An Algorithm to Find Sums of Consecutive Powers of Primes

Cathal O’Sullivan
Mathematics, Statistics, and Actuarial Science Department
Computer Science and Software Engineering Department
Butler University, Indianapolis IN, USA
cbosulli@butler.edu

Jonathan P. Sorenson
Computer Science and Software Engineering Department
Butler University, Indianapolis IN, USA
jsorenso@butler.edu

Aryn Stahl
Physics and Astronomy Department
Computer Science and Software Engineering Department
Butler University, Indianapolis IN, USA
anstahl@butler.edu

April 26, 2022

Abstract

We present and analyze an algorithm to enumerate all integers $n \leq x$ that can be written as the sum of consecutive $k$th powers of primes, for $k > 1$. We show that the number of such integers $n$ is asymptotically bounded by a constant times

$$c_k \frac{x^{2/(k+1)}}{(\log x)^{2k/(k+1)}},$$

where $c_k$ is a constant depending solely on $k$, roughly $k^2$ in magnitude. This also bounds the asymptotic running time of our algorithm.

We also present some computational results, using our algorithm, that imply this bound is, at worst, off by a constant factor near 0.6.

Our work extends the previous work by Tongsomporn, Wananiyakul, and Steuding [6] who examined consecutive sums of squares of primes.
1 Introduction

Let $S_k(x)$ denote the set of integers $n \leq x$ that can be written as a sum of the $k$th powers of consecutive primes. For example, $5^3 + 7^3 + 11^3 = 1799$ is an element of $S_3(2000)$. Let $S_k(x)$ be $\#S_k(x)$.

In this paper,

- We describe an algorithm that, given $k$ and $x$, produces the elements of $S_k(x)$ along with their representation. Its running time is linear in such representations; in practice this is linear in $S_k(x)$. The algorithm uses $O(kx^{1/k})$ space.

- In §3 we show that $S_k(x) \ll c_k x^{2/(k+1)} \log x^{2k/(k+1)}$, where $c_k = (k^2/(k - 1)) \cdot (k + 1)^{1-1/k}$. This is a generalization of a bound for $S_2(x)$ proven by [6]. Their bound is explicit and ours is not. This is also an upper bound on the number of arithmetic operations used by our algorithm.

- In §4 we apply our new algorithm to find exact values of $S_k(x)$ for various $x$ and $k$, and give some examples of integers that can be written as sums of consecutive powers of primes in more than one way. Note that $S_2(5000)$ was computed by [6]; see also sequence A340771 at the On-Line Encyclopedia of Integer Sequences (OEIS.org) [1].

We begin by describing our algorithm in the next section.

2 The Algorithm

Given as input a bound $x$ and integer exponent $k > 1$, our algorithm produces the elements of the set $S_k(x)$ as follows.

Let $p_1 = 2, p_2 = 3, \cdots, p_\ell$ denote the primes, and let $\pi(y)$ denote the number of primes $\leq y$. By the prime number theorem (see, for example, [4]), $\pi(y) \sim y / \log y$, and thus $p_n \sim n \log n$.

We assume all arithmetic operations take constant time. In practice, all our integers are at most 128 bits, or roughly 38 decimal digits.

1. Find the primes up to $x^{1/k}$.

   This step is not the bottleneck, so the Sieve of Eratosthenes is sufficient, taking $O(x^{1/k} \log \log x)$ time. See also [2, 5].
2. Compute the prefix sum array \( f[] \), where \( f[0] = 0 \) and \( f[i] := p_1^k + p_2^k + \cdots + p_i^k \) for all \( i \leq \pi(x^{1/k}) \), so that \( f[i+1] = f[i] + p_{i+1}^k \).

Note that the value of the largest entry in the array is bounded by \( x^{1+1/k} \).

Using a binary algorithm for integer exponentiation, this takes time \( O(\pi(x^{1/k}) \log k) \), which is smaller than the asymptotic bound given for Step 1. Storing \( f[] \) uses \( O(kx^{1/k}) \) bits of space.

3. Loops to enumerate \( S_k(x) \):

\[
\text{for } b := 0 \text{ to } \pi(x^{1/k}) - 1 \text{ do:}
\]

\[
\text{for } t := b + 1 \text{ to } \pi(x^{1/k}) \text{ do:}
\]

\[
\text{n := } f[t] - f[b];
\]

\[
\text{if } n > x \text{ break the } t \text{ loop,}
\]

\[
\text{else output}(n, p_{b+1});
\]

The time this step takes is proportional to the number of \((n, p_{b+1})\) pairs that are output. This, in turn, we bound in Theorem 3.1 below, at \( O(c_k x^{2/(k+1)} / (\log x)^{2k/(k+1)}) \) time.

We output pairs \((n, p_{b+1})\) in case a specific value of \( n \) gets repeated. If we have repeats for \( n \), the \( p_{b+1} \) values will be different, and \( p_{b+1} \) is the first prime in the sequence of powers of primes to generate \( n \), allowing us to quickly construct two (or more) representations of \( n \) as \( k \)th powers of consecutive primes.

In practice, we found repeated values of \( n \) to be quite rare.

**Example**

Let us compute \( S_3(1000) \) for an example.

1. We find the primes up to \( 1000^{1/3} = 10 \), so 2, 3, 5, 7.

2. We compute the prefix array \( f[] \) as follows:

|   | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 0 | 8 | 35 | 160 | 503 |

3. We generate the \( f[t] - f[b] \) values, and hence \( S_3(1000) \), as follows:

\[
b = 0 : (8, 2), (35, 2), (160, 2), (503, 2)
\]

\[
b = 1 : (27, 3), (152, 3), (495, 3)
\]

\[
b = 2 : (125, 5), (468, 5)
\]

\[
b = 3 : (343, 7)
\]
3 Analysis

In this section we prove the following theorem, which provides an upper bound on \( S_k(x) \). Because of the nature of the proof, it also gives a bound on the running time of our Algorithm in \( \S 2 \) as it counts the number of integer pairs \((n, p_b+1)\) output by the algorithm. Because pairs with the same \( n \) value are not corrected for, the bound below is a potential overcounting.

**Theorem 3.1** For \( k > 1 \) we have

\[
S_k(x) \ll c_k \frac{x^{2/(k+1)}}{(\log x)^{2k/(k+1)}},
\]

where \( c_k = \frac{k^2}{(k-1)} \cdot (k+1)^{1-1/k} \).

Note that \( c_k \sim k^2 \) for large \( k \).

In [6] they prove the explicit bound

\[
S_2(x) \leq 28.4201 \frac{x^{2/3}}{(\log x)^{4/3}}.
\]

We also have the trivial lower bound \( S_k(x) \geq \pi(x^{1/k}) \sim kx^{1/k}/\log x \) by the prime number theorem.

Our proof follows the same lines as in [6]. We begin by partitioning the members of \( S_k(x) \) by the number of prime powers \( m \) in their representative sum. Define

\[
S_{k,m}(x) = \# \{ n \leq x : \exists \ell \geq 0 : n = p_{\ell+1}^k + \cdots + p_{\ell+m}^k \}
\]

so that \( S_k(x) \leq \sum_{m=1}^M S_{k,m}(x) \) for a sufficiently large, and as yet unknown value \( M = M(x,k) \). We might not have equality here, as a specific integer \( n \) may have more than one representation as a sum of \( k \)th powers of consecutive primes, and so therefore be counted in more than one \( S_{k,m}(x) \) term. In practice, such integers are very rare; we elaborate on this in \( \S 4 \).

**Lemma 3.2 ([6])**

\[
S_{k,m}(x) \leq \pi((x/m)^{1/k}).
\]

**Proof:** Let \( n = S_{k,m}(x) \). We have

\[
m p_n^k \leq p_n^k + p_{n+1}^k + \cdots + p_{n+(m-1)}^k \leq x.
\]

Thus \( m p_n^k \leq x \), or \( p_n \leq (x/m)^{1/k} \), or \( S_{k,m}(x) = n \leq \pi((x/m)^{1/k}) \). \( \square \)

Next, we need an estimate for \( M \).
Lemma 3.3

\[ M(x, k) \sim (k + 1) \frac{x^{1/(k+1)}}{(\log x)^{k/(k+1)}} \]

Proof: We have

\[ \sum_{n=1}^{M} p_n^k \leq x < \sum_{n=1}^{M+1} p_n^k. \]

Using the asymptotic estimate \( p_M \sim M \log M \) from the prime number theorem and using the methods from [3, §2.7] we have

\[
\begin{align*}
\sum_{n=1}^{M} p_n^k & \sim \sum_{p \leq M \log M} p^k = \int_2^{\log M} t^k d\pi(t) \\
& \sim \int_2^{\log M} \frac{t^k}{\log t} dt \\
& \sim \frac{1}{\log M} \int_2^{\log M} t^k dt,
\end{align*}
\]

and so we have

\[ x \sim \frac{(M \log M)^{k+1}}{(k+1) \log M}. \]

Taking the log of both sides gives us \( (k + 1) \log M \sim \log x \). We then obtain that

\[ M \sim (k + 1)(x \log x)^{1/(k+1)}/\log x. \]

We are now ready to prove Theorem 3.1.

Proof: We have

\[
S_k(x) \leq \sum_{m=1}^{M} S_{k,m}(x) \leq \sum_{m=1}^{M} \pi((x/m)^{1/k}).
\]

By the prime number theorem and our lemmas, we have

\[
S_k(x) \leq \sum_{m=1}^{M} \pi((x/m)^{1/k}) \sim \sum_{m=1}^{M} \frac{k(x/m)^{1/k}}{\log(x/m)} \\
\sim \frac{k}{\log x} \sum_{m=1}^{M} m^{-1/k} \sim \frac{k}{\log x} \frac{M^{1-1/k}}{1-1/k}
\]
Plugging in our estimate above for $M$ gives this bound for $S_k(x)$:

$$kx^{1/k} \frac{k}{\log x} \frac{k}{k-1} \left( \frac{x^{1/(k+1)}}{(\log x)^{k/(k+1)}} \right)^{1-1/k}.$$ 

A bit of algebra simplifies the exponents to complete the proof.

Note that $\lim_{k \to \infty} (1/k) \cdot (k + 1)^{1-1/k} = 1$. \hfill □

4 Empirical Results

In this section we give some of our empirical results. This is not everything we have—the interested reader is encouraged to contact the second author for copies of the data or source code.

4.1 Tightness of Theorem 3.1

Although the proof for the upper bound given in Theorem 3.1 is straightforward, it does seem to give a very good asymptotic approximation for $S_k(x)$. Here we present values of $S_k(x)$ for $k = 2, 3, 5, 10, 20$ for $x$ up to $10^{38}$, which is close to the limit for 128-bit hardware integer arithmetic.

| $x$   | $S_2(x)^*$ | $c_2x^{2/3}/(\log x)^{4/3}$ | Ratio |
|-------|------------|-----------------------------|-------|
| $10^4$ | 37         | 91                          | 0.41  |
| $10^4$ | 132        | 288                         | 0.46  |
| $10^5$ | 519        | 994                         | 0.52  |
| $10^6$ | 1998       | 3619                        | 0.55  |
| $10^7$ | 7840       | 13680                       | 0.57  |
| $10^8$ | 31372      | 53142                       | 0.59  |
| $10^9$ | 126689     | 210816                      | 0.60  |
| $10^{10}$ | 517191   | 850276                      | 0.61  |
| $10^{11}$ | 2132474  | 3475655                     | 0.61  |
| $10^{12}$ | 8867094  | 14365431                    | 0.62  |
| $10^{13}$ | 37153225 | 59928838                    | 0.62  |
| $10^{14}$ | 156713533 | 251993659                   | 0.62  |
| $x$  | $S_3(x)^*$ | $c_3 x^{1/2}/(\log x)^{3/2}$ | Ratio |
|------|------------|-------------------------------|-------|
| $10^4$ | 29 | 64 | 0.45 |
| $10^6$ | 186 | 350 | 0.53 |
| $10^8$ | 1297 | 2276 | 0.57 |
| $10^{10}$ | 9568 | 16291 | 0.59 |
| $10^{12}$ | 73575 | 123930 | 0.59 |
| $10^{14}$ | 584184 | 983460 | 0.59 |
| $10^{16}$ | 4769563 | 8049501 | 0.59 |
| $10^{18}$ | 39796129 | 67459048 | 0.59 |
| $10^{20}$ | 338386013 | 575975457 | 0.59 |

| $x$  | $S_5(x)^*$ | $c_5 x^{1/3}/(\log x)^{5/3}$ | Ratio |
|------|------------|-------------------------------|-------|
| $10^4$ | 10 | 29 | 0.34 |
| $10^5$ | 38 | 78 | 0.48 |
| $10^6$ | 127 | 239 | 0.53 |
| $10^8$ | 479 | 796 | 0.60 |
| $10^{10}$ | 1639 | 2799 | 0.59 |
| $10^{12}$ | 6053 | 10237 | 0.59 |
| $10^{14}$ | 22938 | 38570 | 0.59 |
| $10^{16}$ | 87959 | 148735 | 0.59 |
| $10^{18}$ | 343199 | 584305 | 0.59 |
| $10^{20}$ | 1359330 | 2330551 | 0.58 |
| $10^{22}$ | 5451410 | 9413947 | 0.58 |
| $10^{24}$ | 22107170 | 38435468 | 0.58 |
| $10^{26}$ | 90459929 | 158370804 | 0.57 |
| $10^{28}$ | 373421607 | 657762277 | 0.57 |

| $x$  | $S_{10}(x)^*$ | $c_{10} x^{2/11}/(\log x)^{20/11}$ | Ratio |
|------|---------------|-------------------------------------|-------|
| $10^{12}$ | 21 | 44 | 0.47 |
| $10^{14}$ | 56 | 104 | 0.54 |
| $10^{16}$ | 154 | 262 | 0.59 |
| $10^{18}$ | 439 | 696 | 0.63 |
| $10^{20}$ | 1187 | 1917 | 0.62 |
| $10^{22}$ | 3304 | 5434 | 0.61 |
| $10^{24}$ | 9744 | 15756 | 0.62 |
| $10^{26}$ | 28290 | 46520 | 0.61 |
| $10^{28}$ | 84393 | 139440 | 0.61 |

| $x$  | $S_{20}(x)^*$ | $c_{20} x^{2/21}/(\log x)^{40/21}$ | Ratio |
|------|---------------|-------------------------------------|-------|
| $10^{22}$ | 15 | 27 | 0.55 |
| $10^{24}$ | 21 | 40 | 0.51 |
| $10^{26}$ | 36 | 63 | 0.57 |
| $10^{28}$ | 66 | 99 | 0.66 |
| $10^{30}$ | 105 | 160 | 0.65 |
| $10^{32}$ | 171 | 262 | 0.65 |
| $10^{34}$ | 232 | 367 | 0.63 |
Note that in all the tables above, the $S_k(x)$ column includes duplicates, so that it is a potential overcounting. Duplicates are very rare, which we will see in a moment, so including them does not much affect the accuracy of the estimates from Theorem 3.1.

4.2 Duplicates

We found 40 values of $n \leq x = 10^{12}$ that have multiple representations as sums of consecutive squares of primes. The smallest such number is 14720439, which can be written as

$$941^2 + 947^2 + 953^2 + 967^2 + 971^2 + 977^2 + 983^2 + 991^2 + 997^2 + 1009^2 + 1013^2 + 1019^2 + 1021^2 + 1031^2 + 1033^2$$

and as

$$131^2 + 137^2 + 139^2 + 149^2 + 151^2 + 157^2 + 163^2 + 167^2 + 173^2 + 179^2 + 181^2 + 191^2 + 193^2 + 197^2 + 199^2 + 211^2 + 223^2 + 227^2 + 229^2 + 233^2 + 239^2 + 241^2 + 251^2 + 257^2 + 263^2 + 269^2 + 271^2 + 277^2 + 281^2 + 283^2 + 293^2 + 307^2 + 311^2 + 313^2 + 317^2 + 331^2 + 337^2 + 347^2 + 349^2 + 353^2 + 359^2 + 367^2 + 373^2 + 379^2 + 383^2 + 389^2 + 397^2 + 401^2 + 409^2 + 419^2 + 421^2 + 431^2 + 433^2 + 439^2 + 443^2 + 449^2 + 457^2 + 461^2 + 463^2 + 467^2 + 479^2 + 487^2 + 491^2 + 499^2 + 503^2 + 509^2 + 521^2 + 523^2 + 541^2 + 547^2 + 557^2 + 563^2 + 569^2 + 571^2 + 577^2 + 587^2 + 593^2 + 599^2 + 601^2 + 607^2 + 613^2 + 617^2 + 619^2 + 631^2 + 641^2 + 643^2 + 647^2.$$

To find these, we sorted the output of our algorithm from §2, and then used the `uniq -D` unix/linux command to suss out the duplicates.

We found no integers that can be written as the sum of consecutive powers of primes in more than one way for any power larger than 2. We searched for cubes up to $10^{18}$, fifth powers up to $10^{27}$, and tenth and twentieth powers up to $10^{38}$. This search requires computing $S_k(x)$ and not just $S_k(x)$; note that it is much faster to compute just $S_k(x)$ in practice because outputting the elements of $S_k(x)$ to a text file slows down the computation considerably.

We found exactly one example with differing powers:

$$23939 = 23^2 + 29^2 + 31^2 + 37^2 + 41^2 + 43^2 + 47^2 + 53^2 + 59^2 + 61^2 + 67^2 = 17^3 + 19^3 + 23^3.$$

We conclude this subsection with the list of 40 integers $\leq 10^{12}$ that can be written as sums of squares of consecutive primes in two ways. For each
such integer in the table below, we list the starting primes for each of their two ways to sum.

| n         | Prime 1 | Prime 2 |
|-----------|---------|---------|
| 14720439  | 131     | 941     |
| 16535628  | 1123    | 569     |
| 34714710  | 2389    | 401     |
| 40741208  | 131     | 653     |
| 61436388  | 569     | 809     |
| 603346308 | 401     | 919     |
| 1172360113| 3701    | 4673    |
| 1368156941| 1367    | 16519   |
| 1574100889| 3623    | 613     |
| 1924496102| 11657   | 2803    |
| 1989253499| 3359    | 613     |
| 2021860243| 3701    | 4297    |
| 6774546339| 11273   | 47513   |
| 9770541610| 1663    | 7243    |
| 1223085963| 10177   | 2777    |
| 12311606487| 28603   | 3257    |
| 12540842446| 11087   | 479     |
| 14513723777| 1663    | 6323    |
| 26423329489| 1709    | 32401   |
| 38648724198| 2777    | 6967    |

| n         | Prime 1 | Prime 2 |
|-----------|---------|---------|
| 47638558043| 28097   | 65731   |
| 50195886916| 479     | 6857    |
| 50811319031| 2039    | 21283   |
| 56449248367| 2803    | 4127    |
| 86659250142| 4561    | 53609   |
| 105146546059| 29587   | 6599    |
| 119789313426| 31847   | 42299   |
| 125958414196| 16763   | 26183   |
| 134051910100| 183047  | 4397    |
| 159625748030| 1367    | 3301    |
| 169046403821| 183829  | 19717   |
| 263787548443| 47297   | 62347   |
| 330881994258| 11161   | 2039    |
| 438882621700| 16763   | 20369   |
| 507397259055| 643     | 75013   |
| 572522061248| 18427   | 44371   |
| 687481319598| 16139   | 338461  |
| 780455791261| 3257    | 7057    |
| 847632329089| 184003  | 7523    |
| 854350226239| 14821   | 6599    |
4.3 Initial Elements of $S_k$

We wrap up the presentation of our computations with the first few elements of each of the $S_k$ sets we computed.

$S_2$:  
4 9 13 25 34 38 49 74 83 87 121 169 170 195 204 208 289 290 339 361

$S_3$:  
8 27 35 125 152 160 343 468 495 503 1331 1674 1799 1826 1834 2197 3871 3996 4023

$S_5$:  
32 243 275 3125 3368 3400 16807 19932 20175 20207 1331 1674 1799 1826 1834 2197 3871 3996 4023

$S_{10}$:  
1024 59049 60073 9765625 9824674 9825698 282475249 292240874 292299923 292300947

$S_{20}$:  
1048576 3486784401 3487832977 95367431640625 95370918425026 95370919473602 79792266297612001 79887633729252626 79887637216037027 79887637217085603

5 Future Work

We have several ideas for future work:

- Our primary goal is to parallelize our algorithm from §2 to extend our computations. For larger powers, this will also mean using multiple-precision integer arithmetic using, for example, GMP.

- A more careful proof of Theorem 3.1 might give explicit upper bounds, or perhaps an asymptotic constant. If such a constant exists, it appears to be near 0.6.

- Is there a power $k > 2$ for which there are integers with multiple representations as sums of powers of consecutive primes? We have not found any as of yet.

Acknowledgements

The authors are grateful to Frank Levinson for his support of computing research infrastructure at Butler University.
References

[1] The On-Line Encyclopedia of Integer Sequences. https://oeis.org.

[2] A. O. L. Atkin and D. J. Bernstein. Prime sieves using binary quadratic forms. Mathematics of Computation, 73:1023–1030, 2004.

[3] Eric Bach and Jeffrey O. Shallit. Algorithmic Number Theory, volume 1. MIT Press, 1996.

[4] G. H. Hardy and E. M. Wright. An Introduction to the Theory of Numbers. Oxford University Press, 5th edition, 1979.

[5] Jonathan P. Sorenson. Two compact incremental prime sieves. LMS Journal of Computation and Mathematics, 18:675–683, 2015.

[6] Janyarak Tongsomporn, Saeree Wananiyakul, and Jörn Steuding. Sums of consecutive prime squares. Integers, 22, 2022. A9.