VOL. 107

2007

NO. 1

PSEUDOPRIME CULLEN AND WOODALL NUMBERS

ВY

FLORIAN LUCA (México) and IGOR E. SHPARLINSKI (Sydney)

Abstract. We show that if a > 1 is any fixed integer, then for a sufficiently large x > 1, the *n*th Cullen number $C_n = n2^n + 1$ is a base *a* pseudoprime only for at most $O(x \log \log x/\log x)$ positive integers $n \le x$. This complements a result of E. Heppner which asserts that C_n is prime for at most $O(x/\log x)$ of positive integers $n \le x$. We also prove a similar result concerning the pseudoprimality to base *a* of the Woodall numbers given by $W_n = n2^n - 1$ for all $n \ge 1$.

1. Introduction. Let a > 1 be an integer. A base a pseudoprime is a composite integer n such that $n | a^n - a$. Note that the set consisting of primes and base a pseudoprimes is of asymptotic density zero, with main contribution coming from the primes (see Section 3 in [1]).

Distribution of pseudoprimes in various sequences has been considered in [5]. Here, we consider this question for two classical sequences, namely the sequence of *Cullen numbers* $C_n = n2^n + 1$ and the sequence of *Woodall* numbers $W_n = n2^n - 1$. We recall that C. Hooley [4] proves that C_n is prime for at most o(x) of positive integers $n \leq x$. This bound has been improved to $O(x/\log x)$ by Heppner [3] (see also [7] and [8]). Results on the counting function of the set of primes $p \leq x$ such that C_p is prime can be found in [3] and [6]. All the results mentioned above for the sequence $\{C_n\}_{n\geq 1}$ apply to the sequence $\{W_n\}_{n\geq 1}$ as well. We define

 $\mathcal{C}(x) = \{ n \le x : C_n \text{ is a base } a \text{ pseudoprime} \},\$

and show the following result which complements the aforementioned estimates.

THEOREM 1. We have the bound

 $#\mathcal{C}(x) = O(x \log \log x / \log x).$

Our approach uses the method of Hooley [4] in the refined form given by Heppner [3], but also includes some new arguments which allow us to study pseudoprime values of C_n rather than prime values.

²⁰⁰⁰ Mathematics Subject Classification: 11A07, 11B83, 11N25.

Key words and phrases: pseudoprime number, Cullen number, Woodall number.

Furthermore, one can easily check (see also the comment at the end of Section 3 of Chapter 8 of [4]) that with the method of [4] one can study primes in several more sequences, like $n2^n - 1$ or $2^n + n^2$. Quite on the contrary, our treatment of pseudoprime values is quite specific to the sequence C_n . However, for the sequence of Woodall numbers, setting

 $\mathcal{W}(x) = \{n \le x : W_n \text{ is a base } a \text{ pseudoprime}\},\$

we use even more ingredients and prove the following (somewhat weaker) estimate.

THEOREM 2. We have the bound

$$#\mathcal{W}(x) = O(x(\log\log x)^2/\log x).$$

2. Notation. Throughout this paper, we use the Vinogradov symbols \gg and \ll with their usual meaning. We recall that the conditions $U \ll V$ and $V \gg U$ are both equivalent to the assertion that U = O(V). The constants implied by them may depend on the base a.

For a positive real number x, we use $\log x$ for the maximum between the natural logarithm of x and 1. Furthermore, for every positive integer k, we write $\log_k x$ for the k-fold iteration of $\log x$.

The letters p, q and r always denote prime numbers, and the letters k and n always denote positive integers.

3. Proof of Theorem 1. It is shown in Lemma 2 of [3] that for any squarefree positive integer k, the number $N_k(x)$ of solutions of the congruence

 $C_n \equiv 0 \pmod{k}$ with $n \le x$

satisfies the estimate

1)
$$N_k(x) = \frac{x}{k} + O(\varphi(k))$$

(here, as usual, $\varphi(k)$ stands for the Euler function of k).

 Put

(

$$y = x^{1/10}$$
 and $z = (\log x)^{10}$.

The argument from the proof of Satz 1 in [3] (based on the estimate (1) and on the Selberg sieve) shows that the cardinality of the set $\mathcal{D}(x)$ of those $n \leq x$ such that C_n is free of primes from the interval [z, y] satisfies the bound

$$#\mathcal{D}(x) \ll x \prod_{z \le q \le y} \left(1 - \frac{1}{q}\right) + \sum_{d < y^2} \mu^2(d) 3^{\omega(d)} \varphi(d),$$

where $\omega(d)$ is the number of distinct prime factors of d. Since $\varphi(d)3^{\omega(d)} \leq$

 $d^2 \leq y^4$, we derive, by the Mertens formula, that

(2)
$$\#\mathcal{D}(x) \ll x \frac{\log z}{\log y} + y^6 \ll \frac{x \log_2 x}{\log x}.$$

From now on, we deal only with those $n \leq x$ such that $p \mid C_n$ for some prime $p \in [z, y]$. As we have noticed, for each fixed prime $p \in [z, y]$, there are x/p + O(p) integers $n \leq x$ with $C_n \equiv 0 \pmod{p}$. Let us denote by $\ell(p)$ the multiplicative order of a modulo p. Let $\mathcal{E}(x)$ be the set of $n \leq x$ which are divisible by at least one prime $p \in [z, y]$ with $\ell(p) \leq p^{1/3}$. We then have

(3)
$$\#\mathcal{E}(x) \ll x \sum_{\substack{p \in [z,y] \\ \ell(p) \le p^{1/3}}} \frac{1}{p} + y^2.$$

For a sufficiently large positive real t, putting

$$W = \prod_{1 \le s \le t^{1/3}} (a^s - 1),$$

we see that

$$\sum_{\substack{p \leq t \\ \ell(p) \leq p^{1/3}}} 1 \leq \omega(W) \ll \log W \ll t^{2/3}.$$

Therefore, by partial summation, we obtain

(4)
$$\sum_{\substack{p \in [z,y]\\ \ell(p) \le p^{1/3}}} \frac{1}{p} \ll z^{-1/3} \ll (\log x)^{-1}.$$

Thus, we derive from (3) that

(5)
$$\#\mathcal{E}(x) \ll x(\log x)^{-1}$$

Let $||k||_2$ denote the 2-adic part of k, that is, $||k||_2 = 2^s$, where the integer s is defined by $2^s | k$ and $2^{s+1} \nmid k$.

Let $\mathcal{F}(x)$ be the set of $n \leq x$ which are divisible by at least one prime $p \in [z, y]$ with $||p - 1||_2 \geq p^{1/6}$. By partial summation

$$\sum_{\substack{p \in [z,y] \\ \|p-1\|_2 \ge p^{1/6}}} \frac{1}{p} \ll \sum_{\substack{k \in [z,y] \\ \|k\|_2 \ge k^{1/6}}} \frac{1}{k} \ll z^{-1/6} \ll (\log x)^{-1}.$$

Thus, as before, we obtain

$$\#\mathcal{F}(x) \ll x(\log x)^{-1}.$$

Finally, let $\mathcal{G}(x) = \mathcal{C}(x) \setminus (\mathcal{D}(x) \cup \mathcal{E}(x) \cup \mathcal{F}(x)).$

We see that for every $n \in \mathcal{G}(x)$, there is a prime $p \in [z, y]$ with $p | C_n$ such that $\ell(p) > p^{1/3}$ and $||p - 1||_2 < p^{1/6}$.

For every such p, we see that if $n \in \mathcal{G}(x)$, then, since $p | C_n$ and C_n is a base a pseudoprime, we have $\ell(p) | n2^n$. Thus, $gcd(n, \ell(p)) > p^{1/6} \ge z^{1/6}$.

Hence, let us fix some integer $d \in [z, y]$ and some prime $p \in [z, y]$ with $d \mid p-1$. Then a slight modification of the argument from Section 3 of Chapter 8 of [4] which has led to a somewhat weaker version of (1), also implies that there are x/dp+O(p/d) integers $n \leq x$ such that $d \mid n$ and $p \mid C_n$. Indeed, we need to count the number of positive integers $m \leq x/d$ with $p \mid C_{dm}$. Let f = (p-1)/d. Writing m = kf + r in the unique way with integers k and r in the ranges $1 \leq r < f$, $0 \leq k \leq (x/d-r)/f = x/(p-1) - r/f$, we see that

$$0 \equiv C_{dm} \equiv (k(p-1) + dr)2^{k(p-1)+dr} + 1 \equiv (dr - k)2^{dr} + 1 \pmod{p}.$$

Thus, for every fixed value for r, the value of k is uniquely determined modulo p, and thus takes no more than x/(p-1)p + O(1) values in the above range. Hence, the total number of pairs (k, r) (and therefore the total number of m also) does not exceed

$$\frac{xf}{(p-1)p} + O(f) = \frac{x}{dp} + O\left(\frac{p}{d}\right),$$

as claimed.

Summing up over all such possibilities for d and p, we get

$$\#\mathcal{G}(x) \ll x \sum_{d \in [z,y]} \frac{1}{d} \sum_{\substack{p \in [z,y] \\ p \equiv 1 \, (\text{mod } d)}} \frac{1}{p} + O(y^2 \log y).$$

Clearly,

$$\sum_{\substack{p \in [z,y]\\p \equiv 1 \pmod{d}}} \frac{1}{p} \ll \sum_{\substack{k \le y\\k \equiv 1 \pmod{d}}} \frac{1}{k} \ll \frac{\log y}{d}$$

Hence,

$$\#\mathcal{G}(x) \ll x \log y \sum_{d \in [z,y]} \frac{1}{d^2} + O(y^3) \ll xz^{-1} \log y + x^{1/5} \log x \ll x(\log x)^{-1},$$

which finishes the proof.

4. Proof of Theorem 2. As in Lemma 2 of [3], for any squarefree positive integer k, the number $M_k(x)$ of solutions of the congruence

 $W_n \equiv 0 \pmod{k}$ with $n \le x$,

satisfies the estimate

(6)
$$M_k(x) = x/k + O(\varphi(k)).$$

We now set

$$y = x^{1/10}$$
 and $z = \exp(144(\log_2 x)^2)$.

As before, for a positive integer k coprime to a we use $\ell(k)$ to denote the multiplicative order of a modulo k. Further, for an odd positive integer k, we write $\kappa(k)$ for the multiplicative order of 2 modulo k.

We define the set $\mathcal{R}(x)$ of $n \leq x$ such that W_n is

- either free of primes in the interval [z, y],
- or divisible by at least one prime $p \in [z, y]$ with $\min\{\kappa(p), \ell(p)\} \le p^{1/3}$.

As in the proof of Theorem 1 (see the bounds (2) and (5) on $\#\mathcal{D}(x)$ and $\#\mathcal{E}(x)$, respectively), we obtain the bound

(7)
$$\#\mathcal{R}(x) \ll x \frac{\log z}{\log y} \ll \frac{x(\log_2 x)^2}{\log x}.$$

We now set

$$w = \exp(\sqrt{\log z}) = (\log x)^{12},$$

and use P(k) to denote the largest prime divisor of the positive integer k with the convention that P(1) = 1.

We let $\mathcal{S}(x)$ be the set of $n \leq x$ such that $n \notin \mathcal{R}(x)$ and W_n is divisible by at least one prime $p \in [z, y]$ with $\min\{P(\kappa(p)), P(\ell(p))\} \leq w$.

In particular, since $n \notin \mathcal{R}(x)$, we see that every "forbidden" prime factor p of n is such that $p \equiv 1 \pmod{d}$ for some $d \geq p^{1/3}$ and $P(d) \leq w$ (we take $d = \kappa(p)$ if $P(\kappa(p)) \leq P(\ell(p))$ and $d = \ell(p)$ otherwise). By estimate (6), we derive

$$\#\mathcal{S}(x) \ll x \sum_{\substack{z^{1/3} \le d \le y \\ P(d) < w}} \sum_{\substack{p < d^3 \\ p \equiv 1 \, (\text{mod } d)}} \frac{1}{p} + y^3.$$

We now recall the bound

(8)
$$\sum_{\substack{p < t \\ p \equiv 1 \pmod{k}}} \frac{1}{p} \ll \frac{\log_2 t}{\varphi(k)},$$

which is uniform in $2 \le k \le t$ and which follows from the Brun–Titchmarsh theorem after simple calculations (see, for example, inequality (3.1) in [2]). Applying (8) with k = d and $t = d^3$, and using the fact that $\varphi(d)/d \gg (\log_2 d)^{-1} \gg (\log_2 x)^{-1}$ for $d \in [1, x]$, we obtain

(9)
$$\sum_{\substack{z^{1/3} \le d \le y \\ P(d) < w}} \sum_{\substack{p < d^3 \\ p \equiv 1 \pmod{d}}} \frac{1}{p} \ll \sum_{\substack{z^{1/3} \le d \le y \\ P(d) < w}} \frac{\log_2 d}{\varphi(d)} \ll (\log_2 x)^2 \sum_{\substack{z^{1/3} \le d \le y \\ P(d) < w}} \frac{1}{d}.$$

For $2 \leq s \leq t$, we write

$$\Psi(t,s) = \#\{n \le t : P(n) \le s\}$$

It is known (see Chapter III in [9]) that the inequality

(10)
$$\Psi(t,s) \ll t \exp(-u/2), \quad \text{where } u = \frac{\log t}{\log s},$$

holds uniformly in $2 \le s \le t$. Furthermore, for s = w and $t \in [z^{1/3}, y]$, we have

$$u = \frac{\log t}{\log s} \ge \frac{\log(z^{1/3})}{\log w} = \frac{1}{3}\sqrt{\log z} = 4\log_2 x.$$

It now follows easily from (10) by partial summation that

$$\sum_{\substack{z^{1/3} \le d \le y \\ P(d) < w}} \frac{1}{d} \ll \exp(-2\log_2 x) \sum_{d < x} \frac{1}{d} \ll (\log x)^{-1},$$

which together with the estimate (9) gives

(11)
$$\#\mathcal{S}(x) \ll x(\log_2 x)^2(\log x)^{-1}.$$

We now let $\mathcal{T}(x)$ be the set of $n \leq x$ not in $\mathcal{R}(x) \cup \mathcal{S}(x)$ such that W_n has a factor $p \in [z, y]$ with $q = P(\ell(p)) \geq w$ but $\kappa(q) < q^{1/3}$. Fix a prime $q \geq w$ with $\kappa(q) < q^{1/3}$. For every $p \in [z, y]$ with $p \equiv 1 \pmod{q}$, the number of $n \leq x$ such that $p \mid W_n$ is x/p + O(p). Hence, summing up over all possible values of q and p, we get

$$\#\mathcal{T}(x) \le x \sum_{\substack{w \le q \le y \\ \kappa(q) < q^{1/3}}} \sum_{\substack{z \le p \le y \\ p \equiv 1 \pmod{q}}} \frac{1}{p} + O(y^3) \ll x \log_2 x \sum_{\substack{w \le q \le y \\ \kappa(q) < q^{1/3}}} \frac{1}{q} + x^{3/10},$$

where in the above estimate we have applied again the bound (8) to estimate the inner sums. An argument identical to the one which leads to the estimate (4) (just change a to 2 in the two estimates preceding (4)) shows that

$$\sum_{\substack{w \le q \le y \\ \kappa(q) < q^{1/3}}} \frac{1}{q} \ll w^{-1/3},$$

therefore

(12)
$$\#\mathcal{T}(x) \ll x(\log_2 x)w^{-1/3} + x^{3/10} \ll x(\log x)^{-1}$$

For a prime p we define

$$q_p = P(\ell(p)).$$

Now let $\mathcal{U}(x)$ be the set of $n \leq x$ not in $\mathcal{R}(x) \cup \mathcal{S}(x) \cup \mathcal{T}(x)$ such that W_n is a multiple of some prime $p \in [y, z]$ with

$$\operatorname{gcd}(\operatorname{lcm}[\kappa(p), \ell(p)], \kappa(q_p)) \ge \kappa(q_p)^{1/2}.$$

Let p be such a "special" prime and let

$$d_p = \gcd(\operatorname{lcm}[\kappa(p), \ell(p)], \kappa(q_p))$$

Thus, $d_p \ge \kappa(q_p)^{1/2}$. Since $n \notin \mathcal{R}(x) \cup \mathcal{S}(x) \cup \mathcal{T}(x)$, we have

 $\min\{\kappa(p), \ell(p)\} \ge p^{1/3} \ge z^{1/3}, \quad q_p \ge w, \quad \kappa(q_p) \ge w^{1/3}.$

Note that $d_p | \operatorname{lcm}[\kappa(p), \ell(p)] | p - 1$. Furthermore, $q_p | \ell(p) | p - 1$, and since $d_p | \kappa(q_p) | q_p - 1$, we deduce that d_p and q_p are coprime and are both divisors of p - 1. Thus, $d_p q_p | p - 1$. Furthermore, we also have $q_p \leq \kappa(q_p)^3 \leq d_p^6$.

We now fix $d \ge w^{1/6}$ and then a prime $q \le d^6$ with $q \equiv 1 \pmod{d}$.

For primes $p \in [z, y]$ with $d_p = d$ and $q_p = q$ we see that $p \equiv 1 \pmod{dq}$. Furthermore, once p is fixed, the number of $n \leq x$ such that $p \mid W_n$ is x/p + O(p). Summing up first over all p, then over all q, and then over all d, we get

$$\#\mathcal{U}(x) \le x \sum_{w^{1/6} \le d} \sum_{\substack{q \le d^6 \\ q \equiv 1 \, (\text{mod } d)}} \sum_{\substack{p \le y \\ p \equiv 1 \, (\text{mod } dq)}} \frac{1}{p} + O(y^4).$$

Applying the estimate (8) twice to estimate the inner sums above and using also the minimal order of the Euler function $\varphi(d)/d \gg (\log_2 x)^{-1}$ in the interval [1, x], we get

$$\sum_{w^{1/6} \le d} \sum_{\substack{q \le d^6 \\ q \equiv 1 \pmod{d}}} \sum_{\substack{p \le y \\ p \equiv 1 \pmod{dq}}} \frac{1}{p} \\ \ll \sum_{w^{1/6} \le d} \sum_{\substack{q \le d^6 \\ q \equiv 1 \pmod{d}}} \frac{\log_2 y}{q\varphi(d)} \ll \sum_{w^{1/6} \le d} \frac{\log_2 x}{\varphi(d)} \sum_{\substack{q \le d^6 \\ q \equiv 1 \pmod{d}}} \frac{1}{q} \\ \ll \log_2 x \sum_{w^{1/6} \le d} \frac{\log_2 d}{\varphi(d)^2} \ll (\log_2 x)^4 \sum_{w^{1/6} \le d} \frac{1}{d^2} \ll (\log_2 x)^4 w^{-1/6}.$$

Thus,

(13)
$$\#\mathcal{U}(x) \ll x(\log_2 x)^4 w^{-1/6} + x^{2/5} \ll x(\log x)^{-1}$$

Now let $\mathcal{V}(x)$ be the set of $n \leq x$ which do not belong to $\mathcal{R}(x) \cup \mathcal{S}(x) \cup \mathcal{T}(x) \cup \mathcal{U}(x)$ such that W_n is a base *a* pseudoprime.

Let $p \in [z, y]$ be such that $p \mid W_n$. We may assume furthermore that

$$\ell(p) \ge p^{1/3}, \quad q_p = P(\ell(p)) \ge w, \quad \kappa(q_p) \ge q_p^{1/3} \ge w^{1/3},$$

and that

$$d_p = \gcd(\operatorname{lcm}[\kappa(p), \ell(p)], \kappa(q_p)) < \kappa(q_p)^{1/2}$$

Let $f_p = \kappa(q_p)/d_p > \kappa(q_p)^{1/2} \ge w^{1/6}$ and put $m_p = \operatorname{lcm}[\kappa(p), \ell(p)]$.

We now count the number of $n \leq x$ such that $p | W_n$ and W_n is a base a pseudoprime. Let α be some positive integer and assume that $n \equiv \alpha \pmod{m_p}$. Then $n \equiv \alpha \pmod{\kappa(p)}$, therefore $2^n \equiv 2^\alpha \pmod{p}$. Since $p | W_n$, we get $n2^{\alpha} \equiv 1 \pmod{p}$, therefore $n \equiv 2^{-\alpha} \pmod{p}$. Hence, the residue of

n modulo m_p determines the residue of *n* modulo pm_p . Furthermore, since $p | W_n | a^{W_n} - a$, and *p* is large, we get $p | a^{W_n - 1} - 1$, therefore $\ell(p) | W_n - 1$. Thus, $q_p | \ell(p) | 2(n2^{n-1} - 1)$, and since $q_p > 2$, we get $q_p | n2^{n-1} - 1$. Since $q_p | \ell(p) | m_p$, it follows that *n* is already determined modulo q_p , and, in fact, $n \equiv \alpha \pmod{q_p}$. The above congruence now implies that $2^{n-1} \equiv \alpha^{-1} \pmod{q_p}$, which determines uniquely $n - 1 \mod \ell(q_p)$. Thus, *n* is determined modulo $lcm[pm_p, \ell(q_p)] = pm_p f_p$. To summarise, the congruence class of *n* modulo m_p determines *n* modulo $pm_p f_p$.

For each congruence class modulo m_p , there are therefore no more than $x/pm_pf_p+O(1)$ values of n in $\mathcal{V}(x)$. Summing up over all the residue classes modulo m_p , we deduce that the number of $n \in \mathcal{V}(x)$ which are multiples of p does not exceed $x/pf_p + O(m_p) \leq x/pw^{1/6} + O(y)$. Summing up over all the prime values of $p \in [z, y]$, we get

(14)
$$\#\mathcal{V}(x) \le \frac{x}{w^{1/6}} \sum_{p \in [z,y]} \frac{1}{p} + O(y^2) \ll \frac{x \log_2 x}{w^{1/6}} + x^{1/5} \ll x (\log x)^{-1},$$

which together with the estimates (7), (11), (12) and (13) completes the proof. \blacksquare

5. Remarks. As we have seen, the proof of Theorem 1 used the particular shape of the Cullen numbers, and in particular, the fact that $C_n - 1 = n2^n$ is a number which, multiplicatively, looks almost like a power of 2. On the other hand, the proof of Theorem 2 does not use this structure and achieves almost the same bound. It is now clear that the proof of Theorem 2 can be adapted to study pseudoprimality of other numbers of this shape, such as $2^n + n^2$ or, in general, $2^n f(n) \pm 1$ or $2^n \pm f(n)$, where $f(X) \in \mathbb{Q}[X]$ is any integer-valued polynomial with rational coefficients. In order to achieve this, one first needs analogues of Lemma 2 from [3] for these more general Cullen type numbers. We give no further details and leave this as an open problem. Following Heppner [3] and Mil'uolo [6], it would also be of interest to study the pseudoprimality of numbers of the type $p2^p \pm 1$, where p is prime. We leave this as a problem for further study as well.

REFERENCES

- [1] R. Crandall and C. Pomerance, *Prime Numbers: A Computational Perspective*, 2nd ed., Springer, New York, 2005.
- [2] P. Erdős, A. Granville, C. Pomerance and C. Spiro, On the normal behavior of the iterates of some arithmetic functions, in: Analytic Number Theory, Birkhäuser, Boston, 1990, 165–204.
- [3] F. Heppner, Über Primzahlen der Form $n2^n + 1$ bzw. $p2^p + 1$, Monatsh. Math. 85 (1978), 99–103.

- [4] C. Hooley, *Applications of Sieve Methods to the Theory of Numbers*, Cambridge Univ. Press, Cambridge, 1976.
- [5] F. Luca and I. E. Shparlinski, *Pseudoprime values of the Fibonacci sequence, polynomials and the Euler function*, Indag. Math., to appear.
- [6] A. Mil'uolo, On primes in sparse sequences, Vestnik Moskov. Univ. Ser. I Mat. Mekh. 1987, no. 2, 75–77 (in Russian); English transl.: Moscow Math. Univ. Bull. 42 (1987), no. 2, 63–65.
- [7] G. J. Rieger, Über Lipschitz-Folgen, Math. Scand. 45 (1979), 168–176.
- [8] H.-B. Sieburg, Siebmethoden bei dünnverteilten Primzahlfolgen, Arch. Math. (Basel) 39 (1982), 322–330.
- [9] G. Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, Cambridge Univ. Press, Cambridge, 1995.

Instituto de MatemáticasDepartment of ComputingUniversidad Nacional Autónoma de MéxicoMacquarie UniversityC.P. 58089, Morelia, Michoacán, MéxicoSydney, NSW 2109, AustraliaE-mail: fluca@matmor.unam.mxE-mail: igor@ics.mq.edu.au

Received 4 April 2006

(4747)