Constructing a ten billion factor Carmichael number

Steven Hayman and Andrew Shallue *Illinois Wesleyan University*

Introduction

We have constructed a Carmichael number with a record number of prime factors. Such constructions are showcases for recent advances in subset-product algorithms that work on problems of very high density [2, 4, 6]. Our new contribution is to show that these techniques can be applied to a subgroup of the natural group that arises from the standard Erdős construction.

Recent records

Here are Carmichael numbers with a record number of prime factors. We start with Löh and Niebuhr [5] since their backtracking algorithm vastly outperformed previous methods.

Löh and Niebuhr (1996): $1,101,518$ prime factors

Alford and Grantham (2003): $19,565,300$ prime factors (*)

Hayman and Shallue (Oct 2011): $1,021,449,117$ prime factors

Hayman and Shallue (Nov 2011): $10,333,229,505$ prime factors (**)

(*) In addition, a Carmichael number with k prime factors for all $3 \le k \le 19565220$.

(**) Stored as ASCII with one prime per line, this requires nearly 311 gigabytes.

Erdős construction

A positive integer n is Carmichael if it is a Fermat pseudoprime to the base a for all a coprime to n. For constructions we use the following equivalent definition.

Definition 1 (Korselt criterion) *An integer n is Carmichael if it is squarefree and $p-1 \mid n-1$ for all primes p dividing n.*

The following construction of Erdős is widely used both in theory [1] and in practice [7].

1. Choose $\Lambda = \prod_{i=1}^{r} q_i^{h_i}$ where $q_1 \dots q_r$ are the first r primes in order and the h_i are all at least 1 and non-increasing.
2. Construct the set $\mathcal{P} = \{p \text{ prime} \ : \ p-1 \mid \Lambda, p \nmid \Lambda\}$. Let $N = |\mathcal{P}|$.
3. Construct Carmichael n as a product of primes in $\mathcal{P}$ in one of two ways:
   (a) Find a subset $\mathcal{S}$ of $\mathcal{P}$ such that
   $$\prod_{p \in \mathcal{S}} p \equiv 1 \mod \Lambda \ .$$
   Then by Definition 1 $n = \prod_{p \in \mathcal{S}} p$ is Carmichael.
   (b) Alternatively, let $b \equiv \prod_{p \in \mathcal{P}} p \mod \Lambda$ and find a subset $\mathcal{T}$ of $\mathcal{P}$ such that
   $$\prod_{p \in \mathcal{T}} p \equiv b \mod \Lambda \ .$$
   Then $n = \prod_{p \in \mathcal{P} \setminus \mathcal{T}} p$ is Carmichael.

Subset product problem

Let G be an abelian group, and let $a_1, \dots, a_n$ be elements of G. The *subset product problem* is to find a subset of the a_i that product to the identity in G (more generally, any element of G).

Definition 2 *The* density *of a subset product problem is given by*

$$\frac{n}{\log_2(|G|)} \ .$$

Solutions will be rare unless density is greater than 1. Problems with density 1 are the most difficult. Problems with greater density are easier, both in the sense that new algorithms are applicable and in the sense that existing algorithms perform better.

Kuperberg idea

Discovered independently by Flaxman and Przydatek [2], though they didn't extend to make the best possible improvement.

Assumptions: Group G has a sequence of $\sqrt{\log_2 |G|}$ subgroups G_k with factor groups of size $2^{\sqrt{\log_2 |G|}}$. Have a_i distributed independently and symmetrically (i.e. $\Pr[a_i = x] = \Pr[a_i = x^{-1}]$ in G), and $n > 2^{2\sqrt{\log_2 |G|}}$.

Algorithm: At level k, have $b_i \in G_k$ where b_i are products of the a_i. Pair b_i so their products are in G_{k+1}. At level $k = \sqrt{\log_2 |G|}$ a product is the identity, and this is the solution.

Complexity: Will succeed with high probability while requiring $O(2^{2\sqrt{\log_2 |G|}})$ group operations and space for $O(2^{2\sqrt{\log_2 |G|}})$ group elements.

New contribution

Inspiration from [5]: elements of $\mathcal{P}$ not uniformly distributed modulo Λ. Instead distributed symmetrically, with $p \equiv 1 \bmod q_i^{h_i}$ more likely. Note the proportion of divisors of Λ divisible by $q_i^{h_i}$ is $\frac{1}{h_i+1}$.

Definition 3 *Define a subgroup G' of G as $(\mathbf{Z}/\hat{\Lambda}\mathbf{Z})^{\times}$ with $\hat{\Lambda} = \prod_{i=1}^{m} q_i^{h_i}$, where*

$$m = \min_{1 \le j \le r} j \text{ such that } N \cdot \prod_{i=j+1}^{r} \frac{1}{h_i+1} > (\log \Lambda) 4^{\sqrt{\sum_{i=1}^{j} h_i \log q_i}} \ .$$

While constructing $\mathcal{P}$ we pick out elements in G', and the claim is that there are enough for the Kuperberg algorithm to still work. Under reasonable assumptions the complexity in terms of N is

$$2^{O(\sqrt{(\log N)(\log\log N)^2})} \ .$$

Carmichael with ten billion prime factors

$$\Lambda = 2^{16} \cdot 3^7 \cdot 5^5 \cdot 7^4 \cdot 11^3 \cdot 13^2 \cdot 17^2 \cdot 19^2 \cdot 23^2 \cdot 29^2 \cdot 31 \cdot 37 \cdot 41 \cdot 43 \cdot 47 \cdot 53 \cdot 59 \cdot 61 \cdot 67 \cdot 71 \cdot 73 \cdot 79 \cdot 83 \cdot 89 \cdot 97$$

With this choice $N = 10333230324$. It took 98 seconds to find an appropriate subset of size 819 to exclude. The resulting Carmichael number has 10333229505 prime factors and its decimal expansion would have 295 billion digits.

With $\Lambda \approx 2^{191}$ applying Kuperberg directly would take approximately $2^{2\sqrt{191}} = 2^{28}$ operations.

In practice algorithm chose $m = 17$ so

$$\hat{\Lambda} = 2^{16} \cdot 3^7 \cdot 5^5 \cdot 7^4 \cdot 11^3 \cdot 13^2 \cdot 17^2 \cdot 19^2 \cdot 23^2 \cdot 29^2 \cdot 31 \cdot 37 \cdot 41 \cdot 43 \cdot 47 \cdot 53 \cdot 59 \ .$$

With $\hat{\Lambda} \approx 2^{141}$ we instead anticipate $2^{2\sqrt{141}} = 2^{24}$ operations. Indeed, the algorithm succeeded with only 16 million primes in G'.

Higher order Carmichael numbers

Holy Grail problem is to construct higher order Carmichael numbers, of which the $620 problem of Pomerance, Selfridge, and Wagstaff is a special case. See [3] for a definition and the following equivalent condition.

Theorem 1 *Composite n is a Carmichael of order m if and only if n is square free and for every prime divisor p of n and for every $1 \le r \le m$, there exists $i \ge 0$ with $n \equiv p^i \mod (p^r - 1)$.*

Build the set

$$S_m(\Lambda) = \{p \text{ prime}, p^r - 1 \mid \Lambda \text{ for all } 1 \le r \le m\}$$

and find a subset that products to 1 modulo Λ. More difficult since density is lower.

References

[1] W. R. Alford, Andrew Granville, and Carl Pomerance, *There are infinitely many Carmichael numbers*, Ann. of Math. (2) **139** (1994), no. 3, 703–722.

[2] Abraham D. Flaxman and Bartosz Przydatek, *Solving medium-density subset sum problems in expected polynomial time*, STACS 2005, Lecture Notes in Comput. Sci., vol. 3404, Springer, Berlin, 2005, pp. 305–314.

[3] Everett W. Howe, *Higher-order Carmichael numbers*, Math. Comp. **69** (2000), no. 232, 1711–1719.

[4] Greg Kuperberg, *A subexponential-time quantum algorithm for the dihedral hidden subgroup problem*, SIAM J. Comput. **35** (2005), no. 1, 170–188 (electronic).

[5] Günter Löh and Wolfgang Niebuhr, *A new algorithm for constructing large Carmichael numbers*, Math. Comp. **65** (1996), no. 214, 823–836.

[6] David Wagner, *A generalized birthday problem (extended abstract)*, Advances in Cryptology – CRYPTO 2002, Lecture Notes in Comput. Sci., vol. 2442, Springer, Berlin, 2002, pp. 288 – 303.

[7] Ming Zhi Zhang, *A method for finding large Carmichael numbers*, Sichuan Daxue Xuebao **29** (1992), no. 4, 472–479.