Carmichael Numbers: Proof of Infinitude and Algorithmic Generation

Dale Liu

July 13, 2024

Abstract

In this paper, we begin by establishing a few fundamental properties of Carmichael numbers. We then show the details and motivation behind Alford, Granville, and Pomerance's proof that there are infinitely many Carmichael numbers. Additionally, we present a few bounds on these numbers, including an analogue of Bernard's postulate for Carmichael numbers. Finally, we analyze a few algorithms designed to generate and check large Carmichael numbers.

Acknowledgements

The author would like to thank Sawyer Dobson for his guidance and mentorship throughout the process of writing this paper. His advice has been invaluable towards the creation of this paper. The author would also like to thank Simon Rubinstein-Salzedo for his lectures on paper writing and math talks. He created an environment that inspired the author to explore more into mathematical paper writing.

1 Introduction

Fermat's little theorem states that for any prime p and integer a, coprime to p, we must have

$$a^{p-1} \equiv 1 \pmod{p}. \tag{1}$$

Naturally, the question arose as to whether p could be a composite integer and still satisfy the equation. These composite integers are known as the Carmichael numbers.

Definition 1.1. A composite number n is a Carmichael number if for any integer a coprime to n, we have $a^{n-1} \equiv 1 \pmod{n}$.

These numbers were named after Robert Carmichael, an American mathematician who began an in-depth study of these numbers in 1910. Carmichael noted that the smallest such number was 561. These numbers are important because they are composite numbers able to pass primality tests such as the Fermat primality test for all possible bases. This makes them useful in public key encryption systems such as the RSA Algorithm.

For most of the 20th century, a big unsolved question regarding the Carmichael numbers was whether there were an infinite amount of them. It wasn't until 1994, when Alford, Granville, and Pomerance proved that this was true. In this paper, we will explain their proof, as well as their motivation behind it. We will also explore a few interesting properties of the Carmichael numbers, and a few algorithms used to identify these numbers.

2 Properties of Carmichael Numbers

First we will show a few proofs of some properties of Carmichael numbers. These proofs will only require elementary techniques.

Theorem 2.1 (Korselt's Criterion)**.** *A composite integer $n > 2$ is a Carmichael number if and only if n is squarefree and for all primes p dividing n, $(p-1) \mid (n-1)$. [Kor99]*

Proof. We will first show that if n is a Carmichael number, then n must be squarefree. Denote $v_p(x)$ as the largest integer e such that $p^e \mid x$. For the sake of contradiction, assume $v_p(n) \geq 2$ for some prime p, and let $k = v_p(n)$. We can now write $n = p^k n'$. We will use the Chinese Remainder Theorem to get a contradiction.
Since $(p^k, n') = 1$, by the Chinese Remainder Theorem, there must be an integer $a \leq n$ such that

$$a \equiv p+1 \pmod{p^k} \tag{2}$$

$$a \equiv 1 \pmod{n'} \tag{3}$$

It follows that $(a, n) = 1$, so by the definition of Carmichael numbers, $a^{n-1} \equiv 1 \pmod{n}$. Since $k \geq 2 \implies p^2 \mid n$, we can take this modulo p^2 instead and get $(p+1)^{n-1} \equiv 1 \pmod{p^2}$. By the binomial theorem,

$$1 \equiv (p+1)^{n-1} \tag{4}$$

$$\equiv p^{n-1} + \binom{n-1}{1} p^{n-2} + \cdots + \binom{n-1}{n-2} p^1 + \binom{n-1}{n-1} p^0 \tag{5}$$

$$\equiv (n-1)p + 1 \pmod{p^2} \tag{6}$$

Since $n - 1 \equiv -1 \pmod{p^2}$, we have $1 - p \equiv 1 \pmod{p^2}$, which is a contradiction, thus $k = 1$.

Next, we will show if n is a Carmichael number, then $(p-1) \mid (n-1)$ for every prime $p \mid n$. Since n is squarefree, $(p, n/p) = 1$. Now, by the Chinese Remainder Theorem, we can pick an integer a such that $a \equiv 1 \pmod{n/p}$ with a being a primitive root modulo p, so $(a, n) = 1$. Then, $a^{n-1} \equiv 1 \pmod{n}$. We can reduce both sides modulo p to get $a^{n-1} \equiv 1 \pmod{p}$, and since a is a primitive root modulo p, the order of $a \pmod{p}$ is $p - 1$, which implies that $(p-1) \mid (n-1)$.

Now we will show that if n is a squarefree integer, and $(p-1) \mid (n-1)$ for every prime p that divides n, then n is a Carmichael number. For any integer a, where $(a, n) = 1$, we must have $(a, p) = 1$ for all primes p dividing n. By Fermat's Little Theorem, $a^{p-1} \equiv 1 \pmod{p}$, and since $(p-1) \mid (n-1)$, $a^{n-1} \equiv 1 \pmod{p}$ for all primes p that divide n, and with the condition that n is squarefree, we have $a^{n-1} \equiv 1 \pmod{n}$. By definition, this means n must be a Carmichael number. □

Alwin Korselt proved this theorem in 1899, 11 years before Robert Carmichael began an in-depth study of these numbers. Unlike Carmichael, Korselt was unable to find any numbers that satisfied this criterion, which is why they are called the Carmichael numbers.

Corollary 2.1 (Chernick)**.** *If k is a positive integer and $6k+1, 12k+1$, and $18k+1$ are all prime, then $(6k+1)(12k+1)(18k+1)$ is a Carmichael number. [Che39]*

Proof. We will use Korselt's Criterion to verify Chernick's construction of Carmichael numbers. Let $n = (6k+1)(12k+1)(18k+1)$. n must be squarefree because it is the product of three distinct primes. Now, we need to show $n-1$ is divisible by $6k, 12k$, and $18k$, so that n satisfies Korselt's Criterion. We can do this by expanding the product

$$n - 1 = (6k+1)(12k+1)(18k+1) - 1 = 36k(36k^2 + 11k + 1)$$

And since $36k \mid n-1$, it follows that $6k, 12k$, and $18k$ must also divide $n-1$, so n must be a Carmichael number. □

The smallest number that can be constructed this way is $1729 = 7 \cdot 13 \cdot 19$ which is indeed a Carmichael number. By the Hardy-Littlewood k-tuples conjecture, there should be infinitely many integers k such that $6k+1, 12k+1$, and $18k+1$ are all prime, which implies that there are infinitely many Carmichael numbers.

Proposition 2.1. *All Carmichael numbers are odd.*

Proof. We will prove this is true by contradiction. Assume $n > 2$ is an even Carmichael number. Now, let $a = n-1$. Since $(n, n-1) = 1$, by definition, $a^{n-1} \equiv 1 \pmod{n} \implies (-1)^{n-1} \equiv 1 \pmod{n}$. But since n is even, $1 \equiv (-1)^{n-1} \equiv -1 \pmod{n}$, which gives us a contradiction. □

Proposition 2.2. *Every Carmichael number n has at least three prime factors.*

Proof. Let $n = pq$ for primes p and q. Since n is squarefree, p and q are distinct. Without loss of generality, assume $p > q$. By Korselt's Criterion, we have $(p-1) \mid (n-1)$, so

$\frac{n-1}{p-1} = \frac{pq-1}{p-1} = q + \frac{q-1}{p-1}$ must be an integer. But since $p > q$, this implies $0 < \frac{q-1}{p-1} < 1$, so $(p-1) \nmid (n-1)$. Therefore, n must have more than two prime factors. □

Proposition 2.3. *Every prime factor of a Carmichael number n is less than $\sqrt{n}$.*

Proof. Let p be a prime factor of n. From Korselt's criterion, we have $p-1 \mid n-1$, so $0 \equiv n-1 \equiv \frac{np}{p} - 1 \equiv \frac{n}{p} - 1 \pmod{p-1}$. We can't have $n = p$, therefore $\frac{n}{p} - 1 \geq p-1 \implies n \geq p^2$. Since n is squarefree, $n \neq p^2 \implies n > p^2$. □

3 There are infinitely many Carmichael numbers

For most of the 20th century, it was believed that the list of Carmichael numbers may be infinitely extended, but no one could come up with a proof. In 1994, William Alford, Andrew Granville, and Carl Pomerance published a paper proving that there are infinitely many Carmichael numbers. The proof stated that there is a finite value c, such that if $x \geq c$, then the number of Carmichael numbers less than x is greater than $x^{2/7}$ [AGP94]. In short, if x is large enough, the number of Carmichael numbers less than x is at least $x^{2/7}$.

The idea was to construct a large number L along with a set of k distinct primes such that for each prime p, we have $(p-1) \mid L$. Now, take a subset of the k primes, multiply them together, and let this product be P. If $P \equiv 1 \pmod{L}$, then P is a Carmichael number from Korselt's criterion because $(p-1) \mid L \mid P-1$ for every prime p that divides P. This simplifies the problem to proving that there are infinitely many such products P as L approaches infinity.

3.1 The sets $\mathcal{E}$ and $\mathcal{B}$

The motivation behind defining the sets $\mathcal{E}$ and $\mathcal{B}$ which we will see later on is to construct the sets $\mathcal{Q}$ and $\mathcal{P}$ with certain properties, in Subsection 3.3. These sets would then be used to construct the set $\mathcal{P}' = \mathcal{P} \backslash \mathcal{Q}$, which we will use to generate Carmichael numbers using the products of various subsequences. This last part will involve some Group Theory, which is why the next subsection is dedicated towards that topic.

Definition 3.1. Denote the function $C(x)$ as the number of Carmichael numbers less than x.

Definition 3.2. $\pi(x)$ is the number of primes p less than x.

Definition 3.3. $\pi(x, y)$ is the number of primes p less than x for which $p-1$ has no prime factors exceeding y.

Definition 3.4. $\pi(x, d, a)$ is the number of primes p less than x such that $p \equiv a \pmod{d}$.

De la Vallée Poussin proved that as x approaches infinity,

$$\pi(x, d, a) \sim \pi(x)/\varphi(d), \tag{7}$$

given that $(a, d) = 1$ (In this case, φ represents Euler's totient function).

Definition 3.5. Denote $\mathcal{E}$ as the set of numbers $E \in (0, 1)$ for which there exists positive numbers $x_{_E}$ and $\gamma_{_E}$ such that $\pi(x, x^{1-E}) \geq \gamma_{_E}\pi(x)$ for all $x \geq x_{_E}$.

Proposition 3.1. *If $E \in \mathcal{E}$ then $(0, E] \subset \mathcal{E}$.*

Proof. Let $E' \in (0, E]$, since $E \geq E'$, $\pi(x, x^{1-E'}) \geq \pi(x, x^{1-E})$. Now, if we let $\gamma_{_{E'}} = \gamma_{_E}$, the inequality $\pi(x, x^{1-E'}) \geq \gamma_{_{E'}}\pi(x)$ holds for all values $x \geq x_{_E}$, so if we set $x_{_{E'}} = x_{_E}$, we are done. □

Proposition 3.2. *The interval $(0, 1 - (2\sqrt{e})^{-1}) \subset \mathcal{E}$. [Fri89]*

Definition 3.6. Denote $\mathcal{B}$ as the set of numbers $B \in (0, 1)$ for which there exists a positive number $x_{_B}$ and a positive integer $D_{_B}$ such that if $x \geq x_{_B}$, $(a, d) = 1$ and $1 \leq d \leq min(x^B, y/x^{1-B})$, then $\pi(y, d, a) \geq \frac{\pi(y)}{2\varphi(d)}$ whenever d is not divisible by any member of a set with $D_{_B}$ integers, each of which exceeds $log(x)$.

Alford, Granville, and Pomerance [AGP94] proved that the interval $(0, 5/12) \subset \mathcal{B}$ using a bound on the zeros of Dirichlet L-functions.

Theorem 3.1 ([AGP94])**.** *Let $B \in \mathcal{B}$. There exists a number β_x such that if $x \geq \beta_x$ and L is a squarefree integer not divisible by any prime exceeding $x^{(1-B)/2}$ and for which $\sum_{prime\ p|L} \leq (1 - B)/32$, then there exists a positive integer $k \leq x^{1-B}$, relatively prime to L, such that*

$$\#\{d \mid L : dk + 1 \leq x, dk + 1 \text{ is prime}\} \geq \frac{2^{-(D_B+2)}}{\ln(x)}\#\{d \mid L : 1 \leq d \leq x^B\} \tag{8}$$

3.2 Group Theory

A group G is defined from a set and an operation, say $\oplus$. The operation must be associative, and the set must contain an identity element, say i. Associativity states that for any 3 elements a, b, c in G, $(a \oplus b) \oplus c = a \oplus (b \oplus c)$ holds. For every element a in G, the identity element satisfies $a \oplus i = a$. Every element of the group must also have an inverse element. The inverse element of element a is defined to be an element in the group b, such that $a \oplus b = i$. An example of a group is the set of positive real numbers with the multiplication operation. The

identity element is 1 since any number multiplied with 1 is equivalent to itself. The inverse of a number n is $\frac{1}{n}$ because $n \cdot \frac{1}{n}$ is equal to the identity element.

Group Theory plays a big role in this proof because the problem of finding Carmichael numbers can be simplified into finding a subsequence of elements in a cleverly constructed group with product equal to the identity.

Proposition 3.3. *If G is a group with m elements, then any sequence of m elements of the group contains a subsequence whose product is 1 (the identity).*

Proof. Let the sequence be $g_1, g_2, \cdots, g_m$. If no subsequence has product 1, then the m products $g_1, g_1g_2, \cdots, g_1g_2 \cdots g_m$ can only contain $m-1$ distinct values. This implies that at least two subsequences have the same product. Let these two products be $g_1g_2 \cdots g_i$ and $g_1g_2 \cdots g_j$. Without loss of generality, assume $j > i$. Now, the product $g_{i+1}g_{i+2} \cdots g_j$ is equal to 1. □

Definition 3.7. An abelian group is a group that has a commutative group operation.

If G is communative, $a \oplus b = b \oplus a$ must hold, where $\oplus$ denotes the operation of the group. An example of an abelian group is the set of positive integers less than 10 that are relatively prime to 10, with multiplication modulo 10 being the operation. Since multiplication is a commutative operation, this group is abelian. We can also check that this set satisfies the definition of a group. The operation is indeed associative, and the identity element is 1. The inverse of the elements $\{1, 3, 7, 9\}$ are $\{1, 7, 3, 9\}$ respectively.

Definition 3.8. The order of an element g in a group G is the smallest positive integer m such that g^m is equivalent to the identity.

For example, let G $= \{1, 3, 7, 9\}$ be the set of positive integers modulo 10 that are relatively prime to 10. This can also be rewritten as $(\mathbb{Z}/10\mathbb{Z})^*$. The element 3 has order 4 because $3^4 \equiv 1 \pmod{10}$, and 4 is the smallest such value. Lagrange's theorem states that the order of any element divides the total number of elements in the group.

Definition 3.9. For a finite group G, denote $n(G)$ as the length of the longest sequence of (not necessarily distinct) elements in G for which no nonempty subsequence has product the identity.

Theorem 3.2. *If G is a finite abelian group and m is the maximal order of an element in G, then* $n(G) < m(1 + \ln(|G|/m))$.

Proposition 3.4. *If G is a finite abelian group and let $r > t > n(G)$ be integers. Then any sequence of r elements in G contains at least $\binom{r}{t}/\binom{r}{n(G)}$ distinct subsequences of length at most t and at least $t - n(G)$ whose product is equal to the identity.*

Proof. Let R be a sequence of r elements of G. Since $r > n(G)$, there must be some subsequence of R whose product is the identity. Let S be the longest such subsequence with length s. Then $s \geq r - n(G)$ because if not, $R \backslash S$ has at least $n(G)$ elements which means it contains a subsequence with product equal to the identity. This subsequence can then be appended into S to create a longer subsequence, which contradicts the maximality of S.

Let T be any subsequence of S with length $t - n(G)$, and call the product of the sequence g. Since the product of S is the identity, the elements of $S \backslash T$ has a product of g^{-1}. Let U be the smallest subsequence of $S \backslash T$ whose product is g^{-1}. U cannot contain any subsequence that has product equal to the identity since U is the smallest subsequence, so the length of U must be less than or equal to $n(G)$. Now, let $V = T \cup U$, and clearly, V is a subsequence of S, and also R. The product of the elements in V is the identity, and V has size at most t and at least $t - n(G)$.

The number of pairs of sequences (T, U) is at least the number of ways of choosing T, which is $\binom{s}{t-n}$. The maximum amount of pairs (T, U) that have the same sequence $V = T \cup U$ is at most $\binom{|V|}{t-n} \leq \binom{t}{t-n} = \binom{t}{n}$. Thus, the amount of distinct subsequences V is at least

$$\binom{s}{t-n}/\binom{t}{n} \geq \binom{r-n}{t-n}/\binom{t}{n} = \binom{r}{t}/\binom{r}{n}, \tag{9}$$

which proves the proposition.

□

3.3 Main theorem

Theorem 3.3 ([AGP94])**.** *For each $E \in \mathcal{E}$, $B \in \mathcal{B}$, and $\varepsilon > 0$ there exists a number x_{EB} such that $C(x) \geq x^{EB-\varepsilon}$ if $x \geq x_{EB}$.*

Since there are positive reals in both sets $\mathcal{E}$ and $\mathcal{B}$, this theorem would imply that there are infinitely many Carmichael numbers because as x approaches infinity, $x^{EB-\varepsilon}$ also approaches infinity if we set $\varepsilon < EB$.

Proof. Let $\theta = \frac{1}{1-E}$, and define a parameter $y \geq 2$. Denote $\mathcal{Q}$ as the set of primes $p \in (\frac{y^\theta}{\ln(y)}, y^\theta]$ for which $p - 1$ is free of prime factors exceeding y. It is known that $\pi(y^\theta) > \frac{y^\theta}{2\ln(y^\theta)}$ for any

sufficiently large y, so by the definition of $\mathcal{E}$,

$$|\mathcal{Q}| \geq \gamma_E \frac{y^\theta}{2\ln(y^\theta)}. \tag{10}$$

Let L be the product of all primes $p \in \mathcal{Q}$. We have,

$$\ln(L) \leq |\mathcal{Q}| \ln(y^\theta) \leq \pi(y^\theta) \ln(y^\theta) \leq 2y^\theta \tag{11}$$

since $\frac{2y^\theta}{\ln(y^\theta)} \geq \pi(y^\theta)$. Now, let $\lambda(L)$, also known as the Carmichael function, be the least common multiple of the numbers $p-1$ for each prime $p \mid L$. If a is the largest integer such that $p^a \mid \lambda(L)$, then $p^a \leq y^\theta$ since $p \leq y$. Now, if we let p^{a_p} be the largest power of p such that $p^{a_p} \leq y^\theta$, then

$$\lambda(L) \leq \prod_{p \leq y} p^{a_p} \leq \prod_{p \leq y} y^\theta = y^{\theta\pi(y)} \leq e^{2\theta y} \tag{12}$$

for all values of y that are sufficiently large. Now that we have a bound for the function $\lambda(L)$, we can relate this to the group theory theorems we stated earlier. Let G be the group $(\mathbb{Z}/L\mathbb{Z})^*$, which is the set of integers modulo L that are relatively prime to L. By Theorem 3.2,

$$n(G) < \lambda(L)\left(1 + \ln\left(\frac{\varphi(L)}{\lambda(L)}\right)\right) \leq \lambda(L)(1 + \ln(L)) \leq e^{3\theta y}. \tag{13}$$

Now, let $\alpha = \frac{\varepsilon\theta}{4B}$ and let $x = e^{y^{\alpha+1}}$. Since

$$\sum_{\text{prime } p|L} \frac{1}{p} \leq \sum_{\frac{y^\theta}{\ln(y)} < p < y^\theta} \frac{1}{p} \leq \frac{2\ln(\ln(y))}{\theta \ln(y)} \leq \frac{1-B}{32} \tag{14}$$

for all large y, we can apply Theorem 3.1 for B, x, and L. Let k be an integer coprime to L, let $\mathcal{P}$ be the set of primes $p \leq x$ with $p = dk + 1$ for some divisor d of L, and let the set $\mathcal{H} = \{d \mid L : 1 \leq d \leq x^B\}$. We know there exists a k such that

$$|\mathcal{P}| \geq \frac{2^{-(D_B+2)}}{\ln(x)} |\mathcal{H}|. \tag{15}$$

Since the product of any $r \coloneqq \frac{\ln(x^B)}{\ln(y^\theta)} = \frac{B\ln(x)}{\theta\ln(y)}$ distinct prime factors of L is a divisor $d \leq x^B$ of L, so

$$|\mathcal{H}| \geq \binom{\omega(L)}{r} \geq \left(\frac{\omega(L)}{r}\right)^r \geq \left(\frac{\gamma_E y^\theta}{2B\ln(x)}\right)^r = \left(\frac{\gamma_E y^{\theta-1-\alpha}}{2B}\right)^r \tag{16}$$

where $\omega(L)$ is defined as the number of distinct prime factors of L. Since $\frac{(\theta-1-\alpha)B}{\theta} = EB - \frac{\varepsilon}{4}$,

$$|\mathcal{P}| \geq \frac{2^{-(D_B+2)}}{\ln(x)} \left(\frac{\gamma_E y^{\theta-1-\alpha}}{2B}\right)^{\left(\frac{B\ln(x)}{\theta\ln(y)}\right)} \geq x^{EB-\varepsilon/3} \tag{17}$$

for all large values of y, and D_B is the set defined in Definition 3.6. Now let $\mathcal{P}' = \mathcal{P} \backslash \mathcal{Q}$. Since $|\mathcal{Q}| \leq y^\theta$, we have

$$|\mathcal{P}'| \geq x^{EB-\varepsilon/2} \tag{18}$$

for all sufficiently large y. Every element in $\mathcal{P}'$ is relatively prime to L, so we may view $\mathcal{P}'$ as a subset of the group G which was previously defined as $(\mathbb{Z}/L\mathbb{Z})^*$. If $\mathcal{S}$ is a subset of $\mathcal{P}'$ with more than one element, and

$$\Pi(\mathcal{S}) \coloneqq \prod_{p \in \mathcal{S}} p \equiv 1 \pmod{L}, \tag{19}$$

then $\Pi(\mathcal{S})$ is a Carmichael number. To see why this is true, first note that each element in P' is 1 mod k, so $\Pi(\mathcal{S}) \equiv 1 \pmod{k}$, and since $(k, L) = 1$, we have $\Pi(\mathcal{S}) \equiv 1 \pmod{kL}$. Now, for every $p \in \mathcal{P}'$, we have $p - 1 \mid kL$ because $\mathcal{P}' \subset \mathcal{P}$, and thus $\Pi(\mathcal{S})$ satisfies Korselt's criterion.

Now let $t = e^{y^{\alpha/2+1}}$. From Proposition 3.4, the number of subsets S where $\Pi(\mathcal{S}) \equiv 1 \pmod{L}$ and $|\mathcal{S}| \leq t$ is at least

$$\binom{|\mathcal{P}'|}{t} \Big/ \binom{|\mathcal{P}'|}{n(G)} \geq \left(\frac{|\mathcal{P}'|}{t}\right)^t / |\mathcal{P}'|^{n(G)} \geq \left(x^{EB-\varepsilon/2}\right)^{t-n(G)} t^{-t} \geq x^{t(EB-\varepsilon)} \tag{20}$$

for all large values of y. We know that the upper bound of $\Pi(\mathcal{S})$ is x^t, we can let $X = x^t$ and see that $C(X) \geq X^{EB-\varepsilon}$ for all sufficiently large values of y. Since X is defined by a function of y, so the values of X and y are bijective, thus $C(X) \geq X^{EB-\varepsilon}$ for all sufficiently large values of X. □

Since $\mathcal{E}$ is an open set, there must be some value $E' \in \mathcal{E}$ such that $E' > E$, where E is the number defined in Theorem 3.3. This way, if we set $\varepsilon = B(E' - E)$, we can see $C(X) \geq X^{EB}$ for all sufficiently large values of X. We know that $(0, 1 - (2\sqrt{e})^{-1}) \subset \mathcal{E}$ and $(0, 5/12) \subset \mathcal{B}$, so any value of $EB < \frac{5}{12}(1 - \frac{1}{2\sqrt{e}}) = 0.2903$ works, for example, $2/7 = 0.2857 < 0.2903$, so $C(X) \geq x^{2/7}$ for sufficiently large values of X.

4 Further bounds on Carmichael numbers

In 2005, Glyn Harman [Har05] improved the lower bound to $C(x) > x^{0.332}$, and 3 years later, improved this again to $C(x) > x^{1/3}$. Before a lower bound was proved, many mathematicians including Knödel and Erdős, supplied upper bounds on $C(x)$, with the best from Richard Pinch [Pin93] who proved

$$C(x) < x \cdot \exp\left(-\frac{\ln(x)\ln(\ln(\ln(x)))}{\ln(\ln(x))}\right) \tag{21}$$

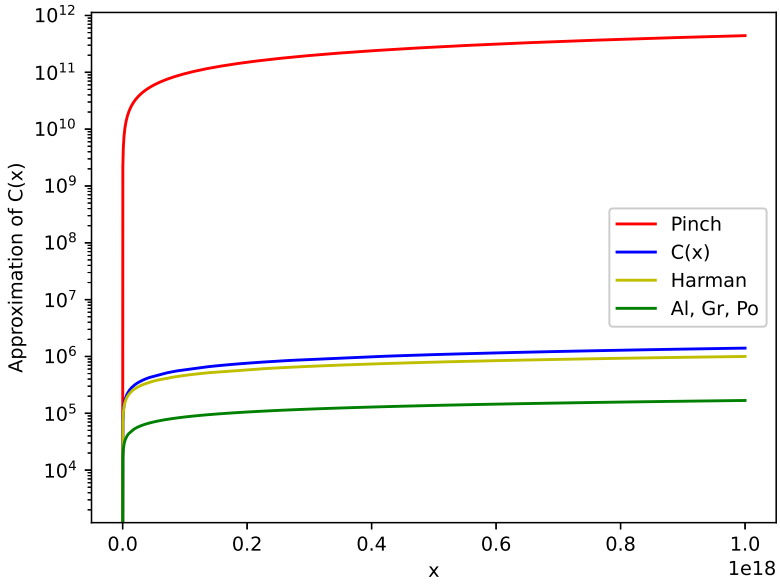

Figure 1: Upper and lower bounds compared to C(x) for $x \leq 10^{18}$

where $\exp(n) = e^n$.

In Figure 1, the bounds given by Pinch, Harman, Alford, Granville, and Pomerance are graphed and scaled logarithmically alongside the function $C(x)$ for $x \leq 10^{18}$, using data from Richard Pinch's website [Pin20]. In total, there are 1,401,644 Carmichael numbers less than 10^{18}. From the graph, we can see that Harman's lower bound of $x^{1/3}$ seems to be a very accurate approximation of $C(x)$ for all $x \leq 10^{18}$.

In Alford, Granville, and Pomerance's paper, they conjectured that a statement similar to Bertrand's postulate could be proven for Carmichael numbers. Bertrand's postulate states that for any integer $n > 1$, there is at least one prime p such that $n < p < 2n$. In 2021, a 17 year old Daniel Larson [Lar21] published a paper proving that for any $\delta > 0$, there is a finite value x_δ such that when $x \geq x_\delta$, there exists at least $\exp\left(\frac{\ln(x)}{(\ln(\ln(x)))^{2+\delta}}\right)$ Carmichael numbers between x and $x + \frac{x}{(\ln(x))^{\frac{1}{2+\delta}}}$. Since $\exp\left(\frac{\ln(x)}{(\ln(\ln(x)))^{2+\delta}}\right) > 1$ for all positive values of δ and x, there must be at least one Carmichael number between every range given in the theorem. Larson built his ideas upon the techniques developed by Yitang Zhang and James Maynard relating to small gaps between primes.

5 Finding larger Carmichael numbers

As the numbers become larger, Carmichael numbers appear less and less. Using Harman's lower bound of $C(x) > x^{1/3}$, the frequency of Carmichael numbers can be approximated with $x^{1/3}/x = x^{-2/3}$ which is a decreasing function. When $x = 10^{18}$, there is approximately one Carmichael number every trillion numbers. This makes it hard to find large Carmichael numbers. Using Korselt's criterion, checking whether a number n is Carmichael is equivalent to finding its prime factorization because after factorization, we can check each prime $p \mid n$ if $(p-1) \mid (n-1)$ holds. Denote $\omega(n)$ as the number of distinct prime factors of n. By the Hardy-Ramanujan theorem [HR17],

$$\omega(n) \sim \ln(\ln(n)) \tag{22}$$

Which makes the running time of checking if Korselt's criterion holds negligible compared to the prime factorization of n.

5.1 Prime factorization algorithms

A simple method to finding prime factors of n is to check every integer from 2 to $\sqrt{n}$ and see if it divides n. If it does, it is a prime factor, and we divide n by this value. If this value divides n more than once, we know n cannot be a Carmichael number since n isn't squarefree. This algorithm has a running time of $O(n^{1/2})$.

Another algorithm is the Sieve factorization algorithm which has a running time of $O(log(n))$ for each query n. The idea is to precompute the smallest prime that divides each integer up to a specified value larger than the largest query. This way, for each query n, simply divide n by its smallest prime factor, and repeat that on the new value of n until n becomes 1. Using Table 1 as an example, we will show how $n = 12$ is factorized. Let $sp[n]$ be the smallest prime that divides n. We know that $sp[12] = 2$ is a factor, so our new value is $12/sp[12] = 6$. Now, $sp[6] = 2$ is the second factor, so our new value is $6/sp[6] = 3$. Our third factor is $sp[3] = 3$, and this is our last factor because $3/sp[3] = 1$. The time complexity is $O(log(n))$ because the most amount of prime factors any number n can have is $log(n)$ (log is in base 2). Although this algorithm is a lot faster than the previous one, the time complexity of precomputing the array $sp[]$ is $O(nlog(n))$, and the auxiliary space is $O(n)$, so for large values of n, this algorithm may not be feasible.

There is another algorithm known as Pollard's rho algorithm [Pol75]. This technique cleverly uses Floyd's cycle-detection algorithm to find prime factors. The idea is to define a

n	2	3	4	5	6	7	8	9	10	11	12
sp[n]	2	3	2	5	2	7	2	3	2	11	2
$\frac{n}{sp[n]}$	1	1	2	1	3	1	4	3	5	1	6

Table 1: Example values for Sieve up to $n = 12$

polynomial $f(x)$ computed modulo n, where n is the number we want to factor, to serve as a pseudorandom generator. Then define two integer sequences a and b such that $a_0 = b_0$. During step $i > 0$ of the algorithm, we set $a_i = f(a_{i-1})$ and $b_i = f(f(b_{i-1}))$. Since f is taken modulo n, there is a finite amount of distinct values in the sequences a and b, so the values must repeat eventually. Once a repeated value occurs, the sequence will cycle. For each prime p that divides n, the sequences $a \pmod p$ and $b \pmod p$ must also cycle, and it's highly likely that its cycles are shorter than the cycles modulo n. By Floyd's algorithm, there exists a step j where $a_j \equiv b_j \pmod p$ for every prime $p \mid n$, so on every step $i > 0$, the rho algorithm checks if $\gcd(|a_i - b_i|, n) > 1$, and if this is true, then $\gcd(|a_i - b_i|, n)$ almost always is a prime factor of n. This will guarantee every prime $p \mid n$ will be found. This works well because for any prime p, the lengths of the cycles of $a \pmod p$ and $b \pmod p$ are expected to be $\sqrt{p}$, so the time complexity to factor n is said to be around $O(n^{1/4})$. There is still an issue with this algorithm because in some cases, $\gcd(|a_i - b_i|, n)$ is not prime. In this case, we need to run the algorithm again on the value $\gcd(|a_i - b_i|, n)$, but with a different polynomial f. We can keep on doing this until the likelihood of getting a non-prime factor is so low that it's negligible.

5.2 Pollard's rho on checking Carmichael numbers

In this section, we implemented Pollard's rho algorithm using C++ to check if a number n is a Carmichael number. First off, if n is even, then it is not a Carmichael number by Proposition 2.1. Then, we run Pollard's rho algorithm on n with a total of 10 different polynomials f. The number of polynomials used can be changed, depending on how accurate you need the algorithm to be. As the number of polynomials used increases, the accuracy increases exponentially, while the run time only increases linearly. After finding all the prime factors, we insert them into a set. If the total amount of factors is greater than the size of the set, n is not squarefree, therefore it's not a Carmichael number. From Proposition 2.2, we know that n must have at least 3 factors to be a Carmichael number, so if the size of the set of factors is less than 3, n is not Carmichael. Finally, we check if each prime p in the set satisfies

$(p-1) \mid (n-1)$. If it does, then n satisfies Korselt's criterion, which means it's a Carmichael number. Our pseudocode is given in Algorithm 1.

6 Further Research

Mathematicians have also studied other types of numbers with similar properties to Carmichael numbers.

6.1 Lucas-Carmichael numbers

Lucas-Carmichael numbers are squarefree numbers n that satisfy $p+1 \mid n+1$ for every prime $p \mid n$. These numbers were named after Édouard Lucas. They share many similar properties with the Carmichael numbers, for example, Proposition 2.1, 2.2, and 2.3 holds for both Carmichael and Lucas-Carmichael numbers. The first five Lucas-Carmichael numbers are 399, 935, 2015, 2915, and 4991. The distribution of these numbers seems to be more compact than that of the Carmichael numbers.

6.2 Quasi-Carmichael numbers

These numbers are a generalization of the Carmichael numbers and the Lucas-Carmichael numbers. A squarefree composite integer n is a Quasi-Carmichael number if every prime p that divides n satisfies $p+b \mid n+b$, with b being any nonzero integer. The subcase of $b=-1$ are the Carmichael numbers and the subcase of $b=1$ are the Lucas-Carmichael numbers. The smallest such Quasi-Carmichael number is 35, and we can see that if $b=-3$, it's true that $2 \mid 32$ and $4 \mid 32$.

6.3 Unsolved problems

There are still many problems relating to Carmichael numbers that are unsolved. The distribution and density of the Carmichael numbers are still not fully understood by mathematicians. It is also difficult to determine the smallest Carmichael numbers with k prime factors as k becomes large. We hope that solving these problems will not only help us understand more about Carmichael numbers, but also reveal more insight on other questions such as the distribution of primes.

References

[AGP94] W. R. Alford, A. Granville, and C. Pomerance. There are infinitely many Carmichael numbers. *Ann. Math. (2)*, 139(3):703–722, 1994.

[Che39] J. Chernick. On Fermat's simple theorem. *Bull. Am. Math. Soc.*, 45:269–274, 1939.

[Fri89] J. B. Friedlander. Shifted primes without large prime factors. *Number Theory and Applications*, pages 393–401, 1989.

[Har05] G. Harman. On the greatest prime factor of $p - 1$ with effective constants. *Math. Comput.*, 74(252):2035–2041, 2005.

[HR17] G. H. Hardy and S. Ramanujan. The normal number of prime factors of a number n. *Quart. J.*, 48:76–92, 1917.

[Kor99] A. Korselt. Probléme chinois, L'Intermédiaire des Mathématiciens. 6:142–143, 1899.

[Lar21] D. Larson. Bertrand's postulate for Carmichael numbers. *Int. Math. Res. Not.*, 2023(15):13072–13098, 2021.

[Pin93] R. G. E. Pinch. The Carmichael numbers up to 10^{15}. *Math. Comput.*, 61(203):381–391, 1993.

[Pin20] R. Pinch. Tables relating to Carmichael numbers. http://www.s369624816.websitehome.co.uk/rgep/cartable.html, 2020. Accessed: 2024-06-29.

[Pol75] J. M. Pollard. A Monte Carlo method for factorization. *BIT, Nord. Tidskr. Inf.-behandl.*, 15:331–334, 1975.

A Appendix

Algorithm 1 Carmichael Number checker using Pollard's Rho Algorithm

```
1: factors = ∅
2: amt = 0
3: function f(x, m, c)
4:     return (x × x + c) mod m
5: procedure rho(n, start, c)
6:     if c > 10 then
7:         factors.insert(n)
8:         amt ← amt + 1
9:         return
10:    if n == 1 then
11:        return
12:    x ← start
13:    y ← start
14:    d ← 1
15:    while d == 1 do
16:        x ← f(x, n, c)
17:        y ← f(f(y, n, c), n, c)
18:        d ← gcd(|x − y|, n)
19:    if d == n then
20:        call rho(n, start, c + 1)
21:    else
22:        call rho(n/d, start, c)
23:        call rho(d, start, c + 1)
24: function is_carmichael(n)
25:    if n mod 2 == 0 then
26:        return false
27:    call rho(n, 2, 1)
28:    if amt > factors.size() or amt < 3 then
29:        return false
30:    for each factor in factors do
31:        if (n − 1) mod (factor − 1) ≠ 0 then
32:            return false
33:    return true
34: input n
35: if is_carmichael(n) then
36:    print "Yes"
37: else
38:    print "No"
```