J. Aust. Math. Soc. **94** (2013), 268–275
doi:10.1017/S1446788712000547

CARMICHAEL NUMBERS IN ARITHMETIC PROGRESSIONS

KAISA MATOMÄKI

(Received 31 August 2011; accepted 18 May 2012; first published online 8 March 2013)

Communicated by I. E. Shparlinski

Abstract

We prove that when $(a, m) = 1$ and a is a quadratic residue $\bmod m$, there are infinitely many Carmichael numbers in the arithmetic progression $a \bmod m$. Indeed the number of them up to x is at least $x^{1/5}$ when x is large enough (depending on m).

2010 *Mathematics subject classification*: primary 11N25; secondary 11A51.
Keywords and phrases: Carmichael number, arithmetic progression.

1. Introduction

Fermat's little theorem states that p divides $b^p - b$ for all integers b whenever p is prime. Composite numbers which satisfy this property are known as Carmichael numbers, named after R. D. Carmichael who began an in-depth study of them around 1910. The first Carmichael number is 561, and Carmichael suggested that perhaps there are infinitely many of them. This conjecture was finally proved in 1994 by Alford *et al.* [2] who even managed to show that there are at least $x^{2/7}$ Carmichael numbers up to x for any large enough x. This lower bound has been improved by Harman [10, 11] to x^{α} with α slightly larger than $1/3$.

It would be interesting to know if Carmichael numbers have similar distributional properties with primes. Two most natural questions are probably distribution in intervals and distribution in arithmetic progressions. Concerning the first question, it is not even known whether there is a Carmichael number between x and $2x$ for every large enough x.

Banks and Pomerance [5] recently studied the second question. They remarked that, for any m, proofs in [2, 10, 11] can be adapted to prove infinitude of Carmichael numbers that are $1 \bmod m$, whereas one does not even know that there exists an m for

Supported by the Academy of Finland grant no. 137883.

which there are infinitely many Carmichael numbers that are not 1 mod m. Banks and Pomerance [5] themselves gave a conditional proof that any arithmetic progression $a \bmod m$ with $(a, m) = 1$ contains infinitely many Carmichael numbers; they had to assume a conjecture on the least prime in an arithmetic progression, more precisely that for certain ξ and any $(b, d) = 1$ there exists a prime $p \ll d^{1+\xi/\log\log d}$ such that $p \equiv b \bmod d$. This conjecture is of course far away from the best available bound $p \ll d^{5.18}$ [13].

Here we will be able to give an unconditional proof in the case where a is a quadratic residue mod m.

THEOREM 1. *Let $(a, m) = 1$ and assume that a is a quadratic residue* mod m. *Then there are infinitely many Carmichael numbers in the arithmetic progression $a \bmod m$. Indeed the number of them up to x is at least $x^{1/5}$ when x is large enough (depending on m).*

This of course in particular implies that for any m for which there is an element in $(\mathbb{Z}/m\mathbb{Z})^*$ of multiplicative order greater than 2, there must indeed be infinitely many Carmichael numbers $\not\equiv 1 \bmod m$. This condition holds for any m which does not divide 24.

It probably is possible to improve the lower bound in the theorem, perhaps up to Harman's $1/3$, but here we wish to keep matters as simple as we can.

A convenient way to characterise Carmichael numbers is Korselt's elementary criterion which states that an integer n satisfies Fermat's property for every base b if and only if n is square-free and $p - 1 \mid n - 1$ for every prime $p \mid n$. We will employ the same strategy as in [2] which has its roots in Erdős's original heuristic [9]. The basic idea is to construct an integer L for which there are many primes p for which $p - 1 \mid L$. Suppose now that some product of these primes, say $C = p_1 \cdots p_k$, is 1 mod L. Then C is a Carmichael number by Korselt's criterion: each $p_i - 1 \mid L \mid C - 1$.

The extra trouble we have here is that we have to find such products C that have the additional property $C \equiv a \bmod m$. Banks and Pomerance [5] handled this additional requirement by instead finding $C \equiv 1 \bmod Lm$ and a prime $p_0 \equiv a \bmod m$ such that $p_0 C$ is also a Carmichael number (a similar idea was apparently independently present in [6]). However, in order to choose such C and p_0, they needed the above-mentioned conjecture. Here we will instead directly find $C \equiv a \bmod m$ through bringing in the Baker–Schmidt theorem [4]. A defect of our method is that it only works for quadratic residue a, the reason for which will be explained in the end of the paper.

2. Preliminary results

In this section we provide lemmas needed in the proof of the theorem and try to describe how they fit into the brief proof sketch given in the introduction.

The number L mentioned in the introduction will be chosen to be a product of certain primes q for which $q - 1$ has only small prime factors. With such a choice the maximal order of an element in $(\mathbb{Z}/L\mathbb{Z})^*$ is relatively small, which will make

https://doi.org/10.1017/S1446788712000547 Published online by Cambridge University Press

finding C easier. Obviously we have to first show that such primes q even exist. Writing $P(n)$ for the largest prime factor of n and

$$\pi_{b,d}(x, y) = |\{\tfrac{1}{2}x \le p \le x\colon\ p \equiv b \bmod d \text{ and } P(p-1) \le y\}|,$$

we can prove the following lemma.

Lemma 2. *Let b and d be fixed coprime integers with $d > 0$. For any $\alpha > 1/2$, there exist $\gamma(\alpha) > 0$ and $x_1(\alpha, d)$ such that*

$$\pi_{b,d}(x, x^\alpha) \ge \gamma(\alpha)\frac{x}{\phi(d)\log x}$$

for all $x \ge x_1(\alpha, d)$.

Proof. We can clearly assume that $\alpha < 2/3$. Choose a positive $\epsilon = \epsilon(\alpha) < \alpha - 1/2$. Now if $p \le x$ is such that $p = 1 + qk$ for some prime $q \in [x^{1-\alpha}, x^{1/2-\epsilon}]$, then $P(p-1) \le x^\alpha$. Each p has at most two such representations. Hence

$$\pi_{b,d}(x, x^\alpha) \ge \frac{1}{2}\sum_{\substack{x^{1-\alpha} < q < x^{1/2-\epsilon}\\ q\in\mathbb{P}}}\ \sum_{\substack{x/2 \le p \le x\\ p\equiv 1 \bmod q\\ p\equiv b \bmod d}} 1.$$

Since $(q, d) = 1$ when x is large enough, the congruence conditions can be combined into a single congruence $\bmod dq$, so, by the Bombieri–Vinogradov theorem,

$$\begin{aligned}\pi_{b,d}(x, x^\alpha) &\ge \frac{1+o(1)}{2\phi(d)}\sum_{\substack{x^{1-\alpha} < q < x^{1/2-\epsilon}\\ q\in\mathbb{P}}}\frac{x}{2\phi(q)\log x} + O(x(\log x)^{-10})\\ &\ge \frac{1}{8}\log\Big(\frac{\frac{1}{2}-\epsilon}{1-\alpha}\Big)\cdot\frac{x}{\phi(d)\log x}\end{aligned}$$

for every large enough x. □

Stronger results in the case $d = 1$ have appeared in the literature, the strongest one [3] working in the range $\alpha > 0.2961$ (albeit with a slightly weaker lower bound). We do not attempt to generalise these results to suit our needs but use this weaker result to keep our argument as transparent as possible.

For a given large L we cannot actually show that there are many primes p for which $p - 1 \mid L$. However, it is possible to show that there must be a relatively small k for which $k \mid p - 1 \mid kL$ for many primes p. The following lemma is a straightforward generalisation of [2, Theorem 3.1] (see also [1, Proposition 1.5] for a similar generalisation with slightly different assumptions).

Lemma 3. *Let $B < 5/12$ and let b and d be coprime integers with $d \ge 1$. Then there exist positive constants $c_0(B)$ and $x_0(B, d)$ such that the following holds. If $x \ge x_0(B, d)$ and if L is a square-free integer which is coprime to d, not divisible by any prime*

https://doi.org/10.1017/S1446788712000547 Published online by Cambridge University Press

exceeding $x^{(1-B)/2}$ *and for which* $\sum_{p|L} 1/p \le (1-B)/40$, *then there exists a positive integer* $k \le x^{1-B}$ *such that* $(k, L) = 1$ *and*

$$|\{l \mid L\colon kl+1 = p \in \mathbb{P},\ p \le x \text{ and } p \equiv b \bmod d\}| \ge \frac{c_0(B)}{\phi(d)\log x}|\{l \mid L\colon 1 \le l \le x^B\}|.$$

Next we turn to generating products which are 1 mod L. For this we need some notation. For a finite (multiplicative) abelian group G, let $n(G)$ denote the length of the longest sequence of (not necessarily distinct) elements of G for which no nonempty subsequence has product the identity. Further, let $\lambda(G)$ be the maximum order of the elements of G. It is easy to see that $\lambda(G) \le n(G) \le |G|$. An important ingredient in [2] was the following better upper bound for $n(G)$ due to van Emde Boas and Kruyswijk [8] and Meshulam [12].

Lemma 4. *One has*

$$n(G) \le \lambda(G)\Big(1 + \log\frac{|G|}{\lambda(G)}\Big).$$

In order to get quantitative results, one needs to know how many identity products there are. To this end one can use the following combinatorial consequence of the previous lemma [2, Proposition 1.2].

Lemma 5. *Let* G *be a finite abelian group and let* $r > t > n(G)$ *be integers. Then any sequence of* r *elements of* G *contains at least* $\binom{r}{t}/\binom{r}{n(G)}$ *subsequences of length at most* t *and at least* $t - n(G)$, *whose product is the identity.*

However, as we want our product to be $a \bmod m$ we are not solely interested in subsequences with product identity but want our products to equal a certain other element of a group. One cannot have a direct generalisation of Lemma 4 with the identity just replaced by any element of G—if a sequence is contained in a subgroup, only elements of that subgroup appear as products. However, there is a generalisation which takes this into account: Baker and Schmidt [4] bound $n(B)$ using a different method which gives a kind of uniform distribution result. From [4, Proposition 1] and Lemma 5 we can easily deduce the following lemma.

Lemma 6. *For any multiplicative abelian group* G *write*

$$s(G) = \lceil 5\lambda(G)^2\Omega(|G|)\log(3\lambda(G)\Omega(|G|))\rceil.$$

Let A *be a sequence of length* n *consisting of nonidentity elements of* G. *Then there exists a nontrivial subgroup* $H \le G$ *such that:*

(i) *if* $n \ge s(G)$, *then, for every* $h \in H$, $A \cap H$ *has a subsequence whose product is* h;
(ii) *if* t *is an integer such that* $s(G) < t < n - n(G)$, *then, for every* $h \in H$, A *has at least* $\binom{n-n(G)}{t-n(G)}/\binom{n}{n(G)}$ *distinct subsequences of length at most* t *and at least* $t - n(G)$ *whose product is* h.

https://doi.org/10.1017/S1446788712000547 Published online by Cambridge University Press

Proof. The first part follows immediately from [4, Proposition 1] by changing from additive to multiplicative setting. Therefore we only have to show that H which satisfies (i) also satisfies (ii). To prove this we can use Lemma 5 after some preparations that resemble the beginning of the proof of that lemma in [2].

Let $h \in H$. Then by (i), there is a subsequence R whose product is h. Let U be the longest such subsequence, with length u, say. Then $u \geq n - n(G)$ since otherwise, by Lemma 4, $A \setminus U$ contains a subsequence whose product is 1 and this subsequence could be appended to U to increase its size. Let $t' = u - t + n(G)$. Then $t' \geq n - t > n(G)$ and $t' < u - s(G) + n(G) < u$, so by Lemma 5 the sequence U contains

$$\binom{u}{t'} \Big/ \binom{u}{n(G)} = \binom{u}{t - n(G)} \Big/ \binom{u}{n(G)} \geq \binom{n - n(G)}{t - n(G)} \Big/ \binom{n}{n(G)}$$

distinct subsequences of length at most t' and at least $t' - n(G)$ whose product is the identity. When any such subsequence is removed from U we obtain a subsequence of A with desired properties, and the claim follows. □

3. Proof of Theorem 1

Let $\theta < 2$, $B < 5/12$ and let y be a large positive parameter. Note that the bounds for θ and B come from Lemmas 2 and 3. Let

$$\mathcal{Q} = \{q \in \mathbb{P} \cap (\tfrac{1}{2}y^{\theta}, y^{\theta}) \colon q \equiv -1 \bmod \phi(m) \text{ and } P(q-1) \leq y\}.$$

Notice that when y is large enough, $(q, m) = 1$ for every $q \in \mathcal{Q}$. By Lemma 2 and a trivial estimate,

$$\gamma(1/\theta)\frac{y^{\theta}}{\phi(\phi(m))\log y^{\theta}} \leq |\mathcal{Q}| \leq \frac{y^{\theta}}{\log y^{\theta}}$$

for all sufficiently large y. Let

$$L = \prod_{q \in \mathcal{Q}} q \quad \text{and} \quad x = e^{y^{1+\delta}}$$

for some small positive constant δ.

By our assumption that a is a quadratic residue mod m, there exists a_0 such that $a_0^2 \equiv a \bmod m$. Applying Lemma 3 and arguing as in [2, Proof of Theorem 4.1], we can find an integer $k \leq x^{1-B}$ coprime to L for which the set

$$\mathcal{P} = \{p \leq x \colon p = dk + 1 \text{ for some } d \mid L \text{ and } p \equiv a_0 \bmod m\} \tag{1}$$

has cardinality at least $x^{1/5+1/500}$ when δ is small enough. Taking $\mathcal{P}' = \mathcal{P} \setminus \mathcal{Q}$,

$$|\mathcal{P}'| \geq x^{1/5+1/1000} = e^{(\frac{1}{5}+\frac{1}{1000})y^{1+\delta}}.$$

Next we borrow some notation from [5]. Let $\mathcal{N}$ be the set of integers n such that $\gcd(n, Lm) = 1$ and $n \equiv 1 \bmod k$ (in particular $\mathcal{P}' \subset \mathcal{N}$). Let G be the subgroup

https://doi.org/10.1017/S1446788712000547 Published online by Cambridge University Press

of $(\mathbb{Z}/kL\mathbb{Z})^* \times (\mathbb{Z}/m\mathbb{Z})^*$ consisting of pairs (α, β) with $\alpha \equiv 1 \bmod k$, and define $\Psi\colon \mathcal{N} \to G$ to be the natural map which takes each integer $n \in \mathcal{N}$ to the pair

$$\Psi(n) = (n \bmod kL, n \bmod m).$$

If now $\mathcal{S} \subset \mathcal{P}'$, $|\mathcal{S}| > 1$ and $n_{\mathcal{S}} = \prod_{p\in\mathcal{S}} p$ is such that $\Psi(n) = (1, a)$, then $n_{\mathcal{S}}$ is a Carmichael number and $n_{\mathcal{S}} \equiv a \bmod m$. Indeed, for any $p \mid n_{\mathcal{S}}$ we have $p \in \mathcal{P}'$, so that $p - 1 \mid kL \mid n - 1$, which implies that $n_{\mathcal{S}}$ is a Carmichael number by Korselt's criterion.

Our last task is to find such subsets $\mathcal{S}$. To this end we will use Lemma 6. Notice that, for sufficiently large y and with our choice of G,

$$\lambda(G) \le \phi(m) \operatorname*{lcm}_{p\mid L}[p-1] \le \phi(m) \prod_{p\le y} y^{\theta} \le \phi(m)(y^{\theta})^{y/\log y} \le e^{2\theta y},$$
$$\log|G| \le \log(Lm) \le 2|\mathcal{Q}| \log(y^{\theta}) \le 2y^{\theta}$$

and

$$\Omega(|G|) = \Omega(\phi(L)\phi(m)) \le 2\log(Lm) \le 4y^{\theta},$$

so that, by Lemmas 4 and 6,

$$n(G) \le e^{3\theta y} \quad \text{and} \quad s(G) \le e^{5\theta y}.$$

We can think of $p \in \mathcal{P}'$ as elements of G through the map Ψ. Let $r \in \mathbb{N}$ be such that

$$r \equiv \begin{cases} 0 & \bmod \lambda((\mathbb{Z}/L\mathbb{Z})^*) \\ 2 & \bmod \phi(m). \end{cases}$$

Such r exists since

$$\gcd(\lambda((\mathbb{Z}/L\mathbb{Z})^*), \phi(m)) = \gcd\Big(\operatorname*{lcm}_{q\in\mathcal{Q}}[q-1], \phi(m)\Big) = \operatorname*{lcm}_{q\in\mathcal{Q}}[\gcd(q-1, \phi(m))] = 2 \qquad (2)$$

by definition of $\mathcal{Q}$.

Now apply Lemma 6 to the sequence $(\Psi(p))_{p\in\mathcal{P}'}$. This lemma asserts that the subgroup H from the lemma contains some member of the sequence, so that at least one $\Psi(p)$ with $p \in \mathcal{P}'$ must be in H. Thus H contains $\Psi(p^r) = (1, a_0^2) = (1, a)$, so that Lemma 6 shows that a desired set $\mathcal{S}$ exists.

To get the quantitative bound, we use part (ii) of Lemma 6 with $t = e^{y^{1+\delta/2}}$. We find that the number of Carmichael numbers $n_{\mathcal{S}} \le x^t$ that are $a \bmod m$ is at least

$$\binom{|P'| - n(G)}{t - n(G)} \Big/ \binom{|P'|}{n(G)} \ge \left(\frac{|P'| - n(G)}{t - n(G)}\right)^{t-n(G)} \Big/ |P'|^{n(G)} \ge |P'|^{t(1-o(1))} \ge x^{t/5}$$

which finishes the proof of Theorem 1. □

https://doi.org/10.1017/S1446788712000547 Published online by Cambridge University Press

One naturally asks whether the argument can be extended to the case where a is a quadratic nonresidue. Arguing as here but replacing the nonexisting a_0 in (1) by a (or any root of a), we would need $r \bmod \phi(m)$ to be odd but $r \bmod \lambda((\mathbb{Z}/L\mathbb{Z})^*)$ to be even, which is impossible because of (2). Of course one could allow more different residue classes mod m in (1), but this does not completely remove the problem. Going back to the proof of our Lemma 6(i) in [4], one sees that it would be enough to show that (extended) $\mathcal{P}'$ cannot be 'almost contained' in a subgroup which does not contain $(0, a)$. This seems very likely but difficult to prove. Thinking of G as $\oplus_{q\in Q}(\mathbb{Z}/\phi(q)\mathbb{Z}) \oplus (\mathbb{Z}/\phi(m)\mathbb{Z})$, an example of a troublesome large subgroup would be

$$\Big\{(a_1, \dots, a_{|Q|+1}) \in \bigoplus_{q\in Q}(\mathbb{Z}/\phi(q)\mathbb{Z}) \oplus (\mathbb{Z}/\phi(m)\mathbb{Z}) \colon \sum_{i=1}^{|Q|+1} a_i \equiv 0 \bmod 2\Big\}.$$

Ekstrom *et al.* [7] have developed the arguments in [2, 5, 6] further to give a conditional proof of the infinitude of elliptic Carmichael numbers. Unfortunately our methods do not seem to be directly helpful in that problem.

Acknowledgements

The author would like to thank Carl Pomerance and the referee for helpful comments on this paper.

References

[1] W. R. Alford, A. Granville and C. Pomerance, 'On the difficulty of finding reliable witnesses', in: *Algorithmic Number Theory (Ithaca, NY, 1994)*, Lecture Notes in Computer Science, 877 (Springer, Berlin, 1994), 1–16.
[2] W. R. Alford, A. Granville and C. Pomerance, 'There are infinitely many Carmichael numbers', *Ann. of Math.* (2) **139**(3) (1994), 703–722.
[3] R. C. Baker and G. Harman, 'Shifted primes without large prime factors', *Acta Arith.* **83**(4) (1998), 331–361.
[4] R. C. Baker and W. M. Schmidt, 'Diophantine problems in variables restricted to the values 0 and 1', *J. Number Theory* **12**(4) (1980), 460–486.
[5] W. D. Banks and C. Pomerance, 'On Carmichael numbers in arithmetic progressions', *J. Aust. Math. Soc.* **88**(3) (2010), 313–321.
[6] A. Ekstrom, 'On the infinitude of elliptic Carmichael numbers', PhD Thesis, University of Arizona, 1999.
[7] A. Ekstrom, C. Pomerance and D. S. Thakur, 'Infinitude of elliptic Carmichael numbers', *J. Aust. Math. Soc.* **92**(1) (2012), 45–60.
[8] P. van Emde Boas and D. Kruyswijk, 'A combinatorial problem on finite Abelian groups III', Afd. Zuivere Wisk., 1969-008, Math. Centrum, Amsterdam.
[9] P. Erdös, 'On pseudoprimes and Carmichael numbers', *Publ. Math. Debrecen* **4** (1956), 201–206.
[10] G. Harman, 'On the number of Carmichael numbers up to x', *Bull. Lond. Math. Soc.* **37**(5) (2005), 641–650.
[11] G. Harman, 'Watt's mean value theorem and Carmichael numbers', *Int. J. Number Theory* **4**(2) (2008), 241–248.

[12] R. Meshulam, 'An uncertainty inequality and zero subsums', *Discrete Math.* **84**(2) (1990), 197–200.
[13] T. Xylouris, 'On the least prime in an arithmetic progression and estimates for the zeros of Dirichlet *L*-functions', *Acta Arith.* **150**(1) (2011), 65–91.

KAISA MATOMÄKI, Department of Mathematics,
University of Turku, 20014 Turku, Finland
e-mail: ksmato@utu.fi

https://doi.org/10.1017/S1446788712000547 Published online by Cambridge University Press