# CARMICHAEL FUNCTIONS

by

Eric Bach

Computer Sciences Technical Report #939

June 1990

# Carmichael Functions

Eric Bach
Computer Sciences Department
University of Wisconsin
Madison, WI 53706
USA

June 15, 1990

**Abstract.** A Carmichael number is a composite number $n$ such that for every $a$ relatively prime to $n$, $a^{n-1} \equiv 1$ modulo $n$. It is unknown whether infinitely many Carmichael numbers exist. The analogous question for polynomials over a finite field with $q$ elements asks if there are infinitely many reducible polynomials $f$ such that if $a$ and $f$ are coprime, then $a^{q^{\deg f}-1} \equiv 1$ modulo $f$. We show this is true, and that such polynomials are rare compared to irreducibles.

Technical Report #939, Computer Sciences Department, University of Wisconsin–Madison.

This research was sponsored by the National Science Foundation, via grant DCR-8552596.

## 1. Introduction.

Composite numbers $n$ such that $a^{n-1} \equiv 1$ modulo $n$ holds for all $a$ relatively prime to $n$ are called Carmichael numbers, or absolute pseudoprimes. The smallest Carmichael number is 561, and by now examples are known into the thousands of digits [3]. The existence of these numbers implies that a naive primality test based only on the converse of Fermat's theorem is doomed to failure.

It is not known if there are infinitely many Carmichael numbers. Heuristics, however, suggest that this should be the case. For example, if $6n+1$, $12n+1$, and $18n+1$ are all prime, then their product is a Carmichael number; a probabilistic argument suggests that this should occur infinitely often.

The purpose of this note is to examine the corresponding question for polynomials over a finite field. That is, when $k$ is a finite field with $q$ elements, we ask if there are infinitely many reducible polynomials $f$ in $k[X]$ with the property that for all $a$ relatively prime to $f$, $a^{q^{\deg f}-1} \equiv 1$ modulo $f$. We characterize such polynomials, and use our characterization to show that they are infinite in number, but rarer than irreducibles.

Our results give an example of an unsolved question about the integers whose analog in $k[X]$ is easy to settle. However, they seem to have no algorithmic applications whatsoever. Indeed, there is already a fast *deterministic* algorithm to test whether a polynomial $f$ in $k[X]$ is irreducible [1], requiring less computation than the evaluation of $a^{q^{\deg f}-1}$ modulo $f$.

We are unaware of any previous work on this problem. Hellegouarch, however, has studied analogs in $k[X]$ to the Miller-Rabin and Solovay-Strassen primality tests [7].

## 2. Notation and Background.

We let $k = \mathbb{F}_q$ denote a finite field with $q$ elements, and $k[X]$ the ring of polynomials in one variable with coefficients in $\mathbb{F}_q$. We let $p$ denote the characteristic of $q$.

Following [7], if $a, f \in k[X]$, we call $f$ a *pseudoprime to base $a$* if $f$ is reducible, and $a^{q^n-1} \equiv 1$ mod $f$. (Here $n$ denotes the degree of $f$.) If $f$ is a pseudoprime to all possible bases, that is, for all $a$ relatively prime to $f$, we call $f$ a *Carmichael function.*

We let $I_q(n)$ denote the number of monic irreducible polynomials in $k[X]$ of degree $n$. We have

$$I_q(n) = \frac{1}{n} \sum_{d|n} \mu(n) q^{n/d},$$

where $\mu$ denotes the Möbius function (see [9], p. 84). This implies that

$$\frac{q^n - q^{n/2+1}}{n} \leq I_q(n) \leq \frac{q^n}{n},$$

from which it follows that $I_q(n)/q^n \sim 1/n$.

We let $d(n)$ denote the number of positive divisors of $n$. Typically, $d(n)$ is around $\log n$. We will need the following upper bound:

$$d(n) \leq n^{O(1/\log\log n)}.$$

(see [8], p. 262). We will call a divisor $d$ of $n$ *proper* if $1 \leq d < n$.

### 3. Characterization.

We first characterize Carmichael functions, with the aid of the following lemma.

**Lemma 3.1.** $q^m - 1$ divides $q^n - 1$ iff $m \mid n$.

*Proof.* Write both numbers in base $q$ and apply long division. ■

**Lemma 3.2.** A reducible polynomial $f$ is a Carmichael function if and only if: i) $f$ is squarefree, and ii) for every irreducible $g$ dividing $f$, $\deg(g) \mid \deg(f)$.

*Proof.* The sufficiency of conditions i) and ii) follows from the Chinese remainder theorem, and Lemma 3.1.

If $f$ is not squarefree, then $k[X]/(f)^*$ is a group whose order is divisible by $p$, and hence contains elements of order $p$. Therefore, $f$ cannot be a Carmichael function. This shows that condition i) is necessary.

Now, let $f = f_1 \dots f_r$ be the irreducible factorization of a squarefree polynomial of degree $n$, and let $n_i$ be the degree of $f_i$. Choose any index $i$ and let $a$ be an element of $k[X]/(f)$ whose homomorphic image generates the multiplicative group $k[X]/(f_i)^*$. If $a^{q^n-1} \equiv 1$ modulo $f$, then $a^{q^n-1} \equiv 1$ modulo $f_i$. Because $a$ is a generator, $q^{n_i} - 1$ divides $q^n - 1$. By Lemma 3.1, $n_i$ divides $n$. This shows the necessity of condition ii). ■

**Theorem 3.3.** If $k$ is a finite field, then $k[X]$ contains infinitely many Carmichael functions.

*Proof.* With one exception (namely, degree 2 for $q = 2$), $\mathbb{F}_q[X]$ contains at least two different irreducible polynomials of each degree. Therefore, for every even $n \geq 6$, there is a Carmichael function of degree $n$ in $\mathbb{F}_q[X]$. ■

We note that $k[X]$ does not contain Carmichael functions of every degree, or even every sufficiently large degree. For, let $r$ be a prime larger than $q$. Then any polynomial of degree

$r$ satisfying condition ii) of Lemma 3.2 must be either irreducible or composed of linear factors, which cannot all be distinct. Therefore there are no Carmichael functions of degree $r$.

Even though there is no *deterministic* polynomial time algorithm to factor polynomials over finite fields, the set of Carmichael functions can be recognized in deterministic polynomial time, by combining a test for repeated factors with distinct-degree factorization. (See [1] for descriptions of these algorithms.) This is contrast with the situation for the integers; there is no known deterministic polynomial-time algorithm to recognize Carmichael numbers.

## 4. Density.

In this section we consider the density of Carmichael functions of degree $n$; we show that as $n \to \infty$, they are strictly rarer than irreducibles. We will use the idea, which goes back at least to Frobenius [6], of comparing the factorization pattern of a polyomial to the cycle structure of a permutation. For this reason, some preliminary remarks on permutations are necessary.

Let $\Sigma_n$ denote the group of permutations on $\{1, \ldots, n\}$. Every element of $\Sigma_n$ can be written in one and only one way as a product of disjoint cycles. Let $\pi$ denote a partition of $n$ containing $m_i$ copies of $n_i$, for $i = 1, \ldots, l$. We let $F_\pi$ denote the fraction of permutations with $m_i$ cycles of length $n_i$. Then

$$F_\pi = \prod_{i=1}^{l} \frac{1}{n_i^{m_i} m_i!}.$$

We illustrate this with an example. Take $n = 7$, and let $\pi = 3 + 2 + 2$. We wish to count the permutations of $\{1, \ldots, 7\}$ whose cycle structure is

$$(* * *)(**)(**)$$

We can fill in the entries in 7! ways. Now we count the number of ways such an assignment gives the same permutation. Without changing the permutation, we can cyclically permute elements in the 3-cycle and in each of the 2-cycles; we can also swap the two 2-cycles. Therefore the number of permutations with this cycle structure is $7!/(3 \cdot 2^2 \cdot 2!)$, so $F_\pi = 1/24$.

For estimates involving the cycle lengths of permutations, it is useful to consider the following "random bisection" process, which generates a random permutation of of $\{1, \ldots, n\}$. Choose a length $i$ uniformly with $1 \leq i \leq n$. Then write down 1, followed by $i - 1$ distinct elements chosen at random from $\{2, \ldots, n\}$. This determines one of the cycles; recursively generate a random permutation of the $n - i$ numbers not appearing in that cycle. (This may be justified using known facts about random permutations; see [4], p. 257.)

With these preliminaries taken care of, we now turn to polynomials.

**Lemma 4.1.** Let $n$ be a fixed positive integer, and let $\pi$ be a partition of $n$, containing $m_i$ copies of $n_i$ for $i = 1, \ldots, l$. Let $S_\pi$ denote the fraction of monic squarefree polynomials in $k[X]$ whose factorization pattern conforms to $\pi$. Then $S_\pi \le F_\pi$.

*Proof.* We have

$$S_\pi = \frac{1}{q^n} \prod_{i=1}^{l} \binom{I_q(n_i)}{m_i} \le \frac{1}{q^n} \prod_{i=1}^{l} \frac{(I_q(n_i))^{m_i}}{m_i!}.$$

Because $I_q(m) \le q^m/m$, this is at most

$$\frac{1}{q^n} \prod_{i=1}^{l} \frac{q^{m_i n_i}}{n_i^{m_i} m_i!} = F_\pi.$$

■

We note that this estimate is very good when $q$ is large; in fact, the fraction of degree $n$ squarefree polynomials with factorization pattern conforming to $\pi$ is asymptotic to $F_\pi$ as $q \to \infty$. Furthermore, $1 - 1/q$ of the degree $n$ polynomials in $\mathbb{F}_q[X]$ are squarefree; thus $F_\pi$ gives the asymptotic fraction of polynomials whose factors match $\pi$.

**Theorem 4.2.** Let $q$ be a prime power. Let $C_q(n)$ denote the fraction of degree $n$ Carmichael functions in $\mathbb{F}_q[X]$. Then

$$C_q(n) \le n^{-2+O(1/\log\log n)}.$$

Furthermore, if we restrict $n$ to even numbers, then

$$C_q(n) = \Omega(n^{-2}).$$

*Proof.* By Lemma 4.1, we have

$$C_q(n) \le \sum_\pi F_\pi,$$

where the sum is taken over all partitions of $n$ composed entirely of proper divisors of $n$. This sum is the probability that the random bisection process always chooses proper divisors of $n$ for cycle lengths; we estimate it as follows. Suppose that lengths $l, m, \ldots$ are chosen, which are all proper divisors of $n$. Then the following two events must take place:

i) $l \le n/2$, and $l \mid n$; call this event $A$.

ii) $l \mid n$; call this event $B$.

Recall that $d(n)$ denotes the number of divisors of $n$. There are $d(n)-1$ proper divisors of $n$, so

$$\Pr[A] = \frac{d(n)-1}{n}.$$

Now, if $l \leq n/2$, there are at least $n/2$ choices for $m$, including all proper divisors of $n$. Therefore

$$\Pr[B \mid A] \leq \frac{d(n)-1}{n/2}.$$

It now follows that

$$C_q(n) \leq \Pr[A \cap B] = \Pr[A]\Pr[B \mid A] = \frac{2(d(n)-1)^2}{n^2}.$$

The first assertion follows from this and the estimate on $d(n)$.

To prove the second assertion, we note that if $n$ is even, then

$$C_q(n) \geq \frac{1}{q^n}\binom{I_q(n/2)}{2} \sim \frac{2}{n^2}.$$

■

By refining our proof, the upper bound of Theorem 4.2 can be sharpened, as follows. Let $r$ denote the smallest prime divisor of $n$. Then we have

$$C_q(n) \leq \frac{r^{r-1}(d(n)-1)^r}{n^r(r-1)!}.$$

This reduces to the bound of Theorem 4.2 when $r = 2$, but is more accurate when $n$ has no small prime divisors. For example, if $n$ is prime, it yields $C_q(n) \leq 1/n!$ , which is the best possible estimate.

It is of interest to compare our estimates with known results about the density of Carmichael numbers. Let $C(x)$ denote the fraction of positive integers $\leq x$ that are Carmichael numbers. Pomerance [11] has shown that

$$C(x) \leq x^{-\frac{\log\log\log x}{\log\log x}(1+o(1))},$$

and conjectures a lower bound of the same type. Thus, $C(x)$ is bounded by a function inversely proportional to a slowly decreasing power of $x$. The lower bound of Theorem 4.2 shows that such an estimate cannot hold in $k[X]$; in this sense, Carmichael functions are more common than Carmichael numbers.

The sum $\sum_\pi F_\pi$, taken over all partitions with proper divisors of $n$ for cycle lengths, is actually the asymptotic fraction of Carmichael functions of degree $n$, as $q \to \infty$. We will denote this asymptotic fraction by $C_\infty(n)$. For moderate values of $n$, one can enumerate the partitions of $n$ (an algorithm using $O(1)$ time per partition appears in [12]), and compute $C_\infty(n)$. Table 1 gives the results of this computation for $n \leq 33$, which displays the extreme irregularity of $C_\infty(n)$ as a function of $n$.

Table 1. The asymptotic density of Carmichael functions.

| $n$ | $C_\infty(n)$ | $n$ | $C_\infty(n)$ |
|---|---|---|---|
| 2 | 0.500000 | 18 | 0.042641 |
| 3 | 0.166667 | 19 | $8.2 \times 10^{-18}$ |
| 4 | 0.416667 | 20 | 0.025069 |
| 5 | 0.008333 | 21 | 0.001241 |
| 6 | 0.383333 | 22 | 0.004214 |
| 7 | 0.000198 | 23 | $3.9 \times 10^{-23}$ |
| 8 | 0.154365 | 24 | 0.043627 |
| 9 | 0.015898 | 25 | 0.000003 |
| 10 | 0.065950 | 26 | 0.002966 |
| 11 | $2.5 \times 10^{-8}$ | 27 | 0.000328 |
| 12 | 0.195361 | 28 | 0.006118 |
| 13 | $1.6 \times 10^{-10}$ | 29 | $1.1 \times 10^{-31}$ |
| 14 | 0.016807 | 30 | 0.015565 |
| 15 | 0.006673 | 31 | $1.2 \times 10^{-34}$ |
| 16 | 0.029023 | 32 | 0.003843 |
| 17 | $2.8 \times 10^{-15}$ | 33 | 0.000140 |

Our function $C_\infty(n)$ is connected with a classic problem in group theory; it is the fraction of elements $x \in \Sigma_n$, excluding $n$-cycles, that satisfy $x^n = 1$. Because $1/n$ of the permutations on $\{1, \ldots, n\}$ are $n$-cycles, our bound may be rephrased in group-theoretic terms as follows:

$$\#\{x \in \Sigma_n : x^n = 1\}/n! = 1/n + O(1/n^{2-\epsilon}),$$

for every $\epsilon > 0$. Many authors have investigated the solutions in finite groups of equations such as $x^d = 1$. (See [5] for a survey.) Although asymptotic formulas for the number of solutions to $x^d = 1$ in $\Sigma_n$ are known (see [2], [10], [13], [14]), these formulas only apply when $d$ is fixed and $n \to \infty$. We are unaware of any analytic estimates that are applicable to problems such as ours, where $d$ grows with $n$. Although Table 1 indicates that any such results would need to take the factorization of $n$ into account, it would be of interest to find the generating function of $C_\infty(n)$ and thereby obtain sharper estimates than those given here.

Although $C_\infty(n) < 1/n$ for sufficiently large $n$, we note that this is not true of all $n$. In particular, $C_\infty(n) \geq 1/n$ for $n = 2, 4, 6, 12, 24$; we believe, but have not proved, that these are only cases.

Another question of interest concerns the quality of the approximation $C_q(n) \doteq C_\infty(n)$. On numerical grounds, we believe that $C_q(n)$ is an increasing function of $q$, for every $n$. If this were true, then one could obtain lower bounds on $C_q(n)$ by computations similar to those that produced Table 1.

## Acknowledgements.

Thanks to Jeffrey Shallit and Louis Solomon for pointing out crucial references.

## References.

1. E. R. Berlekamp, Factoring polynomials over large finite fields, Mathematics of Computation 24, 713-735, 1970.
2. S. Chowla, I. N. Herstein, and W. K. Moore, On recursions connected with symmetric groups I, Canadian Journal of Mathematics 3, 328-334, 1951.
3. H. Dubner, A new method for producing large Carmichael numbers, Mathematics of Computation 53, 411-414, 1989.
4. W. Feller, An Introduction to Probability Theory and its Applications, Volume 1, Third Edition, John Wiley and Sons, 1968.
5. H. Finkelstein, Solving equations in groups: a survey of Frobenius' theorem, Periodica Mathematica Hungarica 9, 187-204, 1978.
6. F. G. Frobenius, Über Beziehungen zwischen den Primidealen eines algebraischen Körpers und den Substitutionen seiner Gruppe, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin, 689-703, 1896. [Reprinted in Gesammelte Abhandlungen, vol. II, pp. 719-733.]
7. Y. Hellegouarch, Loi de réciprocité, critère de primalité dans $\mathbb{F}_q[t]$, Comptes Rendus Mathematiques de l'Academie des Sciences [Canada] 8, 291-296, 1986.
8. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Fifth Edition, Oxford University Press, 1979.
9. K. Ireland and M. Rosen, A Classical Introduction to Modern Number Theory, Springer-Verlag, 1982.

10. L. Moser and M. Wyman, On solutions of $x^d = 1$ in symmetric groups, Canadian Journal of Mathematics 7, 159-168, 1955.

11. C. Pomerance, On the distribution of pseudoprimes, Mathematics of Computation 37, 587-593, 1981.

12. E. M. Reingold, J. Nievergelt, and N. Deo, Combinatorial Algorithms, Prentice-Hall, 1977.

13. L. M. Volynets, Number of solutions of the equation $x^s = e$ in the symmetric group, Matematicheskie Zametki 40, 155-160, 1986. [English translation in Mathematical Notes 40, 586-589, 1987.]

14. H. S. Wilf, The asymptotics of $e^{P(z)}$ and the number of elements of each order in $S_n$, Bulletin of the American Mathematical Society (New Series) 15, 228-232, 1986.