

A CONDITIONAL DENSITY FOR CARMICHAEL NUMBERS

THOMAS WRIGHT

(Received 15 June 2019; accepted 20 October 2019; first published online 13 February 2020)

Abstract

Under sufficiently strong assumptions about the first prime in an arithmetic progression, we prove that the number of Carmichael numbers up to X is $\gg X^{1-R}$, where $R = (2 + o(1)) \log \log \log \log X / \log \log \log X$. This is close to Pomerance's conjectured density of X^{1-R} with $R = (1 + o(1)) \log \log \log X / \log \log X$.

2010 *Mathematics subject classification*: primary 11A51; secondary 11N13.

Keywords and phrases: Carmichael number, Heath-Brown conjecture, least prime in an arithmetic progression.

1. Introduction: bounds

In 1910, Carmichael [3] discovered a numerical construct that now bears his name.

DEFINITION 1.1. A composite number m is a Carmichael number if $m \mid a^m - a$ for all $a \in \mathbb{Z}$.

Of course, if m is prime, it is always the case that $m \mid a^m - a$, so Carmichael numbers are often called *pseudoprimes*. The referee provided an additional historical comment. Carmichael was not the first to discover these numbers. Šimerka [12] discovered the first seven Carmichael numbers in 1885, but he published his work in a Czech journal that was not widely disseminated and his discovery was not known to most mathematicians at the time.

The search for Carmichael numbers was aided by an earlier result of Korselt [8].

KORSELT'S CRITERION. A composite number m is a Carmichael number if and only if m is square-free and $p - 1 \mid m - 1$ for every prime $p \mid m$.

Although Korselt devised the criteria for such a pseudoprime in 1899, he never computed any examples. It took another 11 years before Carmichael found the first of these numbers.

Once Carmichael numbers were discovered, attention quickly turned to the obvious follow-up question.

QUESTION 1.2. Let $C(X)$ be the number of Carmichael numbers up to X . How large is $C(X)$?

The first nontrivial lower bound for $C(X)$ came in 1994, when Alford *et al.* [1] proved that there are infinitely many Carmichael numbers. More specifically, they proved the following result.

THEOREM 1.3 (Alford, Granville and Pomerance, 1994). *Let $C(X)$ be as above. Then*

$$C(X) \gg X^{2/7}.$$

The current best result is by Harman [6].

THEOREM 1.4 (Harman, 2008). *Let $C(X)$ be as above. Then*

$$C(X) \gg X^{0.3336704}.$$

On the other side, upper bounds for $C(X)$ have been studied since the 1950s. In 1953, Knödel [7] proved the following result.

THEOREM 1.5 (Knödel 1953). *There exists a constant $k > 0$ such that*

$$C(X) \ll X e^{-k(\log X \log \log X)^{1/2}}.$$

Improvements by Erdős [5] and Pomerance [11] have brought this bound down.

THEOREM 1.6 (Pomerance 1981). *Let $C(X)$ be as above. Then*

$$C(X) \ll X e^{-\log X \log \log \log X / 2 \log \log X}.$$

Pomerance has conjectured that this upper bound should be close to the correct density for $C(X)$; in the paper mentioned above, he posited the following conjecture.

CARMICHAEL DENSITY CONJECTURE (Pomerance 1981). For $C(X)$ as above,

$$C(X) = X e^{-(1+o(1)) \log X \log \log \log X / \log \log X}.$$

In this paper, we provide further evidence for Pomerance's heuristic by showing that the assumption of a strong conjecture about the first prime in an arithmetic progression will yield a lower bound for $C(X)$ that is very close to his Carmichael density conjecture.

2. Introduction: methods

The result of Alford, Granville and Pomerance established a blueprint for proving density results about Carmichael numbers. The blueprint draws from two conjectures/theorems that we describe below.

In order to describe our conjectures/theorems, let us define the following notation. We will let $P(n)$ denote the largest prime factor of n ; we will let $\pi(x)$ denote the number of primes p up to x ; we will let $\pi(x, y)$ be the number of primes p up to x such that

$P(p-1) < y$; and we will let $\pi(x; d, a)$ denote the number of primes p up to x such that $p \equiv a \pmod{d}$.

With this notation in hand, we can state the two conjectures/theorems. Progress on these propositions is very closely linked to progress on the density of Carmichael numbers.

CONJECTURE/THEOREM 2.1. *Let $E \in (0, 1)$. Then there exists a constant $\gamma_E > 0$ depending only on E such that for all sufficiently large x ,*

$$\pi(x, x^{1-E}) \geq \gamma_E \pi(x).$$

This has been proved for certain values of E but is conjectured to be true for all $E < 1$.

CONJECTURE/THEOREM 2.2. *Let $B \in (0, 1)$. Then there exists a constant D_B depending only on B such that for each sufficiently large x , there exists a set $D_B(x)$ consisting of at most D_B integers where*

$$\pi(y; d, a) \geq \frac{\pi(y)}{2\phi(d)}$$

as long as $(a, d) = 1$, $d < \min\{x^B, y/x^{1-B}\}$ and d is not divisible by any of the integers in D_B . Moreover, all the elements in D_B must be of size at least $\log x$.

Again, this has been proved for some values of B but is conjectured to be true for every $B < 1$.

Theorem 1.3 from [1] can then be restated in the following way.

THEOREM 2.3 (Alford, Granville, Pomerance, 1994). *Choose an E and a B for which Conjectures/Theorems 2.1 and 2.2 are both true. Then*

$$C(X) \gg X^{EB}.$$

In particular, the conjectures/theorems are both proved theorems when $B = 5/12$ and $E = 1 - (2\sqrt{e})^{-1}$.

Subsequent improvements have largely followed this framework, generally by either improving the B and E or by slightly loosening the requirements of the conjectures/theorems to allow for improvement. The most recent result [6] takes $B = 0.4736$ and $E = 0.7039$, leading to the lower bound quoted above.

Of course, we expect that these conjectures are true for $E = 1 - \delta$ for any $\delta > 0$ and $B = 1 - \epsilon$ for any $\epsilon > 0$. Indeed, the first of these two conjectures/theorems would follow from something like the Elliott–Halberstam conjecture (with level of distribution $1 - E$), while the second is a much weaker form of Montgomery’s conjecture on primes in arithmetic progressions (see [10] for further discussion of Montgomery’s conjecture). As such, we are fairly confident that $C(X)$ should be $\gg X^{1-\epsilon}$.

The current paper, however, moves past power loss to draw closer to the cavalcade of logs appearing in Pomerance’s conjecture. To do this, we must take a different tack and replace these two conjectures with a single one.

3. Introduction: new results

To this end, we will invoke a conjecture that has commonly been used in the pursuit of Carmichael numbers and related problems.

HEATH-BROWN CONJECTURE. Let $(a, m) = 1$. There exists some constant A (independent of a and m) such if p is the smallest prime such that $p \equiv a \pmod{m}$ then $p \ll m \log^A m$.

This conjecture was first used by Banks, Ekstrom, Pomerance and Thakur in two papers [2, 4] that would prove (conditionally) that (a) there are infinitely many Carmichael numbers in arithmetic progressions (if the modulus and residue class are coprime) and (b) there are infinitely many composite square-free numbers m such that $p + 1 \mid m + 1$ for any prime p which divides m . Although these results were later made unconditional in [16] and [14], the conjecture has also been used in other Carmichael results, including [15].

Assuming this conjecture, we will prove the following result.

THEOREM 3.1 (Main theorem). *Assume the Heath-Brown conjecture above. Then*

$$C(X) \geq X^{1-(2+o(1)) \log \log \log X / \log \log \log X}.$$

Equivalently, we can say that

$$C(X) \geq X e^{-(2+o(1)) \log X \log \log \log X / \log \log \log X}.$$

This barely misses Pomerance's conjecture; it has an extra iteration of \log in the numerator and denominator of the exponent (as well as an extra 2).

We note that the Heath-Brown conjecture is actually slightly stronger than is needed; we use it as it is written here for purposes of expediency and clarity. In particular, we could prove the same result if we only assumed that the conjecture is true for primes equivalent to $1 \pmod{m}$, and we could even draw a similar result if we allowed the A to go to infinity sufficiently slowly (something like $A = \log \log m$). Moreover, if we were to take an even weaker version of the Heath-Brown conjecture (say, $p \ll m^{1+o(1)}$), our methods would still yield $C(X) = X^{1-o(1)}$.

4. Alford–Granville–Pomerance framework and proof framework

The general framework for finding infinitely many Carmichael numbers, as laid out in [1], is as follows. Again, let $P(n)$ denote the largest prime factor of n . In the traditional framework, one first finds that there are many primes q such that $P(q-1) < q^{1-\delta}$ for some δ . Multiplying these q 's together (apart from some small finite number of q 's, which we omit for exceptional zero reasons) yields a number L that has many factors and a small value for $\lambda(L)$ (where $\lambda(L)$ denotes the largest possible order of an element of $(\mathbb{Z}/L\mathbb{Z})^\times$). From this, one can find some k for which there are a large number of primes p where $p = dk + 1$ for various choices of $d \mid L$; if one has enough such primes p , there must exist some subset of these primes that multiply to $1 \pmod{Lk}$ and hence multiply to give a Carmichael number.

In order to use our conjecture (and remove the other two), we change the way in which we find these q 's. In particular, we will generate q 's in the same way we generate p 's. We take a number J that is the product of many small primes and we find an l such that there are many primes $q = gl + 1$ for various $g \mid J$. This will make $\lambda(L)$ much smaller, giving us more flexibility in generating our Carmichael numbers.

This technique was first used in [17] in a paper that conditionally addressed the question of whether there are infinitely many Carmichael numbers with a fixed number of prime factors. In that paper, the technique was useful because it kept $\lambda(L)$ really small while generating a large number of primes p . In this paper, the use is slightly different. The fact that we keep $\lambda(L)$ small means that we can keep our p 's and our Carmichael numbers relatively small, which allows for better density estimates.

It is interesting, if unfortunate, to note that these new methods cannot yet improve unconditional results for densities of Carmichael numbers. If one were to use our method of construction with the current known bounds and theorems, one would end up with $C(X) \gg X^{B^2}$; however, since the best known bound for B is less than that for E , this is a worse result than the original Alford–Granville–Pomerance result.

5. Finding q 's

First, for some sufficiently large natural number z , let us define J by

$$J = \prod_{\sqrt{z} < r < z, r \text{ prime}} r.$$

Let us consider the set of g 's where $g \mid J$. By the Heath-Brown conjecture, we can assume that the following assertion is true.

LEMMA 5.1. *For any $g \mid J$, there exists a q such that $q = gl + 1$ for some $l < \log^A g$.*

From this, we will winnow down our set of q 's to be a bit more helpful in our Carmichael search. As is standard, we let $\omega(x)$ denote the number of prime factors of x . So, for a given l , let us define the set of primes Q_l such that

$$Q_l = \{q \text{ prime} : q = gl + 1 \text{ for some } g \mid J \text{ with } \omega(g) = \lfloor \log z \rfloor\}.$$

Then we can prove the following result.

LEMMA 5.2. *There exists an $l_0 < \log^{2A} z$ such that*

$$|Q_{l_0}| > z^{\log z - (2+o(1)) \log \log z}.$$

PROOF. We know that the number of divisors of J is $> z/2 \log z$. So the number S of g 's for which $g \mid J$ and $\omega(g) = \lfloor \log z \rfloor$ is

$$S > \binom{z/2 \log z}{\lfloor \log z \rfloor} \gg \left(\frac{z}{2 \log^2 z} \right)^{\log z} = z^{\log z - (2+o(1)) \log \log z}.$$

Additionally, note that $g < z^{\log z}$ for any such g . So by Lemma 5.1, for any such g , there must exist a prime $q = gl + 1$ where

$$l < \log^A(z^{\log z}) = \log^{2A} z = z^{2A \log \log z / \log z}.$$

Since there are at least

$$z^{\log z - (2+o(1)) \log \log z}$$

choices for g and at most

$$\log^{2A} z$$

choices for l , there must exist an l_0 for which

$$|Q_{l_0}| \geq \frac{z^{\log z - (2+o(1)) \log \log z}}{\log^{2A} z} = z^{\log z - (2+o(1)) \log \log z}. \quad \square$$

We note for future reference that for any $q \in Q_{l_0}$,

$$q = gl_0 + 1 \leq z^{\log z + 2A \log \log z / \log z}. \tag{5.1}$$

6. Finding p 's

Now let us define

$$L = \prod_{q \in Q_{l_0}} q.$$

Let $\lambda(L)$ denote the largest possible order of a coprime residue mod L . We will require two pieces of information about L : the size of $\lambda(L)$ and the size of L itself.

LEMMA 6.1. *For λ as defined above,*

$$\lambda(L) < z^{((1+o(1))z + 2A \log \log z) / \log z}.$$

PROOF. For every $q \in Q_{l_0}$, we know that $q - 1 \mid Jl$. Since J is the product of fewer than $(1 + o(1))z / \log z$ primes up to z ,

$$J \leq z^{(1+o(1))z / \log z}.$$

Moreover, as we saw above,

$$l < z^{2A \log \log z / \log z}.$$

Multiplying these together gives us the lemma. □

LEMMA 6.2. *We have*

$$\log L < z^{\log z}.$$

PROOF. Recall that

$$q \leq z^{\log z + 2A \log \log z / \log z}.$$

Since the number of possible q 's is at most $z^{\log z - (2+o(1)) \log \log z}$,

$$L \leq z^{(\log z + 2A \log \log z / \log z)(z^{\log z - (2+o(1)) \log \log z})}.$$

So

$$\log L \leq z^{\log z - (2+o(1)) \log \log z} \left(\log z + 2A \frac{\log \log z}{\log z} \right) \log z < z^{\log z}$$

as required. □

Now, just as we did with J in the previous section, we will use this large product L to construct more primes. Let us consider primes p where $p = dk + 1$ for some $d \mid L$. In particular, for a given k , we define

$$\mathcal{P}_k = \{p = dk + 1 : p \text{ prime}, \omega(d) = z\}.$$

We can use the Heath-Brown conjecture to prove the following result. As before, z is a sufficiently large integer.

THEOREM 6.3. *There exists a $k_0 < (z \log^2 z + 2A(z \log \log z))^A$ such that*

$$|\mathcal{P}_{k_0}| > z^{z \log z - (2+o(1))z \log \log z}.$$

PROOF. First, recall that for every prime $q \mid L$,

$$q \leq z^{\log z + 2A \log \log z / \log z}.$$

If $\omega(d) = z$ then

$$d \leq z^{z \log z + 2Az \log \log z / \log z}.$$

So, for every d , there exists a k such that

$$k < \log^A z^{z \log z + 2Az \log \log z / \log z} = (z \log^2 z + 2A(z \log \log z))^A,$$

and $p = dk + 1$ is prime.

The number of possible d 's for which $\omega(d) = z$ is

$$\begin{aligned} \#\{d \mid L : \omega(d) = z\} &= \binom{z^{\log z - (2+o(1)) \log \log z}}{z} \\ &\geq \left(\frac{z^{\log z - (2+o(1)) \log \log z}}{z} \right)^z = z^{z \log z - (2+o(1))z \log \log z}. \end{aligned}$$

By the pigeonhole principle, there is a $k_0 < (z \log^2 z + 2A(z \log \log z))^A$ such that

$$|\mathcal{P}_{k_0}| \geq \frac{z^{z \log z - (2+o(1))z \log \log z}}{(z \log^2 z + 2A(z \log \log z))^A} = z^{z \log z - (2+o(1))z \log \log z}. \quad \square$$

Note that for a prime $p \in \mathcal{P}_{k_0}$,

$$p \leq (z^{\log z + 2A \log \log z / \log z})^z (z \log^2 z + 2A(z \log \log z))^A = z^{z \log z + (2A+o(1))z \log \log z / \log z}.$$

7. Carmichael number construction

In this section we can finally construct Carmichael numbers. To this end, we recall a theorem of van Emde Boas and Kruyswijk [13] and Meshulam [9]. Let $n(L)$ denote the smallest number such that a sequence of at least $n(L)$ elements in $(\mathbb{Z}/L\mathbb{Z})^\times$ must contain some nonempty sequence whose product is the identity.

THEOREM 7.1 [1, Theorem 1.5 and Proposition 2.1]. *Let $n(L)$ be as above. Then*

$$n(L) < \lambda(L)(1 + \log(\phi(L)/\lambda(L))).$$

Moreover, if $r > t > n(L)$, then any sequence of r elements in $(\mathbb{Z}/L\mathbb{Z})^\times$ contains at least $\binom{r}{t} / \binom{r}{n(L)}$ distinct subsequences of length at least $t - n(L)$ and at most t whose product is the identity.

In order to find the density of Carmichael numbers, we first need to know the size of $n(L)$ and $n(k_0L)$.

LEMMA 7.2. *For $n(L)$ as defined above, $n(L) < z^{2z/\log z}$.*

PROOF. From Lemma 6.1,

$$\lambda(L) < z^{((1+o(1))z+2A \log \log z)/\log z}.$$

Additionally, from Lemma 6.2, $\log L < z^{2 \log z}$. Putting this together gives

$$n(L) < z^{((1+o(1))z+2A \log \log z)/\log z} (1 + z^{2 \log z}) < z^{2z/\log z}. \quad \square$$

LEMMA 7.3. *We have*

$$n(k_0L) < z^{(2+o(1))z/\log z}.$$

PROOF. From the Heath-Brown conjecture, $k_0 < \log^A L$. Since

$$\lambda(k_0L) \leq k_0 \lambda(L) = \lambda(L)^{1+o(1)},$$

the lemma follows easily. □

Now we can use Theorem 7.1 to construct Carmichael numbers. First, we show that our construction actually yields the desired pseudoprimes.

THEOREM 7.4. *For r as defined in Theorem 7.1, let $p_1, p_2, \dots, p_r \in \mathcal{P}_{k_0}$ be distinct primes such that $m = p_1 p_2 \cdots p_r \equiv 1 \pmod{k_0L}$. Then m is a Carmichael number.*

PROOF. By construction, $p_i - 1 = dk_0 \mid Lk_0$ for every $p_i \in \mathcal{P}_{k_0}$. So for every $p_i \mid m$, we have $p_i - 1 = dk_0 \mid Lk_0 \mid m - 1$, which is Korselt’s criterion. □

Now that we know our construction yields Carmichael numbers, the next step is to determine how many such numbers our method yields.

THEOREM 7.5. *Let $C(X)$ denote the number of Carmichael numbers up to X . Then*

$$C(X) \geq X^{1-(2+o(1))\log \log \log \log X / \log \log \log X}.$$

Equivalently, we can say that

$$C(X) \geq X e^{-(2+o(1))\log X \log \log \log \log X / \log \log \log X}.$$

PROOF. From the preceding lemmas, $n(L) < z^{(2+o(1))z/\log z}$. We apply Theorem 7.1 with $t = z^z$ and $r = z^{z \log z - (2+o(1))z \log \log z}$. Clearly, $t < r < |\mathcal{P}_{k_0}|$. Let $I(z)$ denote the number of subsequences whose product is the identity. Then from Theorem 7.1,

$$\begin{aligned} I(z) &\gg \binom{z^z \log z - (2+o(1))z \log \log z}{z^z} \binom{z^z \log z - (2+o(1))z \log \log z}{z^{(2+o(1))z/\log z}} \\ &\gg \left(\frac{z^z \log z - (2+o(1))z \log \log z}{z^z} \right)^{z^z} \left(z^z \log z - (2+o(1))z \log \log z \right)^{z^{(2+o(1))z/\log z}} \\ &\gg (z^z \log z - (2+o(1))z \log \log z)^{z^z - z^{(2+o(1))z/\log z}} \\ &= z^{z^{z+1} \log z - (2+o(1))z^{z+1} \log \log z - z^{1+(2+o(1))z/\log z} \log z + (2+o(1))z^{1+(2+o(1))z/\log z} \log \log z} \\ &= z^{z^{z+1}(\log z - (2+o(1)) \log \log z)}. \end{aligned}$$

We recall from before that for all of the primes p ,

$$p \leq z^{z \log z + (2A+o(1))z \log \log z / \log z}.$$

By our construction $t = z^z$ and a Carmichael number m can consist of at most t such primes p , so

$$m \leq (z^{z \log z + (2A+o(1))z \log \log z / \log z})^{z^z} = z^{z^{z+1}(\log z + (2A+o(1))\log \log z / \log z)}.$$

Define

$$X = z^{z^{z+1}(\log z + (2A+o(1))\log \log z / \log z)},$$

so that

$$\begin{aligned} \log X &= z^{z+1}(\log^2 z + (2A + o(1)) \log \log z), \\ \log \log X &= z \log z + O(\log z), \\ \log \log \log X &= \log z + O(\log \log z), \\ \log \log \log \log X &= \log \log z + O(1). \end{aligned}$$

From this,

$$\begin{aligned} C(X) &\geq z^{z^{z+1}(\log z - (2+o(1)) \log \log z)} \\ &= z^{z^{z+1}(\log z + (2A+o(1))\log \log z / \log z)} z^{-z^{z+1}(2+o(1)) \log \log z} \\ &= X z^{-z^{z+1}(2+o(1)) \log \log z} \\ &= X (X^{-(2+o(1))(\log \log z) / (\log z + (2A+o(1))\log \log z / \log z)}) \\ &= X^{1-(2+o(1))\log \log z / \log z}. \end{aligned}$$

Since $(1 + o(1)) \log \log \log X = \log z$ and $(1 + o(1)) \log \log \log \log X = \log \log z$, this yields

$$C(X) \geq X^{1-(2+o(1))\log \log \log X / \log \log \log X}$$

which is as stated in the theorem. \square

Acknowledgement

I would like to thank the referee for some extremely helpful feedback.

References

- [1] W. R. Alford, A. Granville and C. Pomerance, ‘There are infinitely many Carmichael numbers’, *Ann. of Math. (2)* **139**(3) (1994), 703–722.
- [2] W. D. Banks and C. Pomerance, ‘On Carmichael numbers in arithmetic progressions’, *J. Aust. Math. Soc.* **88**(3) (2010), 313–321.
- [3] R. D. Carmichael, ‘Note on a new number theory function’, *Bull. Am. Math. Soc.* **16** (1910), 232–238.
- [4] A. Ekstrom, C. Pomerance and D. S. Thakur, ‘Infinitude of elliptic Carmichael numbers’, *J. Aust. Math. Soc.* **92** (2012), 45–60.
- [5] P. Erdős, ‘On pseudoprimes and Carmichael numbers’, *Publ. Math. Debrecen* **4** (1956), 201–206.
- [6] G. Harman, ‘Watt’s mean value theorem and Carmichael numbers’, *Int. J. Number Theory* **4**(2) (2008), 241–248.
- [7] W. Knödel, ‘Carmichaelsche Zahlen’, *Math. Nachr.* **9** (1953), 343–350.
- [8] A. Korselt, ‘Problème chinois’, *L’intermédiaire des mathématiciens* **6** (1899), 142–143.
- [9] R. Meshulam, ‘An uncertainty inequality and zero subsums’, *Discrete Math.* **84**(2) (1990), 197–200.
- [10] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory* (Cambridge University Press, Cambridge, 2006).
- [11] C. Pomerance, ‘On the distribution of pseudoprimes’, *Math. Comp.* **37**(156) (1981), 587–593.
- [12] V. Šimerka, ‘Zbytky z arithmetické posloupnosti (On the remainders of an arithmetic progression)’, *Časopis pro pěstování matematiky a fyziky* **14**(5) (1885), 221–225.
- [13] P. Van Emde Boas and D. Kruyswijk, *A Combinatorial Problem on Finite Abelian Groups. III*, *Zuivere Wiskunde*, 1969-008 (Stichting Mathematisch Centrum, Amsterdam, 1969).
- [14] T. Wright, ‘Infinitely many Carmichael numbers in arithmetic progressions’, *Bull. Lond. Math. Soc.* **45**(5) (2013), 943–952.
- [15] T. Wright, ‘Variants of Korselt’s criterion’, *Canad. Math. Bull.* **58**(4) (2015), 869–876.
- [16] T. Wright, ‘There are infinitely many elliptic Carmichael numbers’, *Bull. Lond. Math. Soc.* **50**(5) (2018), 791–800.
- [17] T. Wright, ‘Factors of Carmichael numbers and an even weaker k -tuples conjecture’, *Bull. Aust. Math. Soc.* **99**(3) (2019), 376–384.

THOMAS WRIGHT, 429 N. Church St., Spartanburg, SC 29302, USA
 e-mail: wrighttj@wofford.edu