Advances in tabulating Carmichael numbers

Andrew Shallue and Jonathan Webster

Illinois Wesleyan University and Butler University, USA
ashallue@iwu.edu, jewebste@butler.edu

July 18, MIT, ANTS 2024

Fermat's Little Theorem

Theorem (Fermat)

If p is prime, then $a^p \equiv a \pmod{p}$.

Theorem (Contrapositive of FLT)

If $a^n \not\equiv a \pmod{n}$, then n is composite.

Definition (Carmichael number)

A Carmichael number is a composite integer n satisfying $a^n \equiv a \pmod{n}$ for any a.

Named after Robert Carmichael (1910) by Nicolaas Beeger in 1950.

Why not Šimerka numbers?

In 1885, Vàolav Šimerka found the first 7 Carmichael numbers.

> Poučka tato dle vynálezce řečená Fermatovou jest jednou z nejdůležitějších v neurčité analytice; neudává však charakteristickou známku kmenných čísel, (jíž by se tato ode všech ostatních lišila), ježto podobně i při některých dělitelných číslech bývá. Tak na př. při $561 = 3 . 11 . 17$, $b = 2$ nalezneme
>
> $$2_{10} = -98,\ 2_{20} = 67,\ 2_{40} = 1,\ (2_{40})^{14} = 2_{560} = 1.$$
>
> Tolikéž u čísel
>
> $1105 = 5 . 13 . 17$, $1729 = 7 . 13 . 19$, $2465 = 5 . 17 . 29$,
> $2821 = 7 . 13 . 31$, $6601 = 7 . 23 . 41$, $8911 = 7 . 19 . 67$ a j. v.,
> kdykoli b s modulem nesoudělné jest.

Korselt's Criterion

Theorem (Korselt - 1899)

A composite number n is a Carmichael number if and only if n is squarefree and $(p-1)|(n-1)$ for all prime divisors p of n.

Definition

Let $\lambda(n)$ be the Carmichael function or the reduced totient function. For n an odd square-free integer

$$\lambda(n) = \mathrm{lcm}\{p_1 - 1, p_2 - 1, \ldots, p_k - 1\}$$

where p_i are distinct prime divisors of n.

Theorem

A composite number n is a Carmichael number if and only if n is squarefree and $\lambda(n)|(n-1)$.

Motivation

- Empirical versus Theoretical. Count of Carmichael numbers less than B:
  - Erdős conjectured:
$$B \exp\left(\frac{-k \log B \log\log\log B}{\log\log B}\right)$$
  - Best lower bound due to Harmon is $B^{1/3}$.
  - Empirically, Harmon's result seems more accurate.

- Other similar but harder problems:
  - Lehmer's conjecture: $\phi(n) | (n-1)$
  - PSW number: base 2 Fermat pseudoprime, Fibonnacci pseudoprime, and $n \equiv 2, 3 \pmod{5}$.
  - Williams' number: Carmichael number, an absolute Lucas pseudoprime, and $(d|n) = -1$

OEIS A055553

Count of Carmichael numbers by order of magnitude.

10^3	1	Šimerka (1885)
10^4	7	Šimerka (1885)
10^5	16	???
10^6	43	???
10^7	105	???
10^8	255	???
10^9	646	Swift (1975), Math. Comp.
10^{10}	1547	PSW (1980), Math. Comp.
10^{11}	3605	Guthmann (1992)
10^{12}	8241	Jaesechke (1990), Math. Comp

OEIS A055553 (cont'd)

10^{13}	19279	Keller (1988)
10^{14}	44706	
10^{15}	105212	Pinch (1993), Math. Comp.
10^{16}	246683	Pinch (1998)
10^{17}	585355	Pinch (2005)
10^{18}	1401644	Pinch (2006)
10^{19}	3381806	
10^{20}	8220777	Pinch (2006), ANTS 7 poster
10^{21}	20138200	Pinch (2007)
10^{22}	49679870	Goutier (2022), S.W. (2024)

Problem

Given bound B, tabulate all Carmichael numbers up to B.

Problem

Given an integer P (called a pre-product), determine the finite list of prime-pairs (q, r) such that Pqr is a Carmichael number.

Problem

Determine the computational complexity of either tabulation problem.

How do you tabulate?

We construct $n = p_1 p_2 \dots p_d$ in factored form with $d > 2$ prime factors.

We let

$$P = \prod_{i=1}^{d-2} p_i, q = p_{d-1}, \text{ and } r = p_d$$

so that $n = Pqr$ is a Carmichael number and $\gcd(P, \phi(P)) = 1$.

There are two cases to consider:

- P is small - find q and r at the same time.
- P is large - exhaustive search for q, compute r.

P is small

Theorem (Proposition 2 of Pinch)

There are integers $2 \le D < P < C$ such that, putting $\Delta = CD - P^2$, we have

$$q = \frac{(P-1)(P+D)}{\Delta} + 1, \tag{1}$$

$$r = \frac{(P-1)(P+C)}{\Delta} + 1, \tag{2}$$

$$P^2 < CD < P^2\left(\frac{p_{d-2}+3}{p_{d-2}+1}\right). \tag{3}$$

- CD pairs - Pinch (Math. Comp. 1993)
- $D\Delta$ pairs - S.W. (ANTS 2022)

Asymptotic costs: CD and $D\Delta$ methods

Let p be the largest prime dividing a fixed a pre-product P.

Theorem

The number of CD pairs is $O((P^2 \log P)/p) = O(P^{2-\frac{1}{d-2}} \log P)$.

Theorem

The number of $D\Delta$ pairs is $O(\tau(P-1)P \log P)$.

- Outer loop is the same for both.
- Hybrid method: enter inner loop of cheaper method.
- Choose based on $P^2/(pD)$ and $\tau((P-1)(P+D))$.

Hybrid: $D\Delta$ - CD

We generated Carmichael numbers exceeding the intended bound.

For $P < 7 \cdot 10^7$, the largest Carmichael number found was

$$69999133 \cdot 4899878690750821 \cdot 171493630078866294519097 =$$
$$58\ 82013\ 03152\ 54068\ 53935\ 58087\ 37155\ 82013\ 87008\ 71721.$$

Found with $D = 2$ and $\Delta = 1$.

Hybrid method would choose the lesser amount of work:

- Iterate through 768 candidates for Δ.
- Iterate through approximately $7 \cdot 10^7$ values of C.

Aside: Chernick-like families

Theorem (Chernick (1939))

If $p = 6m+1$, $q = 12m+1$, and $r = 18m+1$ are prime, then pqr is a Carmichael number.

Theorem

If p, $q = p^2 + p - 1$, and $r = \frac{(p^3+p^2-p+1)}{2}$ are prime, then pqr is a Carmichael number.

Example

$$99999437 \cdot 9999887500316407 \cdot 499991560047488910931993 =$$
$$499\ 99518\ 00193\ 60158\ 52677\ 14742\ 32654\ 10341\ 56276\ 99201$$

P is large

Given a large P, exhaustively consider all primes $q \in (p, \sqrt{B/P})$.

Given P and q, use two conditions on $r-1$:

- $r - 1 \mid Pq - 1$, and
- $r - 1 \equiv r^\star - 1 \pmod{\lambda(Pq)}$ where $0 < r^\star < \lambda(Pq)$ is the modular inverse of $Pq \pmod{\lambda(Pq)}$.

Pinch's two approaches

Let

$$k = \min\left\{\frac{Pq-1}{\lambda(Pq)}, \frac{B}{Pq\lambda(Pq)}\right\}.$$

If k is small enough, consider $r = r^\star + j\lambda(Pq)$ for $0 \le j \le k$.

Otherwise balance:

- small r: $r = r^\star + j\lambda(Pq)$ for small j
- large r: $r = (Pq-1)/f + 1$ for small f

Complexity: $O\left(\sqrt{\dfrac{Pq}{\lambda(Pq)}}\right)$

New observation: Divisors in Residue Class

Consider $g = \gcd(r^\star - 1, \lambda(Pq))$.

Let $\mathcal{R}_1 = (r^\star - 1)/g$, $\mathcal{L} = \lambda(Pq)/g$, and $\mathcal{P} = (Pq - 1)/g$.

New problem: Find factors of $\mathcal{P}$ that are $\mathcal{R}_1 \pmod{\mathcal{L}}$.

So, $(\mathcal{R}_1 + k_1\mathcal{L})(\mathcal{R}_2 + k_2\mathcal{L}) = \mathcal{P}$ implies $\mathcal{R}_2 \equiv \mathcal{P}\mathcal{R}_1^{-1} \pmod{\mathcal{L}}$.

- Search for small r candidates in $r^\star + j\lambda(Pq)$
- Search for large r candidates in $(Pq - 1)/(\mathcal{R}_2 + j\mathcal{L}) + 1$

Complexity: $O\left(\frac{\sqrt{gPq}}{\lambda(Pq)}\right)$

Complexity: If $\mathcal{L}^2 > \mathcal{P}$, the divisors are $\mathcal{R}_1$ and $\mathcal{P}/\mathcal{R}_2$ and are found in polynomial time.

New observation: Divisors in Residue Class

Due to stronger results, we can factor in polynomial time if:

- $\mathcal{L}^3 > \mathcal{P}$, due to Lenstra.
- $\mathcal{L}^4 > \mathcal{P}$, due to Coppersmith, et al.

Question: Are there even faster algorithms for this problem?

- Find divisors bounded in size.
- Find divisors in a residue class.
- Asymptotic versus practical.

New observation: Worst case inputs

What are the worst-case inputs to this algorithm?

- If $\lambda(Pq) | (Pq - 1)$, then Pq is a Carmichael number.
- This is a general factoring problem (there is no residue class information).

How small can $\lambda(Pq)$ be?

- Arbitrarily small as a power of Pq.

How often do these inputs occur?

- Can use prior tabulations to gauge this count.
- The asymptotic count of Carmichael numbers is an open question.

New observation: Timing Results

How much better is this?

B	Method 1 (SW)	Method 2 (Pinch)
10^{14}	25	25
10^{15}	192	189
10^{16}	1324	1306
10^{17}	8752	8865
10^{18}	55631	56072
10^{19}	361983	364816

Many cases have $Pq\lambda(Pq) > B$ and the improved method is not invoked very often.

Heuristic Cost

Assume that, on average the cost to find r given Pq is polynomial time. We count the number of valid Preproducts Pq.

We say a number is *B-admissible* if

1. P is cyclic, and
2. $Pp_{d-1}^2 < B$.

Theorem (with Sungjin Kim)

The count of B-admissible preproducts P is at most

$$B \exp\left(\left(-\frac{1}{2} + o(1)\right)\sqrt{\log B \log\log B}\right).$$

Continuing Work

- For the small case:
  - Completed small case for preproduct $P < 7.0 \cdot 10^7$.
    - 35985331 were found
    - 1202914 were less than 10^{22}
  - New goal for small case: $P < 10^8$.
  - Asymptotically faster small case?
- New goals for the large case:
  - Faster preproduct generation.
  - Better parallelization scheme.
  - (Asymptotically) fewer preproducts.
  - Better analysis.
  - New tabulation up to 10^{23} (or 10^{24}).

Thank you!