method of proof) does not say anything about the case when $n = p$ is prime. Here, we make some considerations about primes p such that (1) holds for some exponents k, s. Let $\mathcal{B}(x)$ be the set of such primes $p \leq x$. We conjecture that $\mathcal{B}(x)$ is infinite but quite thin. In fact, we conjecture that

$$\log x \ll \#\mathcal{B}(x) \ll (\log x)^3. \tag{3}$$

The rest of this section is devoted to heuristics in support of (3).

To start with, let us note that any number of the form $n = 2^s 3^k + 1$ satisfies (1). The number of such numbers up to x is $O((\log x)^2)$. Almost certainly there are infinitely many primes of the above form but the counting function of them is quite likely very small. Using the heuristic that a random positive integer n is expected to be prime with a probability of $1/\log n$, the expectation that $n = 2^s 3^k + 1$ is prime should be about $O(1/(s+k))$. Summing this up over $n = 2^s 3^k + 1 \leq x$, we get a count of $O(\log x)$. This already suggests that the left inequality in (3) might hold, but there are additional candidates which we now explain.

Let us turn our attention to primes p such that $2^p - 1$ is prime (Mersenne primes). Heuristics of Crandall and Pomerance [4] predict that the expectation that p has the property that $2^p - 1$ is prime is $O(\log p/p)$. Summing this up over $p \leq x$ we get $O(\log x)$. We still have to comment about (1). Well, if 3 is a primitive root modulo $2^p - 1$, then every nonzero residue class is a power of 3. In particular, $(p-1)/2^{\nu_2(p-1)}$ should be a power of 3 modulo $2^p - 1$ so (1) should hold for some k even with $s = 0$ (or any fixed value of s). One can ask whether it is reasonable to conjecture that 3 is a primitive root modulo $2^p - 1$ often enough once $2^p - 1$ is prime. Note that by quadratic reciprocity

$$\left(\frac{3}{2^p - 1}\right) = -\left(\frac{2^p - 1}{3}\right) = -\left(\frac{1}{3}\right) = -1,$$

so 3 is not a quadratic residue modulo $2^p - 1$. Artin's conjecture predicts that 3 should be a primitive root modulo q for a positive proportion of primes q which is given by Artin's constant

$$A = \prod_{q\geq 2}\left(1 - \frac{1}{q(q-1)}\right) \sim 0.37\ldots.$$

The first 13 primes p such that $2^p - 1$ is prime are

$$2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521.$$

We ignore $p = 2$ since for it 3 is not invertible modulo $2^2 - 1$. Of the remaining 12 values of p, 8 of them (namely, $\{3, 5, 7, 17, 19, 89, 107, 521\}$) have the property that 3 is a primitive

root modulo $2^p - 1$, which is a significantly larger proportion then predicted by Artin's constant. So, assuming that for a positive proportion of primes p such that $2^p - 1$ is prime, also 3 is a primitive root, we get additional candidates for the left inequality in (3) to hold.

The Crandall–Pomerance heuristic can be extended to assume that $2^p - 1$ has ℓ distinct prime factors with probability $(\log p)^\ell/(\ell! p)$. Summing this up to x we get $O((\log x)^\ell/\ell!)$. Taking $\ell = 2, 3$, we get a count of $O((\log x)^3)$. We next give a heuristic that there are only finitely many primes p such that $\omega(2^p - 1) = \ell \geq 4$ and (1) holds for some k and s. This heuristic leads to the upper bound in (3). So, write

$$2^p - 1 = q_1^{\alpha_1} \cdots q_\ell^{\alpha_\ell}.$$

We know that $q_i \equiv 1 \pmod{p}$ for $i = 1, \ldots, \ell$. Thus,

$$\lambda(2^p - 1) = \mathrm{lcm}[\lambda(q_i^{\alpha_i}) : 1 \leq i \leq \ell] = \mathrm{lcm}[q_i^{\alpha_i - 1}(q_i - 1) : 1 \leq i \leq \ell] \mid \frac{\phi(2^p - 1)}{p^{\ell - 1}}.$$

In particular, the order of 3 modulo $2^p - 1$ is a divisor of $\phi(2^p - 1)/p^{\ell-1}$. The order of 2 modulo $2^p - 1$ is p. Hence, the cardinality of the multiplicative subgroup $H = \langle 2, 3 \rangle$ generated by 2, 3 modulo $2^p - 1$ divides

$$\frac{\phi(2^p - 1)}{p^{\ell-2}} \Big| \frac{\phi(2^p - 1)}{p^2}$$

assuming $\ell \geq 4$. Thus, the "probability" that some invertible residue class modulo $2^p - 1$ is in the above subgroup should be at most $O(1/p^2)$. Applying this to the fixed class $p - 1$ (or to $(p - 1)/2^{\nu_2(p-1)}$), we get that the probability that (1) holds for some k and s when $\ell \geq 4$ is $O(1/p^2)$, and since $\sum_{p \geq 2} 1/p^2$ is convergent, perhaps there are only finitely many such primes p altogether.

Acknowledgements We would like to express our sincere appreciation for the reviewer's careful reading, beneficial comments and suggestions, and to the editors for a prompt handling of our manuscript. Both authors worked on this paper during a visit to the Max Planck Institute for Software Systems in Saarbrücken, Germany in Spring 2023. They thank Professor J. Ouaknine for the invitation and the Institute for hospitality and support.

References

1. L. Carlitz, Permutations in a finite field. Proc. Amer. Math. Soc. **4**, 538 (1953)
2. F. Luca, S. Sarkar, P. Stănică, "Representing the inverse map as a composition of quadratics in a finite field of characteristic 2", manuscript, (2023)
3. P. Moree, On the divisors of $a^k + b^k$. Acta Arith. LXXX. **3**, 197–212 (1997)
4. R. Crandall, C. Pomerance, *Prime numbers. A computational perspective*, 2nd edn. Springer, New York, (2005). xvi+597 pp