Around Pelikán's conjecture on very odd sequences

Pieter Moree and Patrick Solé

Abstract

Very odd sequences were introduced in 1973 by Pelikán who conjectured that there were none of length ≥ 5. This conjecture was disproved first by MacWilliams and Odlyzko [15] in 1977 and then by two different sets of authors in 1992 [1], 1995 [8]. We give connections with duadic codes, cyclic difference sets, levels (Stufen) of cyclotomic fields, and derive some new asymptotic results on the length of very odd sequences and the number of such sequences of a given length.

1 Introduction

For a given natural number n consider the sequence $a_1 \dots a_n$ with $a_i \in \{0,1\}$ for $1 \leq i \leq n$. For $0 \leq k \leq n-1$, put $A_k = \sum_{i=1}^{n-k} a_i a_{i+k}$. We say that $a_1 \dots a_n$ is a *very odd sequence* if A_k is odd for $1 \leq k \leq n$. By $S(n)$ we denote the number of very odd sequences of length n. Pelikán [23] conjectured that there are no very odd sequences of length $n \geq 5$. However, the sequence 101011100011 (for which the corresponding A_k's are $7,3,3,1,3,3,3,1,1,1,1,1$) disproves this conjecture. Alles [1] showed that if $S(n) > 0$, then also $S(7n-3) > 0$ and thus showed that there are infinitely many counterexamples to Pelikán's conjecture. In his note [1] Alles raised two questions:
(1) Does $S(n) > 0$ imply $n \equiv 0, 1 (\text{mod } 4)$?
(2) Does $S(n) > 0$ imply $S(n) = 2^k$ for some integer k?
These two questions, which also appear as unsolved problem E38 in [6], were positively answered in [8].

All of the above was, however, already known much earlier [15]. For a and b coprime integers, let $\text{ord}_a(b)$ denote the smallest positive integer r such that $a^r \equiv 1 (\text{mod } b)$. Let $\mathcal{P} = \{7, 23, 31, 47, 71, 73, 79, \dots\}$ denote the set of odd primes p for which $\text{ord}_2(p)$ is odd (throughout this paper the letter p will be used to denote primes). MacWilliams and Odlyzko proved that $S(n) > 0$ iff $2n-1$ is composed only from primes in $\mathcal{P}$. (Alternatively we can formulate this as $S(n) > 0$ iff the order of 2 modulo $2n-1$ is odd.) Since $7 \in \mathcal{P}$, Alles result that

Mathematics Subject Classification (2000). 11N64, 94B15, 11N37

$S(n) > 0$ implies $S(7n-3) > 0$, immediately follows. Using the supplementary law of quadratic reciprocity it is easily seen, as already noticed in [15], that $p \in \mathcal{P}$ implies $p \equiv \pm 1 \pmod 8$. Thus if $S(n) > 0$, then $2n-1 \equiv \pm 1 \pmod 8$ and from this it follows that the answer to question (1) is affirmative. In [15] question (2) was answered in the affirmative also; to be more precise it was shown that if $S(n) > 0$ then $X^{2n-1} - 1$ decomposes over $\mathbb{F}_2$ into an odd number of distinct irreducible factors, say $2h+1$ irreducible factors in total and that, moreover, $S(n) = 2^h$. R. van der Veen and E. Nijhuis [35] described and implemented an algorithm to determine all very odd sequences of a given length. Moreover, they showed that, for $n > 1$, a very odd sequence is never periodic.

Let $N(x)$ denote the number of $n \leq x$ for which very odd sequences of length n exist. Let $\mathrm{Li}(x)$ denote the logarithmic integral. By [17, Theorem 2] it follows that

$$\mathcal{P}(x) = \frac{7}{24}\mathrm{Li}(x) + O\left(\frac{x(\log\log x)^4}{\log^3 x}\right), \tag{1}$$

where $\mathcal{P}(x)$ denotes the number of primes $p \leq x$ in $\mathcal{P}$. Using Theorem 4 of [17] it then follows there exists a $c_0 > 0$ such that

$$N(x) = c_0 \frac{x}{\log^{17/24} x}\left(1 + O\left(\frac{(\log\log x)^5}{\log x}\right)\right). \tag{2}$$

Since $\lim_{x\to\infty} N(x)/x = 0$, it follows in particular that Pelikán's conjecture holds true for almost all integers n. The estimates (1) and (2) sharpen the assertions $\mathcal{P}(x) \sim \frac{7}{24}\frac{x}{\log x}$, respectively $N(x) = o(x)$ made in [15].

The analysis of very odd sequences and that of Ducci sequences [3] shows a certain analogy (in the latter case the factorization of $(1+x)^n + 1$ over $\mathbb{F}_2[X]$ plays an important role and there is also a link with Artin's primitive root conjecture).

In Sections 2 and 3 the connection between very odd sequences and coding theory, respectively the Stufe (level) of cyclotomic fields is considered. In Section 4 a formula for $S(n)$ is derived (Proposition 2). In Section 5 the value distribution of $S(n)$ is then considered, using some results related to Artin's primitive root conjecture (Theorem 2 and Theorem 3). It is shown that under the Generalized Riemann Hypothesis (GRH) the preimage of the value 2^e is infinite, provided we can find just one integer n of a certain type satisfying $S(n) = 2^e$ (Theorem 4). This leads us then into a query of finding such integers and motivates the introduction of the notion of solution tableaux (Section 6.1). In the final section it is shown that certain values of $S(n)$ are much more sparse than others.

2 Codes and Pelikán's conjecture

For coding theoretic terminology, we refer to [16].

Lemma 1 *We have $S(n) > 0$ iff there exists a $[2n, n]$ extended binary cyclic code. In particular, there are very odd sequences of length $\frac{p+1}{2}$ for all odd primes $p \equiv \pm 1 \pmod 8$, and of length 2^m for all even integers $m \geq 2$.*

Proof. Write $b_k := a_{k-1}$ and $b(x) := \sum_{k=0}^{n-1} b_k x^k$, with x an indeterminate. With this notation as observed in [15], the definition of a very odd sequence amounts to the polynomial identity $x^{2n-1}+1=(x+1)b(x)b^*(x)$, the $*$ denoting reciprocation. From that, we see that the cyclic code of length $2n+1$ and generator $b(x)$ is self-orthogonal of dimension n. By adding an overall parity-check we obtain a self-dual code of length $2n$.

The two infinite families of lengths correspond to, respectively, quadratic residue codes attached to the prime p and Reed-Muller codes $\mathrm{RM}(\frac{m}{2}, m+1)$ (see [16]). $\square$

We denote by $C(n)$ the number of nonequivalent codes of length $2n$ obtained by the preceding lemma. The celebrated $[24, 12, 8]$ Golay code arises on taking $n = 12$. More generally, we obtain duadic codes with multiplier -1 which have received much attention in the last twenty years [11, 21]. In the special case of n a multiple of 8, these codes are, furthermore, doubly even. There are generalizations over $\mathbb{F}_4$ [22], $\mathbb{F}_q$ [28, 29] and over rings [9, 14]. The following connection with difference sets is proved differently in [29, Thm 6.2.1.] and in a very general setting in [25].

Lemma 2 *If there is a $2-(N, K, \lambda)$ cyclic difference set with K and λ both odd, then there is a very odd sequence of length $(N+1)/2$.*

Proof. We use the ring morphism (given by reduction mod 2) between the algebra $\mathbb{F}_2[x]/(x^N-1)$, where binary cyclic codes of length N live and the group algebra $\mathbb{Z}[C_N]$ which occurs in the following characterization [2, Lemma 3.2, p. 312] of difference sets of the cyclic group C_N :

$$DD^{-1} = K - \lambda + \lambda C_N.$$

The difference set D is then the set of exponents of x in the $b(x)$ of the preceding proof. $\square$

The following construction generalizes Alles's [1], who took $m = 4$. Let a (resp. b) denote a very odd sequence of length n (resp. m). Let $a(X)$ (resp. $b(X)$) denote their associated polynomials in $\mathbb{F}_2[X]$. By $a \otimes b$ we shall mean the sequence of attached polynomial $a(X)b(X^{2n-1})$. Explicitly this amounts to the sequence obtained by taking b, replacing each $b_i \in b$ by a block of m zeros (if $b_i = 0$) or $\mathbf{a}$ (if $b_i \neq 0$), and finally inserting a sequence of $n-1$ zeros between each of these blocks. The following result is given without a proof in [8].

Lemma 3 *If a and b are very odd sequences of respective lengths n and m, then $a \otimes b$ is a very odd sequence of length $2mn - n - m + 1$. If a' and b' are very odd sequence of lengths n and m, then $a \otimes b = a' \otimes b'$ implies $a = a'$ and $b = b'$.*

Proof. We know that, by hypothesis, $Y^{2m-1} + 1 = (Y+1)b(Y)b^*(Y)$. Letting $Y = X^{2n-1}$ and using the hypothesis on a, that is $X^{2n-1}+1 = (X+1)a(X)a^*(X)$, we see that $c(X) := a(X)b(X^{2n-1})$, satisfies $X^{2N-1}+1 = (X+1)c(X)c^*(X)$, for $2N-1 = (2n-1)(2m-1)$, i.e. $N = 2mn - n - m + 1$.

Conversely, writing c(X)=A(X)B(X), with A, B such that

- the order of the roots of A divide $2n-1$

- the order of the roots of B does *not* divide $2n-1$

we see that such a factoring is unique. We claim that $A=a$ and $B(X)=b(X^{2n-1})$ is of that shape. Indeed, if $\beta^{2n-1}=1$ and $b(\beta^{2n-1})=0$ we have that 1 is a root of $b(X)$, contradicting the definition of b. Alternatively note that the second part of the assertion follows from the explicit construction of $a\otimes b$. $\square$

Corollary 1 *We have* $S(2mn-n-m+1)\geq S(m)S(n)$.

3 Connection with the Stufe of a cyclotomic field

Let K be a field. Then the Stufe (or level) of K, $s(K)$, is defined as the smallest number s (if this exists) such that $-1=\alpha_1^2+\cdots+\alpha_s^2$ with all $\alpha_i\in K$. If the Stufe does not exist, it is not difficult to see that there exists an order $\leq$ on K that is compatible with the field operations (i.e. K is orderable). Pfister proved that the Stufe of any field, if it exists, is a power of two. For $m\geq 1$, let $K_m=\mathbb{Q}(e^{2\pi i/m})$. Hilbert proved that $s(K_m)\leq 4$ for $m\geq 3$. Fein et al. [5] proved that $s(K_{2m-1})=2$ iff $\mathrm{ord}_2(2m-1)$ is even. We thus obtain the following lemma.

Lemma 4 *We have* $S(n)>0$ *iff* $s(K_{2n-1})=4$.

Note that if $4|m$, then clearly $s(K_m)=1$ and for m odd we have $K_{2m}=K_m$, thus once we know $s(K_m)$ for every odd m, we know the Stufe of every cyclotomic field. Lemma 4 then shows that knowing for which n we have $S(n)>0$ is equivalent with knowing the Stufe for every cyclotomic field. Interestingly enough, knowing the Stufe of every imaginary quadratic number field is equivalent with knowing which integers can be represented as a sum of three integer squares (Gauss' famous three-squares theorem), see [27]. Let $St_4(x)$ count the number of $m\leq x$ such that $s(K_{2m-1})=4$. Then from Lemma 4 and (2) we infer that

$$St_4(x)=N(x)=c_0\frac{x}{\log^{17/24}x}\left(1+O\left(\frac{(\log\log x)^5}{\log x}\right)\right).$$

4 An explicit formula for $S(n)$

It is not difficult to derive an explicit formula for $S(n)$. To this end we first study the factorization of X^n-1 into irreducibles over $\mathbb{F}_q$, with q a power of a prime p. If $n=mp^e$, with $p\nmid n$, then since $X^n-1=(X^m-1)^{p^e}$ over $\mathbb{F}_q$, we can reduce to the case where $(n,q)=1$. Then we have the following result.

Lemma 5 *Let q be the order of a finite field. If $(n,q)=1$, then X^n-1 factors into*

$$i_q(n)=\sum_{d|n}\frac{\varphi(d)}{\mathrm{ord}_q(d)}\tag{3}$$

distinct irreducible factors over $\mathbb{F}_q$.

Proof. The assumption $(n,q) = 1$ ensures that the irreducibles will be distinct. Let $\Phi_d(X)$ denote the cyclotomic polynomial of degree d. We have $X^n - 1 = \prod_{d|n} \Phi_d(X)$, cf. [13, Theorem 2.45]. It is not difficult to show [13, Theorem 2.47] that the polynomial $\Phi_d(X)$ decomposes into $\phi(d)/\mathrm{ord}_q(d)$ irreducibles over $\mathbb{F}_q$ each having degree $\mathrm{ord}_q(d)$. On combining these results, the proof is then completed. □

The above formula for $i_2(n)$ together with results from [11], then yields the following first few values.

Table 1: $S(n)$ and $C(n)$ for small n

n	$S(n)$	$C(n)$
4	2	1
12	2	1
16	8	2
24	2	1
25	4	1
36	2	1
37	16	2
40	2	1
45	16	2
52	2	1
64	512	30

For $q \nmid n$, let $r_q(n) = \varphi(n)/\mathrm{ord}_q(n)$. Note that $r_q(n) = [(\mathbb{Z}/n\mathbb{Z})^* : \langle q\rangle]$. This quantity is sometimes called the (residual) index of q in $(\mathbb{Z}/n\mathbb{Z})^*$.

Lemma 6
1) *Suppose that* $\delta|d$ *and* $(d,q) = 1$, *then* $r_q(\delta)|r_q(d)$.
2) *If* $r_q(p) = r_q(p^2)$, *then* $r_q(p^e) = r_q(p)$ *for every* $e \geq 1$.

Proof. 1) The natural projection of multiplicative groups $(\mathbb{Z}/d\mathbb{Z})^* \to (\mathbb{Z}/\delta\mathbb{Z})^*$ gives rise to the projection $(\mathbb{Z}/d\mathbb{Z})^*/\langle q\rangle \to (\mathbb{Z}/\delta\mathbb{Z})^*/\langle q\rangle$, and so $r_q(\delta)$ divides $r_q(d)$ as claimed.
2) It is a well-known, and easy to prove, result in elementary number theory that if $\mathrm{ord}_2(p) = n$ and $\mathrm{ord}_2(p^2) = pn$, then $\mathrm{ord}_2(p^e) = p^{e-1}n$ for $e \geq 1$. □

Let $f(X) \in \mathbb{Z}[X]$ be an irreducible monic polynomial over $\mathbb{Q}$. A celebrated result of Dedekind (see e.g. [20, Theorem 4.12]) relates the factorization of $f(X)$ over $\mathbb{F}_p$ to the factorization of the ideal (p) into prime ideals in the ring of integers $\mathcal{O}$ of the quotient field $\mathbb{Q}[X]/f(X)$ (to each irreducible $f_i(X)$ of $f(X)$ over $\mathbb{F}_p$ corresponds a prime ideal lying over (p) in $\mathcal{O}$ of degree $\deg(f_i)$). When we apply this with $f(X) = \Phi_d(X)$ we see that in $\mathbb{Z}[\zeta_d]$, the ring of integers of $\mathbb{Q}(\zeta_d)$, the prime ideal (p) factorizes as $(p) = \mathcal{P}_1 \cdots \mathcal{P}_g$, where each $\mathcal{P}_i$ has degree $\mathrm{ord}_p(d)$ and $g = r_p(d)$ (see e.g. [20, Theorem 4.16]). This interpretation of $r_p(d)$ together with some basic facts from algebraic number theory gives another proof of Lemma 6 in case $q = p$. From the above remarks it follows that $i_p(n)$, for $p \nmid n$, counts the total number of prime ideals (p) factorizes in, in all the cyclotomic subfields of $\mathbb{Q}(\zeta_n)$.

The formula for $i_q(n)$ given in Lemma 5 also arises in some other mathematical contexts, see e.g. [24, 33, 34]. In particular we like to recall the following important result due to Ulmer [33].

Theorem 1 *Let p be a prime, n a positive integer, and d a divisor of p^n+1 that is coprime with 6. Let q be a power of p and let E be the elliptic curve over $\mathbb{F}_q(t)$ defined by $y^2+xy=x^3-t^d$. Then the j-invariant of E is not in $\mathbb{F}_q$, the conjecture of Birch and Swinnerton-Dyer holds for E, and the rank of $E(\mathbb{F}_q(t))$ equals $i_q(d)$.*

The function i_q is neither multiplicative nor additive, but we can still prove something in this direction. The proof will make use of the following proposition, the proof of which is left to the reader. (It requires elementary number theory only.)

Proposition 1 *Suppose that $q\nmid n_1n_2$.*
1) *If $(n_1,n_2)=1$, then $i_q(n_1n_2)\ge i_q(n_1)i_q(n_2)$.*
2) *If $(n_1,n_2)=1$ and $(\mathrm{ord}_q(n_1),\mathrm{ord}_q(n_2))=1$, then $i_q(n_1n_2)=i_q(n_1)i_q(n_2)$.*
3) *If $i_q(p_1^{e_1}\cdots p_s^{e_s})$ is prime, then for every $1\le i\le s$ there exists a $j\ne i$ such that $(\mathrm{ord}_q(p_i^{e_i}),\mathrm{ord}_q(p_j^{e_j}))>1$.*
4) *We have $i_q(n_1n_2)\ge i_q(n_1)+i_q(n_2)-1$.*

Proof. 1+2) If $(d_1,d_2)=1$, then $\mathrm{ord}_q(d_1,d_2)=\mathrm{lcm}(\mathrm{ord}_q(d_1),\mathrm{ord}_q(d_2))$ and so $\mathrm{ord}_q(d_1d_2)\le\mathrm{ord}_q(d_1)\mathrm{ord}_q(d_2)$. If in addition $(\mathrm{ord}_q(n_1),\mathrm{ord}_q(n_2))=1$, then $\mathrm{ord}_q(d_1d_2)=\mathrm{ord}_q(d_1)\mathrm{ord}_q(d_2)$.
3) This is a corollary to part 2.
4) We have

$$i_q(n_1n_2)=\sum_{d|n_1n_2}r_q(d)\ge\sum_{d|n_1}r_q(d)+\sum_{\substack{d|n_2\\ d\ne1}}r_q(n_1d)\ge i_q(n_1)+i_q(n_2)-1,$$

where in the derivation of the latter inequality we used part 1 of Lemma 6. □

The results from [15] as described above together with Lemma 5 yield an explicit formula for $S(n)$.

Proposition 2 *Let i_2 be as in Lemma 5. We have*

$$S(n)=\begin{cases}0 & \text{if } \mathrm{ord}_2(2n-1) \text{ is even;}\\ \sqrt{2}^{\,i_2(2n-1)-1} & \text{if } \mathrm{ord}_2(2n-1) \text{ is odd.}\end{cases}$$

Note that if $\mathrm{ord}_2(2n-1)$ is odd, then for every divisor $d>1$ of $2n-1$ we have that $\phi(d)/\mathrm{ord}_2(d)$ is even. Thus $i_2(2n-1)$ is odd and $S(n)$ is an integer, as a priori it has to be.

Using part 4 of Proposition 1 and the latter proposition, an alternate proof of Corollary 1 is obtained. We like to point out that often $S(mn+(m-1)(n-1))>$

$S(m)S(n)$; it can be shown for example that $S(n^2+(n-1)^2) = S(n)^2 > 0$ iff $\mathrm{ord}_2(2n-1)$ is odd and $2n-1$ is of the form p^k with $\mathrm{ord}_2(p) \neq \mathrm{ord}_2(p^2)$. From this equivalence and (2), we infer that $S(n^2+(n-1)^2) \neq S(n)^2$ for almost all n with $S(n) > 0$. Part 1 of Proposition 1 can also be turned into a, not so elegant, inequality for S.

It is easy to show that $\mathrm{ord}_2(p) = \mathrm{ord}_2(p^2)$ iff $2^{p-1} \equiv 1(\mathrm{mod}\ p^2)$. Primes satisfying the latter congruence are known as *Wieferich primes* and are discussed more extensively in the next section.

5 On the value distribution of S(n)

By Proposition 2 and the remark that if $\mathrm{ord}_2(2n-1)$ is odd, then $i_2(2n-1)$ is odd, we infer that $S(n) \in \{0, 1, 2, 4, 8, 16, 32, \ldots\}$. For v an integer, let $N_v(x)$ denote the number of $n \le x$ for which $S(n) = v$ and let N_v be the corresponding set of natural numbers n.

Proposition 3 *Let $e \ge 1$ be any natural number. Then N_{2^e} is non-empty.*

Proof. Let p be any prime with $p \equiv 3(\mathrm{mod}\ 4)$, $\mathrm{ord}_2(p) = (p-1)/2$ and $2^{\frac{p-1}{2}} \not\equiv 1(\mathrm{mod}\ p^2)$ (the prime 7 will do), then we claim that $(p^e+1)/2 \in N_e$. Our assumptions imply that $r_2(p) = r_2(p^2)$ and thus, by Proposition 6, we infer that $r_2(p^k) = r_2(p) = 2$ for $k \ge 1$. Hence $i_2(p^e) = 2e+1$. Note that $\mathrm{ord}_2(p^e)$ is odd. By Proposition 2 it then follows that $S(\frac{p^e+1}{2}) = 2^e$. □

Corollary 2 *We have $Im(S) = \{0, 1, 2, 4, 8, \ldots\}$.*

Proof. This follows from $\mathrm{Im}(S) \subseteq \{0, 1, 2, 4, 8, \ldots\}$, the proposition, $S(1) = 1$ and $S(2) = 0$. □

The sets N_0 and N_1 are relatively well understood; we have $N_0(x) = [x] - N(x)$, which together with (2) gives a good estimate and furthermore $N_1 = \{1\}$. It remains to deal with N_{2^e} for $e \ge 1$. We propose the following strengthening of Proposition 3.

Conjecture 1 *Let $e \ge 1$ be any natural number. There are infinitely many integers n for which $S(n) = 2^e$.*

Put

$$\mathcal{P}_m = \left\{p > 2 : \frac{p-1}{m} \text{ is odd and } \mathrm{ord}_2(p) = \frac{p-1}{m}\right\}$$

and

$$\mathcal{P}'_m = \{p \in \mathcal{P}_m : 2^{\frac{p-1}{m}} \not\equiv 1(\mathrm{mod}\ p^2)\}.$$

Note that if $p \in \mathcal{P}_m$, but $p \notin \mathcal{P}'_m$, then $\mathrm{ord}_2(p) = \mathrm{ord}_2(p^2)$ and hence p is a Wieferich prime. The proof of Proposition 3 shows that if we could prove that there are infinitely many primes in $\mathcal{P}'_2$ such that, the truth of Conjecture 1 follows. Although it seems very likely there are indeed infinitely many such primes, proving this is quite another matter. A prime being in $\mathcal{P}_2$ is closely being related to the Artin primitive root conjecture and a prime p satisfying

$2^{\frac{p-1}{2}} \equiv 1(\bmod p^2)$, is closely related to Wieferich's criterion for the first case of Fermat's Last Theorem. Wieferich proved in 1909 that if there is a non-trivial solution of $x^p + y^p = z^p$ with $p \nmid xyz$, then $2^{p-1} \equiv 1(\bmod p^2)$ (hence the terminology Wieferich prime). Despite extensive computational efforts, the only Wieferich primes ever found are 1093 and 3511 [4]. Note that $3511 \in \mathcal{P}_2 \backslash \mathcal{P}_2'$ and that $1093 \notin \cup_{m=1}^{\infty}(\mathcal{P}_m \backslash \mathcal{P}_m')$. Heuristics suggest that up to x there are only $O(\log\log x)$ Wieferich primes. If we replace 2 in Wieferich's congruence by an arbitrary integer a (i.e., consider primes p such that $a^{p-1} \equiv 1(\bmod p^2)$), then it is known that on average (the $O(\log\log x)$ heuristic holds true [19], where the averaging is over a. Assuming the abc-conjecture it is known [26] that there are $\gg \log x/\log\log x$ primes $p \le x$ that are non-Wieferich primes. For our purposes it would be already enough to know that there only $o(x/\log x)$ Wieferich primes $\le x$, however, even this is unproved.

The situation regarding $\mathcal{P}_m$ is slightly more promising. Under GRH we can namely prove the following result. Our proof rests on a variation of Hooley's proof of Artin's primitive root conjecture that belongs to a class of variations of this problem dealt with by H.W. Lenstra [10]. This relieves us from the burden of dealing with its analytic aspects.

Theorem 2 (GRH). *Let m be a natural number. Let $\nu_2(m)$ denote the exponent of 2 in m. If m is odd or $\nu_2(m) = 2$, then $\mathcal{P}_m$ is empty. In the remaining cases we have, under GRH,*

$$\mathcal{P}_m(x) = \frac{2A\epsilon_1(m)}{3m^2} \prod_{p|m} \frac{p^2-1}{p^2-p-1} \frac{x}{\log x} + O\left(\frac{x\log\log x}{(\log x)^2}\right), \tag{4}$$

where $A = \prod_p \left(1 - \frac{1}{p(p-1)}\right) \approx 0.3739558136\ldots$ denotes the Artin constant and

$$\epsilon_1(m) = \begin{cases} 1 & \text{if } \nu_2(m) = 1; \\ 2 & \text{if } \nu_2(m) \ge 3. \end{cases}$$

Proof. If m is odd, then $(p-1)/m$ is even and hence $\mathcal{P}_m$ is empty. If $4||m$, then $(p-1)/m$ is odd implies that $p \equiv 5(\bmod 8)$ and $\mathrm{ord}_2(p-1) = (p-1)/m$ implies that $(\frac{2}{p}) = 1$. By the supplementary law of quadratic reciprocity the conditions $(\frac{2}{p}) = 1$ and $p \equiv 5(\bmod 8)$ cannot be satisfied at the same time. Hence $\mathcal{P}_m$ is empty.

For the remainder of the proof we assume GRH. Then it can be shown, cf. [10, 36] that the set of primes p such that $\mathrm{ord}_2(p) = (p-1)/m$ satisfies an asymptotic of the form (4) with constant

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{[\mathbb{Q}(\zeta_{nm}, 2^{1/nm}) : \mathbb{Q}]}. \tag{5}$$

Let f be an arbitrary natural number. If in addition to requiring $\mathrm{ord}_2(p) = (p-1)/m$, we also require $p \equiv 1(\bmod f)$, it is readily seen that the sum in (5) has to be replaced by

$$\delta(f, m) = \sum_{n=1}^{\infty} \frac{\mu(n)}{[\mathbb{Q}(\zeta_f, \zeta_{nm}, 2^{1/nm}) : \mathbb{Q}]}.$$

Let $r=\nu_2(m)$. Note that $(p-1)/m$ is odd iff $p\equiv 1+2^r(\text{mod } 2^{r+1})$. We conclude that (4) holds with constant $\delta(2^r,m)-\delta(2^{r+1},m)$. It is not difficult to see that

$$\delta(2^r,m)-\delta(2^{r+1},m)=\sum_{\substack{n=1\\ 2\nmid n}}^{\infty}\frac{\mu(n)}{[\mathbb{Q}(\zeta_{2nm},2^{1/nm}):\mathbb{Q}]}. \tag{6}$$

Let $s|t$. It is known (cf. [18, Lemma 1]) that

$$[\mathbb{Q}(\zeta_t,2^{1/s}):\mathbb{Q}]=\begin{cases}\varphi(t)s/2 & \text{if } 2|s \text{ and } 8|t;\\ \varphi(t)s & \text{otherwise.}\end{cases} \tag{7}$$

On using the latter formula for the degree, we infer after some tedious calculation that the constant in (6) equals the constant in (4). □

Corollary 3 (GRH). *If e is odd or $4|e$, then N_{2^e} is an infinite set.*

Proof. If $p\in\mathcal{P}_{2e}$, then $S(\frac{p+1}{2})=2^e$. Hence the infinitude of $\mathcal{P}_{2e}$ for e is odd and for $4|e$, implies the result. □

Corollary 3 shows that under GRH Conjecture 1 holds true, provided that e is odd or $4|e$. It is possible to go further than this, but this requires a result going beyond Theorem 2:

Theorem 3 (GRH). *Suppose that $\nu_2(e)\neq 1$. Let a and f be integers such that $4e|f$, $\nu_2(4e)=\nu_2(f)$, $(a,f)=1$ and $a\equiv 1+2e(\text{mod } 4e)$. There exists an integer v such that for all squarefree n we have $a\equiv 1(\text{mod } (f,2en))$ iff $(n,v)=1$. The density of primes p such that $r_2(p)=2e$ and $p\equiv a(\text{mod } f)$ exists, is positive, and is given by*

$$\frac{\epsilon_1(2e)}{\varphi([f,2e])2e}\prod_{p\nmid v}\left(1-\frac{\varphi([f,2e])}{\varphi([f,2ep])p}\right).$$

Corollary 4 (GRH). *Suppose that $\nu_2(2e)\neq 2$, $2\nmid f$, $(e,f)=1$ and $(a,f)=1$. Then the set $\mathcal{P}_{2e}\cap\{p:p\equiv a(\text{mod } f)\}$ contains infinitely many primes.*

Proof of Theorem 3. The existence of the density, denoted by δ, follows by the work of Lenstra [10]. One obtains that

$$\delta=\sum_{n=1}^{\infty}\frac{\mu(n)c_2(a,f,2en)}{[\mathbb{Q}(\zeta_f,\zeta_{2en},2^{1/2en}):\mathbb{Q}]},$$

with

$$c_2(a,f,k)=\begin{cases}1 & \text{if } \sigma_a|_{\mathbb{Q}(\zeta_f)\cap\mathbb{Q}(\zeta_k,2^{1/k})}=\text{id};\\ 0 & \text{otherwise}\end{cases},$$

where σ_a is the automorphism of $\mathbb{Q}(\zeta_f):\mathbb{Q}$ uniquely determined by $\sigma_a(\zeta_f)=\zeta_f^a$. Under the conditions of the result one infers that

$$\mathbb{Q}(\zeta_f)\cap\mathbb{Q}(\zeta_{2en},2^{1/2en})=\mathbb{Q}(\zeta_{(f,2en)})$$

and hence $c_2(a,f,2en)=1$ iff $a\equiv 1(\text{mod } (f,2en))$. Let n be squarefree. Note that there exists an integer v such that $a\equiv 1(\text{mod } (f,2en))$ iff $(n,v)=1$. We infer that $[\mathbb{Q}(\zeta_f,\zeta_{nm},2^{1/nm}):\mathbb{Q}]=[\mathbb{Q}(\zeta_{[f,2en]},2^{1/2en}):\mathbb{Q}]=\varphi([f,2en])2en/\epsilon_1(2e)$, by (7). We thus find that

$$\delta=\frac{\epsilon_1(2e)}{\varphi([f,2e])2e}\sum_{(n,v)=1}\frac{\mu(n)\varphi([f,2e])}{\varphi([f,2en])n}=\frac{\epsilon_1(2e)}{\varphi([f,2e])2e}\prod_{p\nmid v}\left(1-\frac{\varphi([f,2e])}{\varphi([f,2ep])p}\right)>0,$$

where we used that the sum is absolutely convergent and has a summand that is a multiplicative function in n. □

`Remark`. An alternative way of proving Theorem 2 is to use the, Galois-theoretic, method of Lenstra, Moree and Stevenhagen [12], cf. [30]. This yields, a priori, that the density is of the form $(1+\prod_{p|2ft}E_p)A$, where the E_p are (real) character averages and hence $-1\le E_p\le 1$.

Now, under GRH, we can prove some further results regarding Conjecture 1.

Proposition 4 (GRH). *Let e be a natural number. Suppose that $1+2e$ is not a prime number congruent to $5(\text{mod } 8)$, then $S(n)=2^e$ for infinitely many n.*

Proof. If $p\in\mathcal{P}_{2e}$, then $S(\frac{p+1}{2})=2^e$. Hence the infinitude of $\mathcal{P}_{2e}$ for e is odd and for $4|e$, implies the result in case $1+2e\not\equiv 5(\text{mod } 8)$. Next suppose that $1+2e\equiv 5(\text{mod } 8)$ and is not a prime. Then there exist natural numbers e_1 and e_2 with $1+2e=(1+2e_1)(1+2e_2)$ and $1+2e_1\not\equiv 5(\text{mod } 8)$ and $e_2\ge 1$. Suppose that n is such that $2n-1=7^{e_2}q$ with

$$q\not\equiv 1(\text{mod } 6e_1)\text{ and } q\not\equiv 1(\text{mod } 14e_1)\text{ and } q\in\mathcal{P}_{2e_1}. \tag{8}$$

The conditions on q ensure that $q\neq 7$ and, moreover, $(7\text{ord}_2(7),\text{ord}_2(q))=1$ and hence $r_2(7^kq)=r_2(7^k)r_2(q)$ for every k. It follows that

$$i_2(2n-1)=\sum_{d|7^{e_2}}r_2(d)+r_2(q)\sum_{d|7^{e_2}}r_2(d)=(1+2e_1)(1+2e_2)=1+2e,$$

and hence $S(n)=2^e$. Since $e_1\not\equiv 2(\text{mod } 4)$, it follows by Theorem 3 that there are infinitely many primes q satisfying (8). □

Some of the cases left open by the previous results are covered by the following result.

Proposition 5 (GRH). *Suppose that $1+2e=z$ with z a prime satisfying $z\equiv 5(\text{mod } 104)$, then $S(n)=2^e$ for infinitely many n.*

Proof. The n for which $2n-1=7^2p$, where $(\text{ord}_2(p),21)=3$ and $p\in\mathcal{P}_{(z-5)/13}$ will do. Now invoke Theorem 3. □

For the values of $S(n)$ hitherto not covered, the following result can sometimes be applied. Recall that $\omega(n)=\sum_{p|n}1$. We define $\omega_1(n)=\sum_{p||n}1$, i.e. the number of distinct prime factors p of n such that $p^2\nmid n$. Note that $\omega_1(n)\le\omega(n)$.

Theorem 4 (GRH). *Suppose that $S(n)=2^e$ for some n with $\omega_1(2n-1)\ge 1$, then there are infinitely many n for which $S(n)=2^e$.*

Our proof rests on the following exchange principle.

Proposition 6 *Let p and q be odd primes and m be a natural number such that $(m, 2pq) = 1$, $r_2(p) = r_2(q)$ and $(\frac{p-1}{r_2(p)}, \mathrm{ord}_2(m)) = (\frac{q-1}{r_2(q)}, \mathrm{ord}_2(m))$, then $i_2(pm) = i_2(qm)$.*

Proof. The proof is a consequence of the identity

$$\sum_{d|m} r_2(pm) = \sum_{d|m} r_2(p) r_2(m)(\mathrm{ord}_2(p), \mathrm{ord}_2(d)),$$

and the observation from elementary number theory that if a_1, a_2, b and d are natural numbers with $(a_1, b) = d$ and $(a_2, b) = d$, then the equality $(a_1, \beta) = (a_2, \beta)$ is satisfied for all $\beta | b$. □

Proof of Theorem 4. The assumption $\omega_1(2n-1) \geq 1$ implies that $2n - 1 = pm$ with $p \nmid m$. By Theorem 2 there exists a number e with $\nu_2(2e) \neq 2$ such that $p \in \mathcal{P}_{2e}$. Let q be any prime number such that

$$q \in \mathcal{P}_{2e} \text{ and } (\frac{q-1}{r}, \mathrm{ord}_2(m)) = (\frac{p-1}{r}, \mathrm{ord}_2(m)), \tag{9}$$

then by Proposition 6 we infer that $S(qm) = S(pm) = 2^e$. By Theorem 3 there are infinitely many primes q satisfying (9) and thus $S(\frac{qm+1}{2}) = 2^e$. □

6 Value distribution of $S(n)$ on n's with $2n - 1$ squarefree

Before returning to the value distribution of $S(n)$, we address the easier problem of studying the value distribution of $S(n)$ with n restricted to those n for which $2n - 1$ is squarefree. Let $V = \{e \ : \ \mu(2n-1) \neq 0 \text{ and } S(n) = 2^e\}$. Let

$$r_\mu(e) = \max\{\omega(2n-1) \ : \ \mu(2n-1) \neq 0 \text{ and } S(n) = 2^e\}.$$

and

$$\min_\mu(e) = \min\{2n-1 \ : \ \omega(2n-1) = r_\mu(e), \ \mu(2n-1) \neq 0 \text{ and } S(n) = 2^e\},$$

if this exists. In the sequel we will also need the related quantities $r_{\omega_1}(e)$ and $r_\omega(e)$. Let

$$r_{\omega_1}(e) = \max\{\omega_1(2n-1) \ : \ S(n) = 2^e\} \text{ and } r_\omega(e) = \max\{\omega(2n-1) \ : \ S(n) = 2^e\},$$

and $\min_{\omega_1}(e)$ and $\min_\omega(e)$ be defined as the smallest value of $2n - 1$ for which $\omega_1(2n-1) = r_{\omega_1}(e)$, respectively $\omega(2n-1) = r_\omega(e)$ and $S(n) = 2^e$ (if these exist). Note that $r_\mu(e), r_{\omega_1}(e)$ and $r_\omega(e)$ exist and satisfy $r_\mu(e) \leq r_{\omega_1}(e) \leq r_\omega(e)$.

Lemma 7 *We have $r_\mu(e) \leq r_{\omega_1}(e) \leq r_\omega(e) \leq [\frac{\log(2e+1)}{\log 3}]$ for $e \geq 0$.*

Proof. Suppose $S(n) = 2^e$. The number $2n-1$ is composed of only primes $p \in \mathcal{P}$. Write $2n-1 = p_1^{e_1}\cdots p_s^{e_s}$. Then $i_2(2n-1) \geq i_2(p_1\cdots p_s)$. We have $\mathrm{ord}_2(p_i) \leq (p_i-1)/2$ and $\mathrm{ord}_2(d) \leq \phi(d)2^{-\omega(d)}$. Thus

$$i_2(2n-1) \geq i_2(p_1\cdots p_s) \geq \sum_{d|p_1\cdots p_s} 2^{\omega(d)} = 3^s.$$

If $s > [\log(2e+1)/\log 3]$, then it follows that $i_2(2n-1) > 2e+1$ and hence $S(2n-1) > 2^e$. □

We now formulate the main result of this section.

Theorem 5 (GRH). *Given an integer e, it is a finite problem to determine whether there exist $e \in V$. The quantity $r_\mu(e)$ can be effectively computed.*

In order to prove it, it turns out to be fruitful to introduce the notion of a solution tableau.

6.1 Solution tableaux

Let $s \geq 2$. We define a map $\lambda : \mathbb{Z}_{>0}^s \to \mathbb{Z}_{>0}^s$, $(a_1,\ldots,a_s) \to (l_1,\ldots,l_s)$ as follows. We put $l_i = \mathrm{lcm}((a_1,a_i),\ldots,(a_{i-1},a_i),(a_{i+1},a_i),\ldots,(a_s,a_i))$ for $1 \leq i \leq s$. Note that $\nu_p(l_i) = \nu_p(a_i)$ if $\nu_p(a_i) \leq \nu_p(a_j)$ for some $j \neq i$ and $\nu_p(l_i) = \max\{\nu_p(a_1),\ldots,\nu_p(a_{i-1}),\nu_p(a_{i+1}),\ldots,\nu_p(a_s)\}$ otherwise. In particular, if there exist i and j with $i \neq j$ such that

$$\nu_p(a_i) = \nu_p(a_j) = \max\{\nu_p(a_1),\ldots,\nu_p(a_s)\}, \tag{10}$$

then $\nu_p(l_i) = \nu_p(a_i)$. If (10) holds for every prime p (with i and j possibly depending on p), then $(a_1,\ldots,a_s)$ is said to be *realizable.* If $(a_1,\ldots,a_s)$ is realizable, then $\nu_p(l_i) = \nu_p(a_i)$ for every $1 \leq i \leq s$ and every prime p and hence $\lambda(a_1,\ldots,a_s) = (a_1,\ldots,a_s)$. The terminology realizable is motivated by the following result.

Proposition 7 *We have $(m_1,\ldots,m_s) \in \mathrm{Im}(\lambda)$ iff $(m_1,\ldots,m_s)$ is realizable.*

Proof. '$\Rightarrow$'. By assumption $\lambda(a_1,\ldots,a_s) = (m_1,\ldots,m_s)$ for some $(a_1,\ldots,a_s) \in \mathbb{Z}_{>0}^s$. W.l.o.g. assume that $\nu_p(m_1) = \max\{\nu_p(m_1),\ldots,\nu_p(m_s)\}$. Then it follows that $p^{\nu_p(m_1)}|(a_1,a_i)|m_i$ for some $1 < i \leq s$ and so $\nu_p(m_i) = \nu_p(m_1)$.
'$\Leftarrow$'. If $(m_1,\ldots,m_s)$ is realizable, then $\lambda(m_1,\ldots,m_s) = (m_1,\ldots,m_s)$. □

Suppose $(m_1,\ldots,m_s)$ is realizable and that a_i is coprime with $\prod_{j=1}^{i-1} a_j \prod_{j=1}^{s} m_j$ for $i = 1,\ldots,s$, then $\lambda(a_1m_1,\ldots,a_sm_s) = (m_1,\ldots,m_s)$. Thus if $(m_1,\ldots,m_s)$ is in $\mathrm{Im}(\lambda)$, then clearly its preimage is an infinite set. However, if $\lambda(a_1,\ldots,a_s) = (l_1,\ldots,l_s)$, then for $i \neq j$, $(a_i,a_j) = (l_i,l_j)$. Thus from a set of merely s numbers all $s(s-1)/2$ gcd's (a_i,a_j) can be computed. This is the motivation of introducing the map λ as it leads to a simple expression (11) for $i_2(p_1\cdots p_s)$. Note that alternatively $i_2(p_1\cdots p_s)$ can be expressed in terms of $r_2(p_i)$, $1 \leq i \leq s$ and the $s(s-1)/2$ gcd's $(\mathrm{ord}_2(p_i),\mathrm{ord}_2(p_j))$.

Proposition 8 *Given* $(a_1,\ldots,a_s)\in\mathbb{Z}^s_{>0}$, *let* $\lambda(a_1,\ldots,a_s)=(l_1,\ldots,l_s)$. *We have*

$$\mathrm{lcm}(a_{i_1},\ldots,a_{i_k})=\frac{a_{i_1}\ldots a_{i_k}}{l_{i_1}\ldots l_{i_k}}\mathrm{lcm}(l_{i_1},\ldots,l_{i_k}),$$

where $k\ge 2$ *and* $1\le i_1<i_2<\ldots<i_k\le s$.

Proof. It is enough to prove the assertion in case $a_i=p^{e_i}$ with $e_i\ge 0$. W.l.o.g. assume that $e_1\le e_2\le\ldots\le e_s$. On noting that $l_i=p^{e_i}$ for $i<s$ and that $l_s=p^{e_{s-1}}$, the result follows after a simple computation. □

Let $p_1,\ldots,p_s$ be distinct odd primes. Using the latter proposition we infer that

$$i_2(p_1\cdots p_s)=\sum_{v_1=0}^{1}\cdots\sum_{v_s=0}^{1}e_1^{v_1}\cdots e_s^{v_s}\frac{l_1^{v_1}\cdots l_s^{v_s}}{\mathrm{lcm}(l_1^{v_1},\ldots,l_s^{v_s})},\tag{11}$$

with $(e_1,\ldots,e_s)=(r_2(p_1),\ldots,r_2(p_s))$ and $\lambda(\mathrm{ord}_2(p_1),\ldots,\mathrm{ord}_2(p_s))=(l_1,\ldots,l_s)$, where we order the prime factors $p_1\ldots p_s$ in such a way that $e_i\le e_j$ for $i\le j$ and $l_i\le l_j$ if $e_i=e_j$. We say that $\binom{e_1\ \cdots\ e_s}{l_1\ \cdots\ l_s}$ is the *tableau* associated to $m=p_1\cdots p_s$. We say that $\binom{e_1\ \cdots\ e_s}{l_1\ \cdots\ l_s}$ is a *solution tableau* if $\nu_2(e_i)\ne 2$, e_i is even and l_i is odd for every l_i and, moreover, $(l_1,\ldots,l_s)$ is realizable. The value associated to a solution tableau is the quantity in the right hand side of (11). The following result is a consequence of Theorem 3.

Proposition 9 *The tableau associated to a squarefree integer* $2n-1$ *satisfying* $S(n)=2^e$ *is a solution tableau. Given any solution tableau* T, *under GRH, there exist infinitely many squarefree integers* $2n-1$ *such that the associated solution tableau equals* T.

`Example 1`. Let us consider the tableau $\binom{2\ \ 2\ \ 8}{3\ 15\ 5}$. Note that it is a solution tableau corresponding to the value 601 of i_2. Let us try to find an $m=p_1{\cdot}p_2{\cdot}p_3$ having this tableau associated to it. We need to find primes p_1,p_2 in $\mathcal{P}_2$ and p_3 in $\mathcal{P}_8$ such that $(\mathrm{ord}_2(p_1),\mathrm{ord}_2(p_2))=3$, $(\mathrm{ord}_2(p_1),\mathrm{ord}_2(p_3))=1$ and $(\mathrm{ord}_2(p_2),\mathrm{ord}_2(p_3))=5$. Any p_1 in $\mathcal{P}_2$ satisfying $p_1\equiv 1(\mathrm{mod}\ 3)$ will fit the bill. E.g. $p_1=79$. Any p_2 in $\mathcal{P}_2$ satisfying $p_2\equiv 1(\mathrm{mod}\ 15)$ and $p_2\not\equiv 1(\mathrm{mod}\ 13)$ will do, e.g. $p_2=991$. Finally any $p_3\in\mathcal{P}_8$ satisfying $p_3\equiv 2(\mathrm{mod}\ 3)$, $p_3\equiv 1(\mathrm{mod}\ 5)$, $p_3\not\equiv 1(\mathrm{mod}\ 11)$ and $p_3\not\equiv 1(\mathrm{mod}\ 13)$ will do, e.g. $p_3=1721$. We have, as expected, $i_2(79\cdot 991\cdot 1721)=601$. (Despite the numerous conditions on p_3, by Corollary 4 there exist, under GRH, infinitely many such primes.)

Proposition 10 *Let* $r\ge 1$ *be any integer. The set of tableaux associated to the set* $\{m:2\nmid\mathrm{ord}_2(m),\ \mu(m)\ne 0,\ i_2(m)\le r\}$ *is finite.*

Proof. By Lemma 7 we have $\omega(m)\le r_\mu(e)$. Thus the number of entries in a row of an associated tableau is bounded by $r_\mu(e)$. Now fix any $s\le r_\mu(e)$. Note that $i_2(p_1\ldots p_s)\ge e_1\ldots e_s l_1\ldots l_s/\mathrm{lcm}(l_1,\ldots,l_s)$. There are only finitely many possibilities for $e_1,\ldots,e_s$, $l_1,\ldots,l_s$ such that $\binom{e_1\ \cdots\ e_s}{l_1\ \cdots\ l_s}$ is a solution tableau and $i_2(p_1\ldots p_s)\le r$. □

`Example 2`. We try to find all primes $z\equiv 5(\mathrm{mod}\ 8)$ with $z\le 229$ for which there are distinct primes p and q both in $\mathcal{P}$ such that $i_2(p\cdot q)=z$. This leads us to find all z of the above form for which there are e_1, e_2 and w satisfying $1 + 2e_1 + 2e_2 + 4e_1e_2w = z$ with $\nu_2(e_1) \neq 1$, $\nu_2(e_2) \neq 1$ and $w \geq 3$ is odd. There are solutions precisely for $z \in \{101, 157, 197, 269, 317, 349, 421, 509, \ldots\}$. The associated solution tableau is $\binom{2e_1\ 2e_2}{w\ \ w}$. In each case one can find p and q corresponding to this solution tableau. E.g., when $z = 421$ an associated solution tableau is $\binom{2\ 38}{5\ \ 5}$. A solution corresponding to this is $p = 71$ and $q = 174991$.

Proof of Theorem 5. One can effectively determine the solution tableaux corresponding to the given value of e and hence $r_\mu(e)$ (it is the largest number of columns amongst these solution tableaux). Note that by Proposition 10 there are only finitely many associated solution tableaux. If there is no associated solution tableau, then $e \notin V$. If there is a solution tableau, then under GRH, it is possible to find an m corresponding to it (cf. Example 1) and thus $e \in V$.

We now also have the tools to show that the upper bound in Lemma 7 is sharp for infinitely many e.

Proposition 11 (GRH). *Let $s \geq 2$. If $e = (3^s - 1)/2$, then $r_\mu(e) = r_{\omega_1}(e) = r_\omega(e) = s$.*

Proof. A solution tableau with $s+1$ columns clearly does not exist. On the other hand $T = \binom{2}{1} \cdots \binom{2}{1}$ (s columns) is a solution tableau. Now apply Proposition 9. □

`Remark`. If we restrict e to be such that $1 + 2e$ is prime, then the upper bound is far from sharp. This is a consequence of part 3 of Proposition 1.

6.2 The general problem revisited

Theorem 5 has the following more important variant.

Theorem 6 (GRH). *Assume that the number of primes $p \leq x$ such that $ord_2(p)$ is odd and p is a Wieferich prime is $o(x/\log x)$. Given an integer e, it is a finite problem to determine whether or not there exist n with $S(n) = 2^e$ and $\omega_1(2n-1) \geq 1$. The quantities $r_{\omega_1}(e)$ and $r_\omega(e)$ can be effectively computed.*

For clarity, we first consider some examples. It will be convenient to divide the integers n with $\mathrm{ord}_2(n)$ odd in two classes. We say $n = \prod_{i=1}^s p_i^{e_i}$ is of type II if there exists i and j such that $e_i \geq 2$ and $p_i | \mathrm{ord}_2(p_j)$ and of type I otherwise.

`Example 3`. We try to find all primes $z \equiv 5 (\mathrm{mod}\ 8)$ with $z \leq 229$ for which there are distinct primes p and q both in $\mathcal{P}$ such that $i_2(p^r \cdot q^s) = z$, with $r, s \geq 1$.
-Let us assume first that $p^r \cdot q^s$ is of type I. Then we have to find all z of the above form for which there are e_1, e_2, r, s and w satisfying $1 + 2e_1r + 2e_2s + 4e_1e_2rsw = z$ with $\nu_2(e_1) \neq 1$, $\nu_2(e_2) \neq 1$ and $w \geq 3$ odd. On making the substitution $f_1 = e_1r$ and $f_2 = e_2r$, we have to solve $1 + 2f_1r + 2f_2s + 4f_1f_2w$, with $w \geq 3$ is odd. This equation is of the same format as the one arising in Example 2, except that now there are no restrictions on f_1 and f_2. There are solutions precisely for $r \in \{101, \underline{109}, 157, \underline{173}, 197, 269, 317, 349, 421, 509, \ldots\}$. The two underlined numbers did not arise in Example 2. A preimage for 109 is readily found using

Proposition 5. One finds that, e.g., $i_2(7^2 \cdot 73) = 109$. For 173 one finds that the preimage has to be of the form $p^2 \cdot q$ with solution tableau $\binom{2\,8}{5\,5}$ associated to $p \cdot q$. One finds that $71^2 \cdot 1721$ is the smallest preimage of the required format.
-Next assume that $p^r \cdot q^s$ is of type II. In this case $i_2(p^r \cdot q^s) \geq 11 + 4\min\{p, q\}$. We only have to analyze the case where $p \in \{7, 23, 31, 47\}$. The only prime $z \equiv 5 (\bmod\ 8)$ with $z \leq 229$ that is assumed as value turns out to be 197. This can happen only if p or q equals 23, e.g., $i_2(23^3 \cdot 47) = 197$ (we leave the details to the reader).

`Example` 4. We try to find all primes $z \equiv 5 (\bmod\ 8)$ with $z \leq 229$ for which there exists an integer m with $\mathrm{ord}_2(m)$ is odd and $\omega(m) \geq 3$ such that $i_2(m) = z$. In case $\omega(m) \geq 4$ one has $i_2(m) \geq 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 5 = 240$, so assume that $\omega(m) = 3$. In case m is of type I one finds, using solution tableaux, that there is no such z. Indeed, the smallest prime $z \equiv 5 (\bmod\ 8)$ occurring thus is 509. It arises only if $m = p^2 \cdot q \cdot r$ and the solution tableau associated to $p \cdot q \cdot r$ is $\binom{2\,2\,2}{5\,5\,5}$. One finds that, e.g., $i_2(71^2 \cdot 191 \cdot 271) = 509$. The smallest prime $z \equiv 5 (\bmod\ 8)$ occurring for a type II integer is 389. One has, e.g., $i_2(7^2 \cdot 71 \cdot 191) = 389$.

In Example 3 one can see the following exchange principle for integers of type I at work.

Proposition 12 *Suppose that $s \geq 2$ and $(p, \mathrm{ord}_2(m)) = 1$ and q is a prime satisfying $r_2(q) = \sum_{j=1}^{s} r_2(p^j)$ (that is $\mathrm{ord}_2(q) = s r_2(p)$ in case q is not a Wieferich prime), then $i_2(p^s \cdot m) = i_2(q \cdot m)$.*

For the prime q above $r_2(q)$ is even, but it is not necessarily the case that $\mathrm{ord}_2(q)$ is odd.

We say that $\binom{e_1}{l_1} \cdots \binom{e_s}{l_s}$ is a *generalized solution tableau* if $2|e_i$, $2 \nmid l_i$ for $1 \leq i \leq s$ and, moreover, $(l_1, \ldots, l_s)$ is realizable. In order to find an integer m of type I such that $i_2(m) = r$ and $\mathrm{ord}_2(m)$ is odd, we first determine all generalized solution tableaux $\binom{f_1}{l_1} \cdots \binom{f_s}{l_s}$ associated to r. If there are none, there is no solution. If there is a generalized solution tableau and $\nu_2(f_i) \neq 2$ for $1 \leq i \leq s$, the generalized solution tableau is even a solution tableau and a squarefree integer m can be found such that $i_2(m) = r$. W.l.o.g. suppose that $\nu_2(f_i) = 2$ for $1 \leq i \leq t$ and $\nu_2(f_i) \neq 2$ for $t < i \leq s$. Put $e_i = f_i/2$ for $1 \leq i \leq t$ and $e_i = f_i$ for $t < i \leq s$. Then we want to find an integer m of the form $m = p_1^2 \cdots p_t^2 \cdot p_{t+1} \cdots p_s$, where $p_1, \cdots, p_t$ are non-Wieferich primes and $p_1 \cdot p_2 \cdots p_s$ has associated solution tableau $\binom{e_1}{l_1} \cdots \binom{e_s}{l_s}$. Under GRH and the assumption that the number of Wieferich primes with $\mathrm{ord}_2(p)$ odd is $o(x/\log x)$, we are guaranteed that such an integer m exists. It can be found proceeding as in Example 1.

The integers of type II can be dealt with as in Example 3 (and, likewise, there are only finitely many possibilities for the smallest prime in the corresponding number $2n - 1$ and each of these cases can be analysed further using the solution tableau method).

This rather informal discussion can be turned into a formal proof of Theorem 6. We leave the details to the interested reader.

The examples together with some additional arguments then yield the correctness of the following table (for the notation see the beginning of Section 6).

Table 2: Sparse values of $S(n)$

$1+2e$	$\min_\mu(e)$	$\min_{\omega_1}(e)$	$\min_\omega(e)$	$r_\mu(e)$	$r_{\omega_1}(e)$	$r_\omega(e)$
5	–	–	7^2	0	0	1
13	–	–	31^2	0	0	1
29	–	–	631^2	0	0	1
37	–	–	127^2	0	0	1
53	–	–	14327^2	0	0	1
61	–	–	3391^2	0	0	1
101	$7\cdot 631$	$7\cdot 631$	$7\cdot 631$	2	2	2
109	–	$7^2\cdot 73$	$7^2\cdot 73$	0	1	2
149	–	–	11471^2	0	0	1
157	$71\cdot 631$	$71\cdot 631$	$71\cdot 631$	2	2	2
173	–	$71^2\cdot 1721$	$71^2\cdot 1721$	0	1	2
181	–	–	23671^2	0	0	1
197	$151\cdot 919$	$7^3\cdot 151$	$7^3\cdot 151$	2	2	2
229	–	–	248407^2	0	0	1

Indeed, using the method of solution tableaux we found that for $1+2e\le 771$ there exists an integer n with $\omega_1(n)\ge 1$ and $S(n)=2^e$, with the exception of $1+2e=5,13,29,37,53,61,149,181,229$ and the possible exception of $461,541$ and 757. Sometimes the corresponding values of n were quite large, for example, $i_2(7^{14}\cdot 73)=709$. Thus merely computing $i_2(n)$ over a large range of n will leave many values in the image unreached.

7 Analytic aspects

We will see that the scarcity of certain values of $S(n)$ is brought out by analytic number theory.

Note that $N_2(x)=\mathcal{P}_2(2x-1)$ and hence, by Theorem 2, we have under GRH that

$$N_2(x)=A\frac{x}{\log x}+O\left(\frac{x\log\log x}{(\log x)^2}\right).$$

Similarly we see that $n\in N_4$ iff $2n-1=p$ with $p\in\mathcal{P}_4$ or $2n-1=p^2$ with $p\in\mathcal{P}_2'$. On invoking Theorem 2 we obtain that, under GRH,

$$N_4(x)\le A\frac{\sqrt{2x}}{\log x}+O\left(\frac{\sqrt{x}\log\log x}{(\log x)^2}\right).$$

We actually conjecture that equality holds. Likewise, we conclude that $n\in N_8$ if $2n-1=p^3$ with $p\in\mathcal{P}_2'$ or $2n-1=p$ with $p\in\mathcal{P}_6$. Thus, under GRH, we deduce on invoking Theorem 2 that

$$N_8(x)=\frac{8Ax}{45\log x}+O\left(\frac{x\log\log x}{(\log x)^2}\right).$$

The asymptotic behaviour of $N_{16}(x)$ is dominated by the number of prime pairs (p, q) with $p < q$ such that $r_2(p) = r_2(q) = 2$, $(\mathrm{ord}_2(p), \mathrm{ord}_2(q)) = 1$ and $p \cdot q \leq 2x - 1$. We are inclined to believe that this number is a positive fraction of all pairs (p, q) with $p < q$ and such that $p \cdot q \leq 2x - 1$, which, as is well-known [7], grows asymptotically as $2x \log\log x / \log x$. Thus we are tempted to conjecture that

$$N_{16}(x) \sim c_0 \frac{x \log\log x}{\log x},$$

for some positive constant c_0.

The following result shows that the values 2^e for which $r_{\omega_1}(e) = 0$ deserve the predicate *sparse*.

Proposition 13 *If $r_{\omega_1}(e) = 0$, then $N_{2^e}(x) \ll \sqrt{x}$. If $r_{\omega_1}(e) \geq 1$, then*

$$\frac{x}{\log x} \ll N_{2^e}(x) \leq (\zeta(2) + o_e(1)) \frac{x}{\log x} (\log\log x)^{r_{\omega_1}(e)-1},$$

where the lower bound holds under the assumption of GRH.

Proof. In case $r_{\omega_1}(e) = 0$, we are counting a subset of the square-full numbers (a number is squarefull if $p|n$ implies $p^2|n$). As is well-known, cf. [31], the number of squarefull integers $m \leq x$ grows asymptotically as $\zeta(3/2)\sqrt{x}/\zeta(3)$. The proof of the lower bound is a consequence of Proposition 6 and Theorem 3. As for the upper bound, notice that $N_{2^e}(x)$ is bounded above by the number of integers $m \leq 2x-1$ such that $\omega_1(m) \leq r_{\omega_1}(e)$. Let C be an arbitrary positive constant. By the method of Sathe-Selberg [7, 32] we find that uniformly for $1 \leq k \leq C \log\log x$ we have

$$\#\{m \leq x : \omega_1(m) = k\} \sim F\left(\frac{k}{\log\log x}\right) \frac{x}{\log x} \cdot \frac{(\log\log x)^{k-1}}{(k-1)!}, \qquad (12)$$

where

$$F(z) = \frac{1}{\Gamma(z+1)} \prod_p \left(1 + \frac{z}{p} + \frac{1}{p^2 - 1}\right) (1 - \frac{1}{p})^z,$$

where as usual $\Gamma(z)$ denotes the gamma function. From this estimate the result follows on noting that $F(0) = \zeta(2)$. □

We think the upper bound in the latter result is much closer to the truth:

Conjecture 2 *Suppose that $r_{\omega_1}(e) \geq 1$, then*

$$\frac{x}{\log x} (\log\log x)^{r_{\omega_1}(e)-1} \ll N_{2^e}(x) \ll \frac{x}{\log x} (\log\log x)^{r_{\omega_1}(e)-1}, \ x \to \infty.$$

Acknowledgement. We thank Prof. E.W. Clark for pointing out the references given in the proof of Lemma 5 and Prof. Shallit for pointing out reference [34].

Part of the research of the first author was carried out whilst he was temporary assistant professor in the PIONIER-group of Prof. E. Opdam at the University of Amsterdam. He also likes to thank Dr. F. Breuer for bringing [3] to his attention.

References

[1] P. Alles, On a conjecture of Pelikán, *J. Combin. Theory Ser. A* **60** (1992), 312–313.

[2] T. Beth, D. Jungnickel and H. Lenz, *Design theory*, Vol. I, 2nd edition, Encyclopedia of Mathematics and its Applications, **69**, Cambridge University Press, Cambridge, 1999. xx+1100 pp..

[3] F. Breuer, E. Lötter and A. B. van der Merwe, Ducci sequences and cyclotomic polynomials, preprint.

[4] R. Crandall, K. Dilcher and C. Pomerance, A search for Wieferich and Wilson primes, *Math. Comp.* **66** (1997), 433–449.

[5] 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.

[6] R. K. Guy, *Unsolved problems in number theory*, Springer, 1994.

[7] A. Hildebrand, On the number of prime factors of an integer, *Ramanujan revisited* (Urbana-Champaign, Ill., 1987), 167–185, Academic Press, Boston, MA, 1988.

[8] N. F. J. Inglis and J. D. A. Wiseman, Very odd sequences, *J. Combin. Theory Ser. A* **71** (1995), 89–96.

[9] P. Langevin and P. Solé, Duadic codes over $\mathbf{Z}_4$, *Finite Fields Appl.* **6** (2000), 309–326.

[10] H. W. Lenstra Jr., On Artin's conjecture and Euclid's algorithm in global fields, *Invent. Math.* **42** (1977), 201–224.

[11] J.S. Leon, J.M. Masley and V. Pless, Duadic codes, *IEEE Trans. Inform. Theory* **30** (1984), 709–714.

[12] H. W. Lenstra, Jr., P. Moree and P. Stevenhagen, Character sums for primitive root densities, in preparation.

[13] R. Lidl and H. Niederreiter, *Finite fields*, Encyclopedia of Mathematics and its Applications **20**, Addison-Wesley, Reading, MA, USA, 1983. xx+755 pp..

[14] S. Ling and P. Solé, Duadic Codes over $\mathbb{F}_2 + u\mathbb{F}_2$, *Appl. Algebra Engrg. Comm. Comput.* **12** (2001), 365–379.

[15] F. J. MacWilliams and A. M. Odlyzko, Pelikan's conjecture and cyclotomic cosets, *J. Combin. Theory Ser. A* **22** (1977), 110–114.

[16] F. J. MacWilliams and N. J. A. Sloane, *The theory of error correcting codes*, (two volumes), North-Holland, Elsevier Science Publisher, 1977.

[17] P. Moree, On the divisors of $a^k + b^k$, *Acta Arith.* **80** (1997), 197–212.

[18] P. Moree, On the distribution of the order and index of $g(\text{mod } p)$ over residue classes, arXiv:math.NT/0211259, to appear in *J. Number Theory*.

[19] L. Murata, The distribution of primes satisfying the condition $a^{p-1} \equiv 1(\text{mod } p^2)$, *Acta Arith.* **58** (1991), 103–111.

[20] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers*, Second edition. Springer-Verlag, Berlin, 1990. xiv+746 pp..

[21] V. Pless, *Q*-codes, *J. Combin. Theory Ser. A* **43** (1986), 258–276.

[22] V. Pless, J.M. Masley and J.S. Leon, On weights in duadic codes, *J. Combin. Theory Ser. A* **44** (1987), 6–21.

[23] J. Pelikán, Contribution to "Problems", Colloq. Math. Soc. János Bolyai **10**, p. 1549, North-Holland, Amsterdam, 1975.

[24] T. D. Rogers, The graph of the square mapping on the prime fields, *Discrete Math.* **148** (1996), 317–324.

[25] J.J. Rushanan, Duadic codes and difference sets, *J. Combin. Theory Ser. A* **57** (1981), 254–261.

[26] J. H. Silverman, Wieferich's criterion and the *abc*-conjecture, *J. Number Theory* **30** (1988), 226–237.

[27] C. Small, Sums of three squares and levels of quadratic number fields, *Amer. Math. Monthly* **93** (1986), 276–279.

[28] M. Smid, Duadic codes, *IEEE Trans. Inform. Theory* **33** (1987), 432–433.

[29] M. Smid, *Duadic codes*, Master thesis, University of Eindhoven (1986).

[30] P. Stevenhagen, The correction factor in Artin's primitive root conjecture, *J. Théor. Nombres Bordeaux* **15** (2003), 383–391.

[31] D. Suryanarayana, On the order of the error function of the square-full integers, *Period. Math. Hungar.* **10** (1979), 261–271.

[32] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics **46**, Cambridge University Press, Cambridge, 1995. xvi+448 pp..

[33] D. Ulmer, Elliptic curves with large rank over function fields, *Ann. of Math.* **155** (2002), 295-315.

[34] T. Vasiga and J. Shallit, On the iteration of certain quadratic maps over $GF(p)$, *Discrete Math.* **277** (2004), 219–240.

[35] R. van der Veen and E. Nijhuis, Very odd sequences, Report on second year student project (Dutch, partly in English), 2004, University of Amsterdam, available at http://staff.science.uva.nl/∼alejan/project.html.

[36] S.S. Wagstaff, Jr., Pseudoprimes and a generalization of Artin's conjecture, *Acta Arith.* **41** (1982), 141–150.

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

CNRS-I3S, ESSI, Route des Colles, 06903 Sophia Antipolis, France.
e-mail: `ps@essi.fr`