There are infinitely many Carmichael numbers

by

W. R. Alford, Andrew Granville and Carl Pomerance

Department of Mathematics, University of Georgia, Athens, Georgia 30602, U.S.A.

dedicated to Paul Erdős on the occasion of his 80^{th} birthday

Introduction.

On October 18th, 1640, Fermat wrote in a letter to Frenicle, that whenever p is prime, p divides $a^{p-1} - 1$ for all integers a not divisible by p, a result now known as Fermat's 'little theorem'. An equivalent formulation is the assertion that p divides $a^p - a$ for all integers a, whenever p is prime. The question naturally arose as to whether the primes are the only integers exceeding 1 that satisfy this criterion, but Carmichael [Ca1] pointed out in 1910 that 561 ($= 3 \times 11 \times 17$) divides $a^{561} - a$ for all integers a. In 1899, Korselt [Ko] had noted that one could easily test for such integers by using (what we will call)

Korselt's criterion: *n divides $a^n - a$ for all integers a if and only if n is squarefree and $p - 1$ divides $n - 1$ for all primes p dividing n.*

In a series of papers around 1910, Carmichael began an in-depth study of composite numbers with this property, which have become known as *Carmichael numbers*. In [Ca2], Carmichael exhibited an algorithm to construct such numbers and stated, perhaps somewhat wishfully, that *"this list* (of Carmichael numbers) *might be indefinitely extended"*. Indeed, until now, no one has been able to prove that there are infinitely many Carmichael numbers, though it has long seemed highly likely.

In 1939 Chernick noted that if $p = 6m + 1, q = 12m + 1$ and $r = 18m + 1$ are all prime then pqr is a Carmichael number. According to Hardy and Littlewood's widely believed prime k-tuplets conjecture, these should simultaneously be prime infinitely often, which would tell us that there are infinitely many Carmichael numbers.

As yet unpublished computations of Richard Pinch have yielded 8,241 Carmichael numbers up to 10^{12}, 19,279 up to 10^{13}, 44,706 up to 10^{14} and 105,212 up to 10^{15}. On the other hand, numerous authors have supplied upper bounds for $\mathrm{C}(x)$, the number of Carmichael numbers up to x, the best being ([PSW], though also see [Po])

$$\mathrm{C}(x) \le x^{1-\{1+o(1)\} \log\log\log x / \log\log x}$$

for $x \to \infty$. We believe that this upper bound probably gives the true size of $\mathrm{C}(x)$. Our belief can be justified by the heuristic argument in [Po], which is based on ideas of Erdős [Er2].

In this paper we show that $\mathrm{C}(x) > x^\alpha$ for all large x and some positive constant α. In particular, we may take $\alpha = 2/7$. A precise upper bound for allowable values of α in our theorem depends on two other constants that appear in analytic number theory. We now describe these constants.

Let $\pi(x)$ be the number of primes $p \le x$, and let $\pi(x,y)$ be the number of these for which $p-1$ is free of prime factors exceeding y. Let $\mathcal{E}$ denote the set of numbers E in the range $0 < E < 1$ for which there exist numbers $x_1(E)$, $\gamma_1(E) > 0$ such that

$$\pi(x, x^{1-E}) \ge \gamma_1(E)\pi(x) \tag{0.1}$$

for all $x \ge x_1(E)$. Erdős (see [Er1]) proved that there is a small positive number in $\mathcal{E}$. Larger values were subsequently found by Wooldridge, Goldfeld, Pomerance, Fouvry and Grupp, Balog, and Friedlander. Currently the best result known ([Fr]) is that any positive number less than $1 - (2\sqrt{\mathrm{e}})^{-1}$ is in $\mathcal{E}$. Erdős has conjectured that any positive number less than 1 is in $\mathcal{E}$; that is, that $\mathcal{E}$ is the open interval $(0,1)$.

We remark that it is easy to see that if $E \in \mathcal{E}$, then $(0,E] \subset \mathcal{E}$. In addition one can show (using the Brun–Titchmarsh inequality) that if $E \in \mathcal{E}$ then $E' \in \mathcal{E}$ for some $E' > E$. That is, $\mathcal{E}$ is an open interval. We give the proof in Section 5.

Define $\pi(x;d,a)$ to be the number of primes up to x that belong to the arithmetic progression a mod d. The prime number theorem for arithmetic progressions states that

$$\pi(x;d,a) \sim \pi(x)/\varphi(d) \qquad \text{for} \quad x \to \infty, \tag{0.2}$$

provided $(a,d) = 1$, where φ is Euler's function. An important problem in analytic number theory is to enquire into the possible dependence on d and a in this asymptotic relation. For example, may d also tend to infinity as x does and if so, how fast? It is conjectured that (0.2) holds uniformly for all coprime integer pairs a, d with $1 \le d \le x^{1-\varepsilon}$, for any fixed $\varepsilon > 0$. Assuming the Riemann hypothesis for Dirichlet L-functions this conjecture can be proved for the more restricted range $1 \le d \le x^{1/2-\varepsilon}$. However, the strongest unconditional such result known is the Siegel-Walfisz theorem, which asserts that (0.2) holds uniformly for all coprime integer pairs a, d with $1 \le d \le (\log x)^k$, for any fixed k.

If one is prepared to disregard multiples of a possible 'exceptional' modulus, then one can significantly improve the range in the Siegel-Walfisz theorem. In fact, if $\psi(x)$ tends to 0 arbitrarily slowly then (0.2) holds for all coprime integer pairs a and d with $1 \le d \le x^{\psi(x)}$, except possibly for those d which are multiples of some integer $d_1(x)$, which exceeds a power of $\log x$ (see page 55 of [Bo]). If, in addition one is willing to relax the asymptotic relation in (0.2) and settle for a lower bound of the correct order of magnitude, then one can take $1 \le d \le x^B$ for some small $B > 0$. One can get larger values of B by allowing more exceptional moduli. Specifically, let $\mathcal{B}$ denote the set of numbers B in the range $0 < B < 1$ for which there is a number $x_2(B)$ and a positive integer D_B such that if $x \ge x_2(B)$, $(a,d) = 1$ and $1 \le d \le \min\{x^B, y/x^{1-B}\}$ then

$$\pi(y; d, a) \ge \frac{\pi(y)}{2\varphi(d)} \tag{0.3}$$

whenever d is not divisible by any member of $\mathcal{D}_B(x)$, a set of at most D_B integers, each of which exceeds $\log x$. In section 2 we show that the interval $(0, 5/12) \subset \mathcal{B}$, which follows from a bound for the density of zeros of Dirichlet L-functions, due to Huxley [Hu] and Jutila [Ju]. Although no result exactly like Theorem 2.1 has been proved in the literature, it was known to be feasible by the experts.

Our theorem on Carmichael numbers depends intimately on the sets $\mathcal{E}$ and $\mathcal{B}$.

Theorem 1. *For each $E \in \mathcal{E}$ and $B \in \mathcal{B}$ there is a number $x_0 = x_0(E, B)$ such that $\mathrm{C}(x) \ge x^{EB}$ for all $x \ge x_0$.*

Since $(0, 1 - (2\sqrt{\mathrm{e}})^{-1}) \subset \mathcal{E}$ and $(0, 5/12) \subset \mathcal{B}$, we conclude that $\mathrm{C}(x) \ge x^{\beta-\varepsilon}$ for any $\varepsilon > 0$ and all large x depending on the choice of ε, where

$$\beta = (1 - (2\sqrt{\mathrm{e}})^{-1})\frac{5}{12} = .290306\ldots.$$

This implies that, as stated above, $\mathrm{C}(x) > x^{2/7}$ for all large x.

Our argument is based on Erdős's original heuristic [Er2], though with certain modifications. The idea is to construct an integer L for which there are a very large number of primes p such that $p-1$ divides L. Suppose that the product of some of these primes, say $C = p_1 \cdots p_k$, is congruent to 1 mod L. Then C is a Carmichael number, since each $p_j - 1$ divides L which divides $C - 1$, and we may apply Korselt's criterion above. Indeed the more such products we can find, the more Carmichael numbers we will have constructed. How large a set of such primes p must we have to guarantee the existence of such products? We may view these primes p as elements of the group $(\mathbf{Z}/L\mathbf{Z})^*$ of reduced residues mod L. The following result, due to van Emde Boas and Kruyswijk (and extending a theorem independently due to Kruyswijk and Olson), gives a partial answer.

Theorem 2. *If G is a finite abelian group in which the maximal order of an element is m, then in any sequence of at least $m(1 + \log(|G|/m))$ (not necessarily distinct) elements of G, there is a non-empty subsequence whose product is the identity.*

We give a simplified proof of this result in the next section.

So as to be able to apply Theorem 2 to finding Carmichael numbers by our proposed method, we will need to find an integer L, with at least

$$\lambda(L)\left(1+\log\frac{\varphi(L)}{\lambda(L)}\right)\geq\lambda(L)$$

primes p for which $p-1$ divides L. Here, Carmichael's lambda function $\lambda(L)$ (see [Ca1]) is the largest order of an element in $(\mathbf{Z}/L\mathbf{Z})^*$. However the number of such primes p cannot exceed $\tau(L)$, the number of divisors of L (since each such p is 1 plus a divisor of L), and usually $\lambda(L)$ is much larger than $\tau(L)$ (see [EPS]). To avoid this problem we will pick our L so that $\lambda(L)$ is surprisingly small, while, at the same time, there are many primes p for which $p-1$ divides L. To do this, we select L to be the product of certain primes q for which the prime factors of $q-1$ are all at most y. This is how a number $E\in\mathcal{E}$ enters into the proof.

Prachar [Pr] (see [APR]) showed that there are infinitely many integers m with more than $2^{c\log m/\log\log m}$ divisors of the form $p-1$, p prime. Here $c>0$ is some constant that depends on a number $B\in\mathcal{B}$. One cannot do much better, since $\tau(m)\leq 2^{(1+o(1))\log m/\log\log m}$ for all m as $m\to\infty$. Prachar's method is to take a number L which is the product of all of the primes up to some point and show that there is some integer k with $k<L^{c'}$ and with $m=kL$ having many divisors of the form $p-1$. For our purposes, we need $\lambda(kL)$ to be inordinately small in comparison to kL. But the introduction of the mysterious factor k may ruin things, for there is no reason why $\lambda(kL)$ cannot be fairly large, even if we started with an L for which $\lambda(L)$ is very small in comparison to L. In Section 3 we will modify Prachar's method, so that now, given L, we can find an integer k coprime with L such that there are many primes $p\equiv 1 \bmod k$ for which $p-1$ divides kL. The advantage of this over Prachar's construction is that we may still apply Theorem 2 with $G=(\mathbf{Z}/L\mathbf{Z})^*$, since each of these primes p is in the subgroup of $(\mathbf{Z}/kL\mathbf{Z})^*$ of residue classes that are 1 mod k, and this subgroup is isomorphic to $(\mathbf{Z}/L\mathbf{Z})^*$.

As mentioned above, it has been conjectured that $\mathcal{E}=(0,1)$ and that (0.2) holds uniformly for all coprime pairs a,d with $1\leq d\leq x^{1-\varepsilon}$, for any fixed $\varepsilon>0$ (and so $\mathcal{B}=(0,1)$). Assuming these conjectures, we see that Theorem 1 implies Erdős's conjecture that $\mathrm{C}(x)\geq x^{1-\varepsilon}$ for any $\varepsilon>0$ and all sufficiently large x (depending on the choice of ε). Actually, we can show that one need only assume that $\mathcal{B}=(0,1)$, for in Section 5 we will prove the following result.

Theorem 3. *For each $B\in\mathcal{B}$, $(0,B)\subset\mathcal{E}$.*

We remark that, for the proofs of Theorems 1 and 3, one only needs a weaker version of the definition of $\mathcal{B}$, where a is restricted to the value 1. In particular, we record the following result.

Theorem 4. *Let $\varepsilon > 0$. Suppose there is a number x_ε such that*

$$\pi(x; d, 1) \geq \frac{\pi(x)}{2\varphi(d)}$$

for all positive integers $d \leq x^{1-\varepsilon}$, once $x \geq x_\varepsilon$. Then there is a number x'_ε such that $\mathrm{C}(x) \geq x^{1-2\varepsilon}$ for all $x \geq x'_\varepsilon$. In particular, if such an x_ε exists for each $\varepsilon > 0$, then $\mathrm{C}(x) = x^{1-o(1)}$ for $x \to \infty$.

Our proof of Theorem 1 is effective in the sense that if numerical values are given for $\gamma_1(E)$, $x_1(E)$, and $x_2(B)$, then following our arguments, a numerical value for $x_0(E, B)$ can be produced. However, the larger values of E that we now know to be in $\mathcal{E}$ are proved to be in $\mathcal{E}$ via the ineffective Bombieri–Vinogradov theorem. It is possible that Friedlander's theorem that every positive number $E < 1 - (2\sqrt{e})^{-1}$ is in $\mathcal{E}$ could be proved from a weaker, but effective version of this theorem, but we do not take up this issue here. It is interesting to note that Erdős's original proof that $\mathcal{E}$ contains some positive number E uses only Brun's method and is thus effective. Our proof in Section 2, that every positive number $B < 5/12$ is in $\mathcal{B}$, is effective. Further, from our proof of Theorem 3, we thus have that values for $\gamma_1(E)$ and $x_1(E)$ are computable for every positive number $E < 5/12$. We thus have the following theorem.

Theorem 5. *For each number α in the range $0 < \alpha < 25/144$, there is a computable number $x(\alpha)$ such that $\mathrm{C}(x) \geq x^\alpha$ for all $x \geq x(\alpha)$.*

It may also be of interest to actually compute a numerical value for $x(\alpha)$ for some specific $\alpha > 0$, but this may be difficult.

It has long been known how to construct infinitely many pseudoprimes for any given base a (that is, composite numbers n which divide $a^n - a$). The best lower bound in the literature had been [Po] that if $E \in \mathcal{E}$, then the number of base a pseudoprimes up to x is at least

$$\exp\left((\log x)^{\frac{E}{E+1}}\right)$$

for all large x depending on the choice of E and a. Evidently this result is majorized by Theorem 1.

Until now Duparc's problem [Du] as to whether there are infinitely many numbers that are simultaneously pseudoprime to both bases 2 and 3 was unsolved, but this follows from Theorem 1.

Our proof shows there are Carmichael numbers with arbitrarily many prime factors, but we have not been able to show that there are infinitely many Carmichael numbers with a fixed number of prime factors. We cannot show that there are infinitely many Carmichael numbers n divisible by some fixed prime factor, nor even with $\varphi(n)/n < 1 - \varepsilon$ for some fixed $\varepsilon > 0$. Our proof is easily modified to show that there are arbitrarily large sets of

Carmichael numbers such that the product of any subset is itself a Carmichael number. It seems to be difficult to prove a 'Bertrand's postulate for Carmichael numbers', that is, that there is always a Carmichael number between x and $2x$ once x is sufficiently large.

One can modify our proof to show that for any fixed non-zero integer a, there are infinitely many squarefree, composite integers n such that $p-a$ divides $n-1$ for all primes p dividing n. However, we have been unable to prove this for $p-a$ dividing $n-b$, for b other than 0 or 1. Such questions have significance for variants of pseudoprime tests, such as the Lucas probable prime test (see [PSW], [Wi]), strong Fibonacci pseudoprimes (see [LMO]) and elliptic pseudoprimes (see [GP]).

Our proof can also be modified to show that, for any given finite set $\mathcal{S}$ of positive integers, there are infinitely many integers n which are strong pseudoprimes to every base in $\mathcal{S}$, as well as being Carmichael numbers. (We say a positive odd integer n is a "strong pseudoprime to the base a" if n is composite and either $a^u \equiv 1 \bmod n$ or $a^{2^i u} \equiv -1 \bmod n$ for some integer $i < t$, where $n-1 = 2^t u$ and u is odd. It is known that if n is odd and composite, then n fails to be a strong pseudoprime for at least three fourths of the integers a in $\{1, 2, \ldots, n-1\}$.) The primality test programmed into some software packages is based on the given integer passing strong pseudoprime tests to each base in a fixed finite set $\mathcal{S}$. It was widely suspected that no matter how large the set $\mathcal{S}$ is taken, there will always be composite numbers that are passed off as prime by the test. Our result confirms this view and in fact we can show that the number of such integers up to x is greater than $x^{2/7}$, for large x.

We intend to take up these and other questions in a future paper.

Throughout the paper the letters p and q shall always denote primes. The constants $c_1, c_2, \ldots$ are all positive, and will always be assumed to be absolute (not dependent on any variable), as well as computable. We shall use both $|\ |$ and $\#$ to denote cardinality of a set, reserving the latter symbol for sets written with braces.

Acknowledgements. The idea for this paper came to us after seeing a preprint of Zhang Mingzhi [Zh] in which the Erdős heuristic is used to give numerical examples of Carmichael numbers with many prime factors. We are endebted to Ed Azoff, Roger Baker, Brian Boe, Enrico Bombieri, Paul Erdős, John Friedlander, Roger Heath-Brown, Sergei Konyagin, Helmut Maier, Greg Martin, Hugh Montgomery, François Morain, Gary Mullen, Jean-Louis Nicolas, Richard Pinch, John Selfridge, Jeff Shallit, Bob Vaughan and Richard Warlimont for their comments and advice concerning this paper. The second and third authors wish to acknowledge support from NSF grant DMS 90–02538.

1. Subsequence products representing the identity in a group.

If G is a group of order m, then any sequence of m elements of the group contains a subsequence whose product is 1, the identity. For if the sequence is $g_1, g_2, \cdots, g_m$, then the $m+1$ products 1, g_1, $g_1 g_2, \cdots, g_1 g_2 \cdots g_m$ cannot all be distinct (as there are only m distinct group elements) and if none of the latter m products is 1, we get $g_1 \cdots g_i = g_1 \cdots g_j$ for some $i < j$, so that $g_{i+1} \cdots g_j = 1$. This result cannot be improved for $G = C_m$, a cyclic group of order m, since if g is a generator of C_m and $g_1 = g_2 = \cdots = g_{m-1} = g$, then no subproduct is 1.

For a finite group G, let $n(G)$ denote the length of the longest sequence of (not necessarily distinct) elements of G for which no non-empty subsequence has product the identity. Kruyswijk [Ba] and Olson [Ol] independently evaluated $n(G)$ when G is a finite abelian p-group. Baker and Schmidt [BS] gave good upper bounds for $n(G)$ for arbitrary finite abelian groups and for significant generalizations of this problem, and van Emde Boas and Kruyswijk [EK] and Meshulam [Me] each gave the result in Theorem 2. We now restate this theorem and give a simplified proof based on that in [EK].

Theorem 1.1. *If G is a finite abelian group and m is the maximal order of an element in G, then $n(G) < m(1 + \log(|G|/m))$.*

Proof. Let $g_1, g_2, \ldots, g_n$ be a sequence of elements of G and assume that $n \geq m(1 + \log(|G|/m))$. Choose q to be any prime with $q \equiv 1 \bmod m$ and let $\mathbf{F}_q$ denote the field of q elements. If we multiply out the product

$$(a_1 - g_1)(a_2 - g_2) \ldots (a_n - g_n) = \sum_{g \in G} k_g g$$

in the group ring $\mathbf{F}_q[G]$, where $a_1, a_2, \ldots, a_n \in \mathbf{F}_q^*$, and suppose that no subsequence of $g_1, g_2, \ldots, g_n$ has product equal to 1, then $k_1 = a_1 a_2 \ldots a_n$. Thus if we can find $a_1, a_2, \ldots, a_n \in \mathbf{F}_q^*$ such that

$$(a_1 - g_1)(a_2 - g_2) \ldots (a_n - g_n) = 0, \tag{1.1}$$

then $k_1 = 0$ and we have a contradiction; implying that, in fact, there must be a subsequence whose product is 1.

Any character $\chi : G \to \mathbf{F}_q^*$ in the character group $\hat{G}$, may be extended to a ring homomorphism $\chi : \mathbf{F}_q[G] \to \mathbf{F}_q$ by letting $\chi(\sum_{g \in G} k_g g) = \sum_{g \in G} k_g \chi(g)$. From the orthogonality relations for group characters, one can show that if $b \in \mathbf{F}_q[G]$ then $b = 0$ if and only if $\chi(b) = 0$ for all $\chi \in \hat{G}$. Thus, since $\chi(\prod_{i=1}^n (a_i - g_i)) = \prod_{i=1}^n (a_i - \chi(g_i))$, we see that (1.1) holds for a given choice of $a_1, a_2, \ldots, a_n \in \mathbf{F}_q^*$ if

(1.2) *for each $\chi \in \hat{G}$ there exists j, $1 \leq j \leq n$, such that $\chi(g_j) = a_j$.*

Therefore it suffices to show that one may select $a_1, a_2, \ldots, a_n \in \mathbf{F}_q^*$ so that (1.2) holds. To do this, we shall proceed by the "greedy algorithm" of picking a_1 so that $\chi(g_1) = a_1$ holds for as many $\chi \in \hat{G}$ as possible, picking a_2 so that $\chi(g_2) = a_2$ holds for as many of the *remaining* $\chi \in \hat{G}$ as possible, and so on. The key observation is that each $\chi(g_j)$ is an mth root of 1 in $\mathbf{F}_q$, and so can be one of only m different values. Thus if $\mathcal{S}$ is any subset of $\hat{G}$ and g is any element of G, then there is some $a \in \mathbf{F}_q^*$ with $\chi(g) = a$ holding for at least $|\mathcal{S}|/m$ characters $\chi \in \mathcal{S}$. That is, $\chi(g) = a$ does *not* hold for at most $|\mathcal{S}|(1 - 1/m)$ characters $\chi \in \mathcal{S}$. Thus applying the greedy algorithm sequentially to $g_1, g_2, \ldots, g_k$, where $k = [m \log(|G|/m)] + 1$, we may choose $a_1, a_2, \ldots, a_k \in \mathbf{F}_q^*$ so that the residual set of characters $\chi \in \hat{G}$ with $\chi(g_j) \neq a_j$, for each $j = 1, 2, \ldots, k$, has cardinality at most

$$|\hat{G}|(1 - 1/m)^k = |G|(1 - 1/m)^k < |G|\mathrm{e}^{-k/m} < m.$$

Call the remaining characters $\chi_1, \chi_2, \ldots, \chi_r$, where $0 \leq r \leq m - 1$. Since $n \geq k + m - 1 \geq k + r$, we still have $a_{k+1}, a_{k+2}, \ldots, a_{k+r}$ remaining to be chosen. We choose them by letting $a_{k+j} = \chi_j(g_{k+j})$ for $j = 1, 2, \ldots, r$. If $k + r < n$, we may choose the remaining a_j's as arbitrary members of $\mathbf{F}_q^*$. Thus (1.2) holds and the theorem is proved.

Remark. It is reported in [Ol] that at the Midwestern Conference on Group Theory and Number Theory at Ohio State University, April 1966, Davenport asked for the best possible bound in Theorem 1.1, since this gives the largest number of prime (ideal) divisors that can divide an irreducible integer in an algebraic number field with class group G. For this and other applications, it is still of great interest to get the best possible result above. Our argument here may be sharpened to give the bound $m(\gamma + \varepsilon + \log(|G|/m))$ provided m and $|G|/m$ are each sufficiently large (as a function of ε), for any given $\varepsilon > 0$, where γ is the Euler-Mascheroni constant.

The next result allows us to construct many such products.

Proposition 1.2. *Let G be a finite abelian group and let $r > t > n = n(G)$ be integers. Then any sequence of r elements of G contains at least $\binom{r}{t}/\binom{r}{n}$ distinct subsequences of length at most t and at least $t - n$, whose product is the identity.*

Proof. Let R be a sequence of r elements of G. Since $r > n$ there is, by the definition of $n(G)$, some subsequence of R whose product is 1. Let S be the longest such subsequence, with cardinality s, say. Then $s \geq r - n$, since otherwise $R \setminus S$ contains a subsequence whose product is 1, and this subsequence might be appended to S, increasing its size, which contradicts the maximality of S.

Let T be *any* subsequence of S of cardinality $t - n$. If the product of the elements of T is g then the product of the elements of $S \setminus T$ is g^{-1}. Let U be the smallest (possibly empty) subsequence of $S \setminus T$ whose product is g^{-1}. Evidently U has cardinality at most

n else, by hypothesis, there exists a subsequence of U that has product 1 and this can be removed from U to make it smaller.

So $V = T \cup U$ is a subsequence of S (and thus R), in which the product of the elements is 1, and which has size at most $(t-n)+n = t$ and at least $t-n$.

The number of ways of choosing such a pair of sequences (T,U) is at least the number of ways of choosing T and is thus at least $\binom{s}{t-n}$. The maximum possible number of different sequences T which give rise to the same sequence $V = T \cup U$ is at most $\binom{|V|}{t-n} \le \binom{t}{t-n} = \binom{t}{n}$. Therefore the number of different subsequences V that we have created is at least

$$\binom{s}{t-n} \Big/ \binom{t}{n} \ \ge \ \binom{r-n}{t-n} \Big/ \binom{t}{n} \ = \ \binom{r}{t} \Big/ \binom{r}{n}.$$

This completes the proof of Proposition 1.2.

2. Primes in arithmetic progressions.

For each Dirichlet character χ and real numbers σ, T in the ranges $1/2 \le \sigma \le 1, T \ge 0$, let $\mathrm{N}(\sigma, T, \chi)$ be the number of zeros $s = \beta + i\gamma$ of the Dirichlet L-function $\mathrm{L}(s,\chi)$ inside the box $\sigma \le \beta \le 1$ and $|\gamma| \le T$. Let $\mathcal{A}$ be the set of real numbers $A > 2$ for which there exists a number $\gamma_2(A) \ge 1$, such that for all $\sigma \ge 1 - 1/A$ and $T \ge 1$,

$$\mathrm{N}(\sigma, T, d) := \sum_{\chi \bmod d} \mathrm{N}(\sigma, T, \chi) \ \le \ \gamma_2(A)(Td)^{A(1-\sigma)}, \tag{2.1}$$

for all positive integers d. One form of the 'density hypothesis for Dirichlet L-functions' asserts that every number $A > 2$ is in $\mathcal{A}$. The best that is currently known unconditionally is that every $A > 12/5$ is in $\mathcal{A}$; this may be deduced by combining the 'log-free' bound of Jutila [Ju] with a result of Huxley [Hu]. In principle these proofs are 'effective', so that one can compute a value for $\gamma_2(A)$ for each $A > 12/5$. Note that (2.1) cannot hold for any $A \le 2$ (with $\sigma = 1/2$), since the number of zeros of $\mathrm{L}(s,\chi)$ up to height T in the critical strip is of order of magnitude $T\log(Td)$ – see [Da], Chapter 16. In particular, there is a computable constant $c_1 \ge 1$ such that

$$\mathrm{N}(1/2, T, d) \le c_1 T d \log(Td) \tag{2.2}$$

for each integer $d \ge 1$ and number $T \ge 1$. Note that (2.2) gives a better result than (2.1) for fixed σ in the range $1/2 \le \sigma < 1 - 1/A$.

One may easily deduce, from the following result, that if $A \in \mathcal{A}$ then $B \in \mathcal{B}$ for all B satisfying $0 < B < 1/A$. In particular, since $(12/5, \infty) \subset \mathcal{A}$, we have $(0, 5/12) \subset \mathcal{B}$.

Theorem 2.1. *For any given $A \in \mathcal{A}$ and $\varepsilon, \delta > 0$, there exist numbers $x_{\varepsilon,\delta}, D_{\varepsilon,\delta}, \eta_{\varepsilon,\delta} > 0$ such that whenever $x \geq x_{\varepsilon,\delta}$ there is a set $\mathcal{D}_{\varepsilon,\delta}(x)$, of at most $D_{\varepsilon,\delta}$ integers, for which*

$$\left| \sum_{\substack{p \leq y \\ p \equiv a \bmod d}} \log p - \frac{y}{\varphi(d)} \right| \leq \varepsilon \frac{y}{\varphi(d)}$$

whenever d is not divisible by any element of $\mathcal{D}_{\varepsilon,\delta}(x)$, with $(a,d) = 1$ and d in the range $1 \leq d \leq \min\{x^{1/A-\delta}, y/x^{1-1/A+\delta}\}$. Furthermore, every number in $\mathcal{D}_{\varepsilon,\delta}(x)$ exceeds $\log x$, and all, but at most one, exceeds $x^{\eta_{\varepsilon,\delta}}$.

Proof. We shall only prove the result when ε and δ are extremely small (depending on the choice of A), since the result then immediately follows for all larger values of ε and δ. When $1 \leq d \leq \log y$ our result is a consequence of the prime number theorem for arithmetic progressions (see [Da], p. 123, eq. (9) and the following display). From this and the hypothesis, we note that we need only consider values of x, y and d in the ranges

$$\log y \leq d \leq \min\{x,y\}^{1/A-\delta} \quad \text{and} \quad \log^{4/\delta} x < x^{1/2} < dx^{1-1/A+\delta} \leq y \leq \mathrm{e}^x. \tag{2.3}$$

From Chapters 16, 19 and 20 in [Da] we can deduce the following explicit formula for prime numbers in an arithmetic progression. For integers a, d with $(a,d) = 1$, $d \geq 1$ and numbers $y \geq 2$, $T \geq 2$, one has

$$\sum_{\substack{p \leq y \\ p \equiv a \bmod d}} \log p =$$

$$\frac{y}{\varphi(d)} - \frac{1}{\varphi(d)} \sum_{\chi \bmod d} \bar{\chi}(a) \sum_{\substack{\mathrm{L}(\beta+i\gamma,\chi)=0 \\ \beta \geq 1/2,\ |\gamma| \leq T}} \frac{y^{\beta+i\gamma}}{\beta+i\gamma} + O\left(y^{1/2} \log^2(Td) + \frac{y \log^2(Tdy)}{T} \right).$$

The double summation may be bounded by noting that each $|\bar{\chi}(a)| = 1$, $|y^{\beta+i\gamma}| = y^\beta$ and $|\beta + i\gamma| \geq \sqrt{1/4 + \gamma^2} \geq (1 + |\gamma|)/3$. We let $T = x^3$ so that, using the hypothesis, $y \geq dx^{1-1/A+\delta} \geq d^2 x^{1-2/A+2\delta} \geq d^2 x^{2\delta} \log^4 x$, and thus $y^{1/2} \log^2(Td) = O(y/dx^\delta)$. Also $T > x^{3/A} > x^\delta d^3$ whereas $\log(Tdy) = O(\log y) = O(d)$ by (2.3), and thus $y \log^2(Tdy)/T = O(y/dx^\delta)$. Therefore

$$\left| \sum_{\substack{p \leq y \\ p \equiv a \bmod d}} \log p - \frac{y}{\varphi(d)} \right| \leq \frac{3}{\varphi(d)} \sum_{\chi \bmod d} \sum_{\substack{\mathrm{L}(\beta+i\gamma,\chi)=0 \\ \beta \geq 1/2,\ |\gamma| \leq x^3}} \frac{y^\beta}{1+|\gamma|} + O\left(\frac{y}{dx^\delta}\right). \tag{2.4}$$

Write $\sum_\sigma^\alpha$ for a sum over all zeros $\beta + i\gamma$ of $\mathrm{L}(s,\chi)$ and over all characters $\chi \bmod d$, where $\sigma \leq \beta < \alpha$ and $|\gamma| \leq x^3$. (Each $\beta + i\gamma$ is counted with multiplicity equal to the number of these L-functions for which it is a zero.) To estimate the double sum on the right side of (2.4), we use the upper bounds $y^\beta \le y^{1-1/A}$ for $\beta \le 1-1/A$; and $y^\beta \le y$ for $\tau \le \beta \le 1$, where $\tau = 1-\rho/\log x$ and $\rho = (1/\delta)\log(1/\varepsilon\delta)$. In the range $1-1/A \le \beta \le \tau$ we use the identity $y^\beta = y^{1-1/A} + \log y \int_{1-1/A}^{\beta} y^\sigma \mathrm{d}\sigma$. Thus the double sum on the right side of (2.4) is at most

$$\sum\nolimits_{1/2}^{\tau} \frac{y^{1-1/A}}{1+|\gamma|} \;+\; \log y \sum\nolimits_{1-1/A}^{\tau} \frac{1}{1+|\gamma|} \int_{1-1/A}^{\beta} y^\sigma \mathrm{d}\sigma \;+\; \sum\nolimits_{\tau}^{1} \frac{y}{1+|\gamma|}$$

$$\le y^{1-1/A} \sum\nolimits_{1/2}^{1} \frac{1}{1+|\gamma|} \;+\; \log y \int_{1-1/A}^{\tau} y^\sigma \Big(\sum\nolimits_{\sigma}^{1} \frac{1}{1+|\gamma|}\Big) \mathrm{d}\sigma \;+\; y \sum\nolimits_{\tau}^{1} \frac{1}{1+|\gamma|}. \tag{2.5}$$

For any $\sigma \ge 1/2$ we can use partial summation to get

$$\sum\nolimits_{\sigma}^{1} \frac{1}{1+|\gamma|} \le \mathrm{N}(\sigma,1,d) + \frac{\mathrm{N}(\sigma,x^3,d)}{x^3} + \int_1^{x^3} \frac{\mathrm{N}(\sigma,t,d)}{t^2} \mathrm{d}t. \tag{2.6}$$

For any t in the range $1 \le t \le x^3$, (2.2) implies that $N(1/2,t,d)/t \le 4c_1 d \log x$. By inserting this estimate into (2.6), we deduce an upper bound for the first sum in (2.5):

$$\sum\nolimits_{1/2}^{1} \frac{y^{1-1/A}}{1+|\gamma|} \le 4c_1 d y^{1-1/A} \log x \left(2 + \int_1^{x^3} \frac{\mathrm{d}t}{t}\right) \le 16 c_1 y^{1-\delta} \log^2 x \le \frac{y}{\log x} \tag{2.7}$$

for $x \ge x_{\varepsilon,\delta}$, since $d \le y^{1/A-\delta}$ and $y^\delta > \log^4 x$ by (2.3).

If $\sigma \ge 1-1/A$ then $A(1-\sigma) \le 1$, so that for any t in the range $1 \le t \le x^3$, (2.1) implies that $N(\sigma,t,d)/t \le \gamma_2(A) d^{A(1-\sigma)}$. By inserting this estimate into (2.6), we deduce that

$$\sum\nolimits_{\sigma}^{1} \frac{1}{1+|\gamma|} \le \gamma_2(A) d^{A(1-\sigma)} \left(2 + \int_1^{x^3} \frac{\mathrm{d}t}{t}\right) \le 4\gamma_2(A) d^{A(1-\sigma)} \log x.$$

If $\sigma \ge 1-1/2A$ then $A(1-\sigma) \le 1/2$, so that for any t in the range $1 \le t \le x^3$, (2.1) implies that $N(\sigma,t,d) \le \gamma_2(A) d^{A(1-\sigma)} t^{1/2}$. By inserting this estimate into (2.6), we deduce that

$$\sum\nolimits_{\sigma}^{1} \frac{1}{1+|\gamma|} \le \gamma_2(A) d^{A(1-\sigma)} \left(2 + \int_1^{x^3} \frac{\mathrm{d}t}{t^{3/2}}\right) \le 4\gamma_2(A) d^{A(1-\sigma)}.$$

Using the two bounds immediately above, we deduce that the middle term in (2.5) is

$$\begin{aligned} &\le 4\gamma_2(A) d^A \log y \left\{ \log x \int_{1-1/A}^{1-1/2A} \Big(\frac{y}{d^A}\Big)^\sigma \mathrm{d}\sigma \;+\; \int_{1-1/2A}^{\tau} \Big(\frac{y}{d^A}\Big)^\sigma \mathrm{d}\sigma \right\} \\ &\le 4\gamma_2(A) d^A \frac{\log y}{\log(y/d^A)} \frac{y}{d^A} \left\{ \Big(\frac{y}{d^A}\Big)^{-1/2A} \log x + \Big(\frac{y}{d^A}\Big)^{-(1-\tau)} \right\} \\ &\le \frac{4\gamma_2(A)}{\delta A} y \left\{ y^{-\delta/2} \log x + \mathrm{e}^{-\delta A \rho/2} \right\} \le \frac{\varepsilon}{9} y, \end{aligned} \tag{2.8}$$

for $x \geq x_{\varepsilon,\delta}$, since $y/d^A \geq y^{\delta A} \geq x^{\delta A/2}$ and $y^\delta > \log^4 x$ by (2.3).

Define $\mathcal{D}_{\varepsilon,\delta}(x)$ to be the set of integers d' in the range $1 \leq d' \leq x^{1/A-\delta}$ for which there is a primitive character $\chi \bmod d'$ with a zero $\beta + i\gamma$ of $\mathrm{L}(s,\chi)$ satisfying $\beta \geq \tau$ and $|\gamma| \leq \nu := \mathrm{e}^{4A\rho}/\varepsilon^2$. Since d is not divisible by any element of $\mathcal{D}_{\varepsilon,\delta}(x)$ (by hypothesis), the final sum of (2.5) only involves zeros from the region $\beta \geq \tau,\ \nu < |\gamma| \leq x^3$. Thus, using (2.1), the third sum in (2.5) is

$$y \sum\nolimits_{\tau}^{1} \frac{1}{1+|\gamma|} \leq y \frac{N(\tau, x^3, d)}{\nu} \leq \gamma_2(A) y \ \frac{x^{4A(1-\tau)}}{\nu} = \gamma_2(A) y \ \frac{\mathrm{e}^{4A\rho}}{\mathrm{e}^{4A\rho}/\varepsilon^2} \leq \frac{\varepsilon}{9} y$$

for $x \geq x_{\varepsilon,\delta}$, since $d \leq x$ by (2.3). We use this, together with (2.7) and (2.8), to bound (2.5); which we then insert into (2.4) to obtain the estimate of Theorem 2.1.

Theorem 14 of [Bo] states that there exist computable constants $c_2, c_3 > 0$ for which

$$\sum_{d \leq T} \ \sum_{\substack{\chi \bmod d \\ \chi \text{primitive}}} \mathrm{N}(\sigma, T, \chi) \ \leq \ c_2 T^{c_3(1-\sigma)},$$

for all $T \geq 2$, $\sigma \geq 1/2$. The set $\mathcal{D}_{\varepsilon,\delta}(x)$ has cardinality no bigger than the left side of this equation with $T = x^{1/A-\delta}$ and $\sigma = \tau$, which is $\leq c_2 \mathrm{e}^{\rho c_3/A} = D_{\varepsilon,\delta}$.

The lemma of Landau and Page (see page 39 of [Bo] or pages 95 and 96 of [Da]) asserts that there is a computable constant $c_4 > 0$ such that for all $T \geq 2$, there is at most one primitive character χ_1 with modulus $d_1 \leq T$ for which $\mathrm{L}(s,\chi_1)$ has a zero $\beta_1 + i\gamma_1$ satisfying $\beta_1 \geq 1 - c_4/\log T$ and $|\gamma_1| \leq T$. Moreover, if such a zero exists, it satisfies $\gamma_1 = 0$ and $\beta_1 \leq 1 - c_5/(d_1^{1/2} \log^2 d_1)$, where $c_5 > 0$ is some computable constant. We apply this result with $T = x^\eta$ where $\eta = \eta_{\varepsilon,\delta} := c_4/\rho$, so that $1 - c_4/\log(x^\eta) = \tau$. Thus $\mathcal{D}_{\varepsilon,\delta}(x)$ contains at most one number that is $\leq x^\eta$. If this number exists, call it d_1, so that $\tau \leq 1 - c_5/(d_1^{1/2} \log^2 d_1)$, and thus $d_1 \geq \log x$ since $x \geq x_{\varepsilon,\delta}$, which completes our proof of Theorem 2.1.

Remark. It is possible, in principle, to compute a value for $\gamma_2(A)$ from the work of [Hu] and [Ju], for any $A > 12/5$. One may then compute the value of all of the constants in the above proof, starting by ensuring that ε and δ are "sufficiently small", and eventually obtaining values for $x_{\varepsilon,\delta}, D_{\varepsilon,\delta}$ and $\eta_{\varepsilon,\delta}$.

3. Prachar's theorem revisited.

Since the probability that a randomly chosen positive integer below x is prime is about $1/\log x$, one might expect that for all integers $L \geq 1$ and numbers $x \geq 2$,

$$\#\{d|L : d \leq x,\ d+1 \text{ is prime}\} \geq \frac{c}{\log x} \#\{d|L : 1 \leq d \leq x\},$$

for some absolute constant $c>0$. This cannot be precisely true in general: for example, when L is odd. Nevertheless, we can actually prove a statement similar to this.

Theorem 3.1. *Suppose that B is in the set $\mathcal{B}$ defined in the Introduction. There exists a number $x_3(B)$ such that if $x \geq x_3(B)$ and L is a squarefree integer not divisible by any prime exceeding $x^{(1-B)/2}$ and for which $\sum_{\text{prime } q|L} 1/q \leq (1-B)/32$, then there is a positive integer $k \leq x^{1-B}$ with $(k,L)=1$, such that*

$$\#\{d|L : dk+1 \leq x,\ dk+1 \text{ is prime}\} \geq \frac{2^{-D_B-2}}{\log x}\#\{d|L : 1 \leq d \leq x^B\}.$$

Proof. Let $x_3(B) = \max\{x_2(B), 17^{(1-B)^{-1}}\}$. For each $d \in \mathcal{D}_B(x)$ which divides L, we divide some prime factor of d out from L, so as to obtain a number L' which is not divisible by any number in $\mathcal{D}_B(x)$. Thus $\omega(L') \geq \omega(L) - D_B$, where $\omega(m)$ is the number of distinct prime factors of m, and

$$\#\{d|L' : 1 \leq d \leq y\} \geq 2^{-D_B}\#\{d|L : 1 \leq d \leq y\} \tag{3.1}$$

for any $y \geq 1$. To see this, think of a divisor d' of L' as corresponding to a divisor d of L if and only if d' divides d and d/d' divides L/L'. So if $d \leq y$ then the corresponding d' is $\leq y$. Moreover, for any divisor d' of L', the number of divisors d of L which correspond to d' is at most the number of divisors of L/L', which is $\leq 2^{D_B}$.

From (0.3) we see that, for each divisor d of L' with $1 \leq d \leq x^B$, we have

$$\pi(dx^{1-B}; d, 1) \geq \frac{\pi(dx^{1-B})}{2\varphi(d)} \geq \frac{dx^{1-B}}{2\varphi(d)\log(dx^{1-B})} \geq \frac{dx^{1-B}}{2\varphi(d)\log x}, \tag{3.2}$$

since $\pi(y) \geq y/\log y$ for all $y \geq 17$ (see [RS]). Furthermore, since any prime factor q of L is at most $x^{(1-B)/2}$ (by hypothesis), we can use Montgomery and Vaughan's explicit version of the Brun-Titchmarsh theorem [MV], to get

$$\pi(dx^{1-B}; dq, 1) \leq \frac{2dx^{1-B}}{\varphi(dq)\log(x^{1-B}/q)} \leq \frac{4}{\varphi(q)(1-B)}\frac{dx^{1-B}}{\varphi(d)\log x} \leq \frac{8}{q(1-B)}\frac{dx^{1-B}}{\varphi(d)\log x}.$$

Therefore, by (3.2), the number of primes $p \leq dx^{1-B}$ with $p \equiv 1 \bmod d$ and $((p-1)/d, L) = 1$ is at least

$$\begin{aligned}\pi(dx^{1-B}; d, 1) &- \sum_{\text{prime } q|L} \pi(dx^{1-B}; dq, 1)\\ &\geq \left(\frac{1}{2} - \frac{8}{1-B}\sum_{\text{prime } q|L}\frac{1}{q}\right)\frac{dx^{1-B}}{\varphi(d)\log x} \geq \frac{x^{1-B}}{4\log x}.\end{aligned}$$

Thus we have at least

$$\frac{x^{1-B}}{4\log x}\#\{d|L' : 1 \le d \le x^B\}$$

pairs (p,d) where $p \le dx^{1-B}$ is prime, $p \equiv 1 \bmod d$, $((p-1)/d, L) = 1$, $d|L'$ and $1 \le d \le x^B$. Each such pair (p,d) corresponds to an integer $(p-1)/d \le x^{1-B}$ that is coprime to L, and so there is at least one integer $k \le x^{1-B}$ with $(k,L) = 1$ such that k has at least

$$\frac{1}{4\log x}\#\{d|L' : 1 \le d \le x^B\}$$

representations as $(p-1)/d$ with (p,d) as above. Thus for this integer k we have

$$\#\{d|L : dk+1 \le x,\ dk+1 \text{ is prime}\} \ge \frac{1}{4\log x}\#\{d|L' : 1 \le d \le x^B\}$$

and the theorem now follows from (3.1).

4. Carmichael numbers.

In this section we shall prove the following theorem.

Theorem 4.1. *For each $E \in \mathcal{E}$, $B \in \mathcal{B}$ and $\varepsilon > 0$, there is a number $x_4(E,B,\varepsilon)$, such that whenever $x \ge x_4(E,B,\varepsilon)$, we have $\mathrm{C}(x) \ge x^{EB-\varepsilon}$.*

This result appears to be slightly weaker than Theorem 1. However, as we shall see in the next section, $\mathcal{E}$ is an open set. Thus if $E \in \mathcal{E}$, there is some $E' > E$ with $E' \in \mathcal{E}$, so that letting $\varepsilon = (E'-E)B$, we may take $x_0(E,B)$ in Theorem 1 to be $x_4(E',B,\varepsilon)$. That is, Theorem 4.1 and Proposition 5.1 imply Theorem 1.

Proof of Theorem 4.1. Let $E \in \mathcal{E}$, $B \in \mathcal{B}$, $\varepsilon > 0$. Clearly we may assume $\varepsilon < EB$. Let $\theta = (1-E)^{-1}$ and let $y \ge 2$ be a parameter. Denote by $\mathcal{Q}$ the set of primes $q \in (y^\theta/\log y, y^\theta]$ for which $q-1$ is free of prime factors exceeding y. By (0.1),

$$|\mathcal{Q}| \ge \frac{1}{2}\gamma_1(E)\frac{y^\theta}{\log(y^\theta)} \tag{4.1}$$

for all sufficiently large y. Let L be the product of the primes $q \in \mathcal{Q}$; then

$$\log L \le |\mathcal{Q}|\log(y^\theta) \le \pi(y^\theta)\log(y^\theta) \le 2y^\theta, \tag{4.2}$$

for all large y. Now $\lambda(L)$ is the least common multiple of the numbers $q-1$ for those primes q that divide L. Since each such $q-1$ is free of prime factors exceeding y, we know that if the prime power p^a divides $\lambda(L)$ then $p \le y$ and $p^a \le y^\theta$. Thus if we let p^{a_p} be the largest power of p with $p^{a_p} \le y^\theta$, then

$$\lambda(L) \le \prod_{p\le y} p^{a_p} \le \prod_{p \le y} y^\theta = y^{\theta\pi(y)} \le \mathrm{e}^{2\theta y} \tag{4.3}$$

for all large y.

Let G be the group $(\mathbf{Z}/L\mathbf{Z})^*$ and recall the number $n(G)$ defined in Section 1. We conclude from Theorem 1.1, (4.2) and (4.3) that

$$n(G) < \lambda(L)\left(1+\log\frac{\varphi(L)}{\lambda(L)}\right) \le \lambda(L)(1+\log L) \le \mathrm{e}^{3\theta y} \tag{4.4}$$

for all large y.

Let $\delta = \varepsilon\theta/(4B)$ and let $x = \mathrm{e}^{y^{1+\delta}}$. Since

$$\sum_{\text{prime } q|L} \frac{1}{q} \le \sum_{y^\theta/\log y<q<y^\theta} \frac{1}{q} \le 2\frac{\log\log y}{\theta \log y} \le \frac{1-B}{32}$$

for sufficiently large y, we may apply Theorem 3.1 with B, x, L. Thus for all sufficiently large values of y, there is an integer k coprime to L, for which the set $\mathcal{P}$ of primes $p \le x$ with $p = dk+1$ for some divisor d of L, satisfies

$$|\mathcal{P}| \ge \frac{2^{-D_B-2}}{\log x}\#\{d|L : 1 \le d \le x^B\}. \tag{4.5}$$

The product of any

$$u := \left[\frac{\log(x^B)}{\log(y^\theta)}\right] = \left[\frac{B\log x}{\theta \log y}\right]$$

distinct prime factors of L, is a divisor d of L with $d \le x^B$. We deduce from (4.1) that

$$\#\{d|L : 1 \le d \le x^B\} \ge \binom{\omega(L)}{u} \ge \left(\frac{\omega(L)}{u}\right)^u \ge \left(\frac{\gamma_1(E)y^\theta}{2B\log x}\right)^u = \left(\frac{\gamma_1(E)}{2B}y^{\theta-1-\delta}\right)^u.$$

Thus, by (4.5) and the identity $(\theta-1-\delta)B/\theta = EB - \varepsilon/4$, we have

$$|\mathcal{P}| \ge \frac{2^{-D_B-2}}{\log x}\left(\frac{\gamma_1(E)}{2B}y^{\theta-1-\delta}\right)^{\left[\frac{B\log x}{\theta\log y}\right]} \ge x^{EB-\varepsilon/3} \tag{4.6}$$

for all sufficiently large values of y. Now take $\mathcal{P}' = \mathcal{P}\setminus\mathcal{Q}$. Since $|\mathcal{Q}| \le y^\theta$, we have by (4.6) that

$$|\mathcal{P}'| \ge x^{EB-\varepsilon/2} \tag{4.7}$$

for all sufficiently large values of y.

We may view $\mathcal{P}'$ as a subset of the group $G = (\mathbf{Z}/L\mathbf{Z})^*$ by considering the residue class of each $p \in \mathcal{P}'$ modulo L. If $\mathcal{S}$ is a subset of $\mathcal{P}'$ that contains more than one element and if

$$\Pi(\mathcal{S}) := \prod_{p \in \mathcal{S}} p \equiv 1 \bmod L,$$

then $\Pi(\mathcal{S})$ is a Carmichael number. Indeed, every member of $\mathcal{P}'$ is 1 mod k so that $\Pi(\mathcal{S}) \equiv 1 \bmod k$, and thus $\Pi(\mathcal{S}) \equiv 1 \bmod kL$, since $(k, L) = 1$. However if $p \in \mathcal{P}'$ then $p \in \mathcal{P}$ so that $p-1$ divides kL. Thus $\Pi(\mathcal{S})$ satisfies Korselt's criterion.

Let $t = \mathrm{e}^{y^{1+\delta/2}}$. Then, by Proposition 1.2, we see that the number of Carmichael numbers of the form $\Pi(\mathcal{S})$, where $\mathcal{S} \subset \mathcal{P}'$ and $|\mathcal{S}| \le t$, is at least

$$\binom{|\mathcal{P}'|}{[t]} \Big/ \binom{|\mathcal{P}'|}{n(G)} \ \ge \ \left(\frac{|\mathcal{P}'|}{[t]}\right)^{[t]} \Big/ |\mathcal{P}'|^{n(G)} \ge \left(x^{EB-\varepsilon/2}\right)^{[t]-n(G)} [t]^{-[t]} \ge x^{t(EB-\varepsilon)}$$

for all sufficiently large values of y, using (4.4) and (4.7). But each such Carmichael number $\Pi(\mathcal{S})$ so formed is such that $\Pi(\mathcal{S}) \le x^t$. Thus for $X = x^t$ we have $\mathrm{C}(X) \ge X^{EB-\varepsilon}$ for all sufficiently large y. But $X = \exp(y^{1+\delta} \exp(y^{1+\delta/2}))$, so that $\mathrm{C}(X) \ge X^{EB-\varepsilon}$ for all sufficiently large values of X. Since y can be uniquely determined from X, this completes the proof of Theorem 4.1.

5. The sets $\mathcal{E}$ and $\mathcal{B}$.

In this section we prove Theorem 3 and show that $\mathcal{E}$ is an open interval. The second result is particularly easy, being an almost immediate consequence of the Brun-Titchmarsh inequality.

Proposition 5.1. *There is some number E_0 with $0 < E_0 \le 1$ such that $\mathcal{E} = (0, E_0)$.*

Proof. Since Erdős has shown that $\mathcal{E}$ contains numbers $E > 0$ and since we evidently have $(0, E] \subset \mathcal{E}$ for any $E \in \mathcal{E}$, it suffices to show that for any $E \in \mathcal{E}$ there is some $E' > E$ with $E' \in \mathcal{E}$. Let $E \in \mathcal{E}$ and let E' be any number with $E < E' < 1$. By the Brun-Titchmarsh inequality (see [MV]), we get for $x \ge x_1(E)$ that

$$\begin{aligned}\pi(x, x^{1-E'}) &\ge \pi(x, x^{1-E}) - \sum_{x^{1-E'} \le p < x^{1-E}} \pi(x; p, 1) \\ &\ge \gamma_1(E)\pi(x) - \sum_{x^{1-E'} \le p < x^{1-E}} \frac{2x}{\varphi(p)\log(x/p)}.\end{aligned}$$

Now using $\pi(x) \geq x/\log x$ for all $x \geq 17$ (see [RS]), we have for $x \geq x_1(E), x \geq 17$ that

$$\pi(x, x^{1-E'}) \geq \frac{\gamma_1(E)x}{\log x} - \sum_{x^{1-E'} \leq p < x^{1-E}} \frac{2x}{E(p-1)\log x}$$
$$= \frac{x}{\log x}\left(\gamma_1(E) - \frac{2}{E}\sum_{x^{1-E'} \leq p < x^{1-E}} \frac{1}{p-1}\right).$$

By Mertens' theorem, we have

$$\sum_{x^{1-E'} \leq p < x^{1-E}} \frac{1}{p-1} = \log\frac{1-E}{1-E'} + O\left(\frac{1}{(1-E')\log x}\right)$$

for $x > 1$. Thus if E' is taken so close to E that

$$\gamma_1(E) - \frac{2}{E}\log\frac{1-E}{1-E'} > \frac{1}{2}\gamma_1(E),$$

say, then

$$\pi(x, x^{1-E'}) > \frac{1}{3}\gamma_1(E)\frac{x}{\log x},$$

for all large x. We conclude from the prime number theorem that $E' \in \mathcal{E}$, completing the proof of Proposition 5.1.

We now give the proof of Theorem 3.

Proof of Theorem 3. Assume that $B \in \mathcal{B}$ and that $x \geq x_2(B)$. Choose a number δ in the range $0 < \delta < B$ and let $\varepsilon = \delta^2/(20B)$. For each number d in $\mathcal{D}_B(x)$, select some prime factor p_d of d. Let $\mathcal{P}$ be the set of primes in the interval $[x^{\delta/2}, x^{\delta/2+\varepsilon}]$ not equal to any p_d, $d \in \mathcal{D}_B(x)$. Since $\mathcal{P}$ contains all but at most D_B of the primes in this interval, Mertens' theorem implies that

$$\sum_{p \in \mathcal{P}} \frac{1}{p} = \log(1 + \delta/(10B)) + O(1/(\delta \log x)).$$

We deduce that

$$\sum_{p \in \mathcal{P}} \frac{1}{p} \geq \frac{\delta}{20B}, \tag{5.1}$$

for all sufficiently large x.

We shall give a lower bound for $\pi(x, x^{1-B+\delta})$ by counting pairs (q, d), where $q \leq x$ is a prime in the congruence class 1 mod d, and d is an integer in the range $x^{B-\delta} \leq d \leq x^B$, whose every prime factor lies in $\mathcal{P}$. Evidently any such prime q must be counted in

$\pi(x, x^{1-B+\delta})$, but will not be involved in more than $2^{2/\delta}$ such pairs (q, d) (since $q-1$ cannot have more than $2/\delta$ prime factors from $\mathcal{P}$). Thus from (0.3) we have

$$(5.2)\qquad \pi(x, x^{1-B+\delta}) \geq 2^{-2/\delta} \sum_{\substack{x^{B-\delta} \leq d \leq x^B \\ p|d \Rightarrow p \in \mathcal{P}}} \pi(x; d, 1) \geq 2^{-1-2/\delta} \sum_{\substack{x^{B-\delta} \leq d \leq x^B \\ p|d \Rightarrow p \in \mathcal{P}}} \frac{\pi(x)}{\varphi(d)}$$

for all $x \geq x_2(B)$. Let u denote the least integer with $u \geq (B-\delta)/(\delta/2)$ so that

$$B - \delta \leq u\delta/2 \quad \text{and} \quad u(\delta/2 + \varepsilon) < (2B/\delta - 1)(\delta/2 + \varepsilon) = B + \frac{\delta}{10} - \frac{\delta}{2} - \varepsilon < B.$$

Therefore any product, d, of u not necessarily distinct primes from $\mathcal{P}$ satisfies

$$x^{B-\delta} \leq x^{u\delta/2} \leq d \leq x^{u(\delta/2+\varepsilon)} \leq x^B,$$

and so, by (5.1),

$$\sum_{\substack{x^{B-\delta} \leq d \leq x^B \\ p|d \Rightarrow p \in \mathcal{P}}} \frac{1}{d} \geq \frac{1}{u!}\left(\sum_{p\in\mathcal{P}} \frac{1}{p}\right)^u \geq \frac{1}{u!}\left(\frac{\delta}{20B}\right)^u = \gamma_3(B, \delta),$$

say. Since $1/\varphi(d) > 1/d$ we can insert this estimate into (5.2) to deduce that (0.1) holds for $E = B - \delta$ with some number $\gamma_1(E)$ satisfying $\gamma_1(E) \geq 2^{-1-2/\delta}\gamma_3(B, \delta)$. This completes the proof of Theorem 3.

References

[APR] L. M. Adleman, C. Pomerance and R. S. Rumely, "On distinguishing prime numbers from composite numbers", *Ann. of Math.* **117** (1983), 173–206.

[Ba] P. C. Baayen, "Een combinatorisch probleem voor eindige abelse groepen", in "Colloquium Discrete Wiskunde" *MC Syllabus* **5** (1968), Mathematisch Centrum, Amsterdam, pp. 76–108.

[BS] R. C. Baker and W. M. Schmidt, "Diophantine problems in variables restricted to the values 0 and 1", *J. Number Theory* **12** (1980), 460–486.

[Bo] E. Bombieri, "Le grand crible dans la théorie analytique des nombres", *Astérisque* **18** (1987/1974), 103 pp.

[Ca1] R. D. Carmichael, "Note on a new number theory function", *Bull. Amer. Math. Soc.* **16** (1910), 232–238.

[Ca2] R. D. Carmichael, "On composite numbers P which satisfy the Fermat congruence $a^{P-1} \equiv 1 \pmod P$", *Amer. Math. Monthly* **19** (1912), 22–27.

[Da] H. Davenport, "Multiplicative Number Theory" (2nd edn.), Springer-Verlag, New York, 1980.

[Du] H. J. A. Duparc, "On Carmichael numbers", *Simon Stevin* **29** (1952), 21–24.

[EK] P. van Emde Boas and D. Kruyswijk, "A combinatorial problem on finite abelian groups III", Z.W. 1969-008 (Math. Centrum, Amsterdam).

[Er1] P. Erdős, "On the normal number of prime factors of $p-1$ and some other related problems concerning Euler's φ–function", *Quart. J. Math. (Oxford Ser.)* **6** (1935), 205–213.

[Er2] P. Erdős, "On pseudoprimes and Carmichael numbers", *Publ. Math. Debrecen* **4** (1956), 201–206.

[EPS] P. Erdős, C. Pomerance and E. Schmutz, "Carmichael's lambda function", *Acta Arith.* **58** (1991), 363–385.

[Fr] J. B. Friedlander, "Shifted primes without large prime factors", in *Number Theory and Applications* (ed. R. A. Mollin), (Kluwer, NATO ASI, 1989), 393–401.

[GP] D. M. Gordon and C. Pomerance, "The distribution of Lucas and elliptic pseudoprimes", *Math. Comp.* **57** (1991), 825–838.

[Hu] M. N. Huxley, "Large values of Dirichlet polynomials, III", *Acta Arith.* **26** (1974), 435–444.

[Ju] M. Jutila, "On Linnik's constant", *Math. Scand.* **41** (1977), 45–62.

[Ko] A. Korselt, "Problème chinois", *L'intermédiaire des mathématiciens* **6** (1899), 142–143.

[LMO] R. Lidl, W. B. Müller and A. Oswald, "Some remarks on strong Fibonacci pseudoprimes", *Applicable Alg. in Eng., Comm. and Computing* **1** (1990), 59–65.

[Me] R. Meshulam, "An uncertainty inequality and zero subsums", *Disc. Math.* **84** (1990), 197–200.

[MV] H. L. Montgomery and R. C. Vaughan, "The large sieve", *Mathematika* **20** (1973), 119–134.

[Ol] J. Olson, "A combinatorial problem on finite Abelian groups, I", *J. Number Theory* **1** (1969), 8–10.

[Pi] R. G. E. Pinch, "The Carmichael numbers up to 10^{15}", *Math. Comp.* **61** (1993), 381–391.

[Po] C. Pomerance, "Two methods in elementary analytic number theory", in *Number Theory and Applications* (ed. R. A. Mollin), (Kluwer, NATO ASI, 1989), 135–161.

[PSW] C. Pomerance, J. L. Selfridge and S. S. Wagstaff Jr., "The pseudoprimes to $25 \cdot 10^9$", *Math. Comp.* **35** (1980), 1003–1026.

[Pr] K. Prachar, " Über die Anzahl der Teiler einer natürlichen Zahl, welche die Form $p-1$ haben", *Monatsh. Math.* **59** (1955), 91–97.

[RS] J. B. Rosser and L. Schoenfeld, "Approximate formulas for some functions of prime numbers", *Illinois J. Math.* **6** (1962), 64–94.

[Wi] H. C. Williams, "On numbers analogous to the Carmichael numbers", *Canad. Math. Bull.* **20** (1977), 133–143.

[Zh] M. Zhang, "Searching for large Carmichael numbers", Sichuan Daxue Xuebao, *to appear*.