INTEGERS: ELECTRONIC JOURNAL OF COMBINATORIAL NUMBER THEORY 8 (2008), #A61

RESOLUTION OF THE MIXED SIERPIŃSKI PROBLEM

Louis Helm lhelm@seventeenorbust.com

Phil Moore moorep@lanecc.edu

Payam Samidoost payam.samidoost@gmail.com

George Woltman woltman@alum.mit.edu

Received: 2/12/08, Revised: 10/8/08, Accepted: 12/3/08, Published: 12/23/08

Abstract

Recent progress on the Sierpiński problem has resulted in the proof of the following theorem: 78557 is the smallest positive odd integer k such that both $k \cdot 2^n + 1$ and $k + 2^n$ are composite for any positive integer n. An algorithmic enhancement to the fast Fourier transform routines used in this research is described. Prospects for the eventual resolution of both the original and the dual Sierpiński problems are estimated.

1. The Sierpiński Problem

In 1960, Waclaw Sierpiński proved that there are an infinite number of odd positive integers k such that $k \cdot 2^n + 1$ is composite for any integer $n \geq 1$, and posed the problem of finding the smallest such k [22]. Such a k value is now known as a Sierpiński number. In 1962, John Selfridge proved that $k = 78557$ possesses this property (unpublished), a result which follows from the observation that for any value of n, at least one of the seven primes in the covering set $\{3, 5, 7, 13, 19, 37, 73\}$ must divide $78557 \cdot 2^n + 1$ depending upon the value of n modulo 36. No smaller Sierpiński numbers have been discovered, and it is now widely believed that 78557 is indeed the smallest number [11]. Attempts to verify this conjecture by systematic search for a prime value of $k \cdot 2^n + 1$ for each $k < 78557$ were carried out by Selfridge, Baillie, Cormack, Williams, Jaeschke, Keller, Buell, and Young [1, 2, 5, 6, 14, 15, 16, 21]. In 1997, Keller and Ballinger organized a distributed search on the remaining candidates using software written by Gallot, and by 2002, the status of only 17 k values less than 78557 was still unknown [3]. In that year, the distributed computing project Seventeen or Bust was started by Helm and Norris to work on the remaining candidates. To date, this project has eliminated 11 more potential Sierpiński number candidates with the discovery of 11 large primes, leaving the following 6 unresolved k values [12]:

10223	21181	22699
24737	55459	67607.

Table 1: Remaining k for Sierpiński Problem

2. Dual Sierpiński Problem

Replacing the exponent n in the above expression with a negative integer leads to the expression $k \cdot 2^{-n} + 1 = (k + 2^n)/2^n$. It is easily verified that for $k = 78557$, the numerator is divisible by at least one of the seven primes in the same covering set $\{3, 5, 7, 13, 19, 37, 73\}$ again depending upon the residue class of $n \bmod 36$. This leads to the consideration of the dual Sierpiński problem: whether or not $k = 78557$ is the smallest value of k with the property that $k + 2^n$ is always composite. Covering set arguments had been used previously by Erdős [8] and Schinzel, and Sierpiński included in his paper the proof of Schinzel that a finite covering set of the form $k + 2^n$ must also be a covering set of the form $k \cdot 2^n + 1$. The converse implication is also easily proven. Sierpiński also noted that Schinzel had proven that there are an infinite number of values of k with the property that $k + 2^n$ is always composite, similar to the analogous property for Sierpiński numbers. Despite this similarity, the solution of one problem does not necessarily imply the solution of the other. Selfridge noted that a covering set of one sequence could contain a prime element of the other sequence, as noted in the corrigendum to Sierpiński's paper [22]. An even more fundamental problem is that we are not able to rule out the possible existence of infinite covering sets. Guy reports that Erdős had conjectured that all Sierpiński numbers must have finite covering sets [11], but Izotov in 1995 constructed a family of Sierpiński numbers which appear likely to have infinite covering sets. He used the identity $4y^4 + 1 = (2y^2 + 2y + 1)(2y^2 - 2y + 1)$ and chose $k = x^4$ so that $k \cdot 2^n + 1$ is composite by the above identity for $n \equiv 2 \bmod 4$. Then x was chosen so that $k \cdot 2^n + 1$ is covered by the set $\{3, 17, 257, 65537, 641, 6700417\}$ for n not congruent to 2 mod 4. If we also take x divisible by 5 so that 5 does not cover the case $n \equiv 2 \bmod 4$, then there do not appear to be any obvious small finite covering sets. Filaseta, Finch, and Kozek [9] give further examples for this type of covering, including the smallest example with the covering set $\{3, 17, 97, 241, 257, 673\}$, and produce numerical evidence that no small finite covering sets exist. They also prove a conjecture of Chen that for each positive integer r, there are infinitely many Sierpinski numbers that are perfect rth powers, and they conjecture that if k is not a perfect power, then k can only be a Sierpiński number when a finite covering set exists. None of the values currently under investigation in either of these problems is a perfect power.

The dual problem initially appears to be somewhat easier than the original problem, as more values of k are eliminated at small values of n. For example, $31 = 29 + 2^1 = 23 + 2^3$ eliminates two values of k from the dual problem, whereas no two k values can be eliminated from the original problem by the same prime. A few values of $k < 78557$, however, still need searches to large values of n to uncover primes. A complete solution of the dual problem appears to be quite difficult, not only because of the increasing difficulty of searching for larger n, but also because a large probable prime of the form $k + 2^n$ is not easily proven to be a genuine prime. On the other hand though, the elimination of a k value from the dual problem with a large prime or probable prime value can at least be looked upon as strong circumstantial evidence that such a k value is probably not a Sierpiński number. In 1983 Jaeschke searched for primes of the form $k+2^n$ for all unresolved $k < 78557$ from the original problem and all $n \leq 100$, thereby generating a list of exactly sixteen values of k for which no prime of either form was known [14]. This problem of identifying the smallest value of k for which both $k \cdot 2^n + 1$ and $k + 2^n$ are always composite has become known as the mixed Sierpiński problem.

In 2001 and 2002, Samidoost encouraged investigation of the dual Sierpiński problem, coordinating his own findings with those of other investigators [18]. As a result, a large number of primes of the form $k + 2^n$ were discovered [20]. Although more k values are eliminated at small values of n than in the original problem, the dual investigation suffers from the difficulty noted above that the larger discoveries are only known to be probable primes. Whereas large numbers of the form $k \cdot 2^n + 1$ can easily be proven prime using Proth's criterion, there is currently no known practical method of proving primality for large numbers of the dual form. Samidoost's August 2002 discovery of $19249 + 2^{551542}$ with 166,031 decimal digits was in fact the largest known probable prime at that time. As a result, Samidoost identified $k = 28433$ as the only value of $k < 78557$ for which neither a prime of the form $k \cdot 2^n + 1$ nor a prime or probable prime of the form $k + 2^n$ had been discovered, and therefore the only remaining unknown case in the mixed Sierpiński conjecture if one allows probable primes in place of proven primes.

3. Recent Results Establishing the Mixed Sierpiński Theorem

The Seventeen or Bust project has performed over one million probable primality tests over the last 6 years. Many candidates have also been eliminated from primality testing by a distributed sieving effort using a variety of sieving programs. The prime $28433 \cdot 2^{7830451} + 1$ was discovered on December 30, 2004, essentially establishing the mixed Sierpiński conjecture in the weaker sense of allowing probable primes. Recently, this result was strengthened by the discovery of the prime $19249 \cdot 2^{13018586} + 1$ by Seventeen or Bust member Konstantin Agafonov. At 3,918,990 digits, this number is currently the largest known non-Mersenne prime. For the remaining six unresolved k values from the original Sierpiński problem, the following dual primes are known:

$10223 + 2^{19}$	$21181 + 2^{28}$	$22699 + 2^{26}$
$24737 + 2^{17}$	$55459 + 2^{14}$	$67607 + 2^{16389}$.

Table 2: Known Dual Primes

Of the five smallest numbers in this list, none exceed 9 digits and all are easily verified to be genuine primes by trial division. The largest number, discovered in 2002 by Fougeron, has 4934 decimal digits and was recently verified to be prime using the Elliptic Curve Primality Proof (ECPP) program Primo.exe by Martin [17]. The run on a 3000 MHz Pentium IV processor required 26 days. The primality certificate was then verified by Fougeron's program Cert_Val after 27 hours of computation. As a result, the mixed Sierpiński conjecture is now a theorem.

4. FFT Multiplication Optimization

Since this new theoretical result was made possible by the computational progress of Seventeen or Bust [12], one might reasonably ask how the project was able to test so many large prime candidates so quickly. Beyond having over 10,000 volunteers in SB's distributed network, the rate of progress was also accelerated by previously unpublished optimizations in the software's FFT multiplication routines. Seventeen or Bust runs a probable prime test with Irrational Base Discrete Weighted Transform (IBDWT) multiplication code written by Woltman [23] and based on work of Crandall and Fagin [7]. This code is also used in the pfgw.exe client for the dual Sierpiǹski search. The multiplication routines contain highly optimized assembly code and take advantage of cache structure to minimize memory access bottlenecks. Woltman also expanded on the recent work of Percival [19] in adapting his routines to efficiently do computations modulo numbers of the form $k \cdot 2^n \pm c$, thus making his routines of particular value to Seventeen or Bust as well as the dual search. Treatment of the $\pm c$ term essentially follows the scheme of Percival, but the $k \cdot 2^n$ term is treated in a way that allows use of smaller FFT sizes for many exponents than those required by Percival's method. This new scheme involves performing a Mersenne-like DWT on $2^{n+\log k} \pm c$ with weights ranging from 1 to 2 rather than using Percival's weights for k in the FFT word which range from 1 to k, saving approximately $\log_2 k$ bits of weighting data, which become $2\log_2 k$ bits in the inverse FFT word after point-wise squaring. The complication of this scheme is that the "wrap around" data is now divided by k, which is dealt with by multiplying each result word by k before rounding to an integer. Carry out of the top word is done very carefully, but is not a major problem. The procedure creates a result that has been multiplied by k, which is circumvented by dealing with numbers in $x/k \bmod (k \cdot 2^n \pm c)$ format. Therefore, to square x, the transform of x/k is computed and squared component-wise, after which the inverse transform is taken, giving a result x^2/k still in the same format, easily adapted to a long series of squarings such as is performed in a probable prime or Proth test. At the conclusion of the computation, a final multiplication by k converts the result back into the desired form. The net savings in each inverse FFT word amounts to the equivalent of $(\log_2 k)/2$ bits in the original input FFT word, allowing in many cases for the use of smaller FFT sizes for the k values under investigation.

5. Prospects for Resolution of the Sierpiński and Dual Problems

The dual Sierpiński problem is presently concerned with the following five k values for which not even a probable prime is known:

2131	28433	40291
41693	75353.	

Table 3: Remaining k for Dual Sierpiński

Search limits for n on these numbers are currently (Oct. 2008) at 1,400,000, and a distributed search is being coordinated at www.mersenneforum.org. We have recently certified all of the least dual primes for each k value which are smaller than $67607 + 2^{16389}$ to be prime, leaving just 28 k values less than 78557 for which only probable primes are known. The probable primes are presented in the following table:

$77899 + 2^{21954}$	$63691 + 2^{22464}$	$62029 + 2^{24910}$	$22193 + 2^{25563}$
$57083 + 2^{26795}$	$77783 + 2^{26827}$	$34429 + 2^{28978}$	$20273 + 2^{29727}$
$29333 + 2^{31483}$	$19081 + 2^{31544}$	$4471 + 2^{33548}$	$47269 + 2^{38090}$
$26213 + 2^{56363}$	$39079 + 2^{56366}$	$21661 + 2^{61792}$	$10711 + 2^{73360}$
$14717 + 2^{73845}$	$17659 + 2^{103766}$	$7013 + 2^{104095}$	$48527 + 2^{105789}$
$35461 + 2^{139964}$	$60443 + 2^{148227}$	$60947 + 2^{176177}$	$64133 + 2^{304015}$
$37967 + 2^{308809}$	$19249 + 2^{551542}$	$60451 + 2^{983620}$	$8543 + 2^{1191375}$

Table 4: Known Probable Primes

The two largest and the fourth largest probable primes in the table were discovered by the authors in a recent extension of the search, and the largest is now the largest known probable prime. We have also performed strong probable prime tests on each of these 28 candidates to at least 10 different bases. F. Morain has recently proven a 20562 digit number to be prime via ECPP on a distributed network, so an extensive effort would probably suffice to certify the smallest 17 of these 28 numbers as definitely prime. The 11 larger probable primes will most likely have to wait for new breakthroughs in primality testing, to say nothing of the 5 undiscovered probable primes corresponding to the k values in Table 3 presumably waiting to be found.

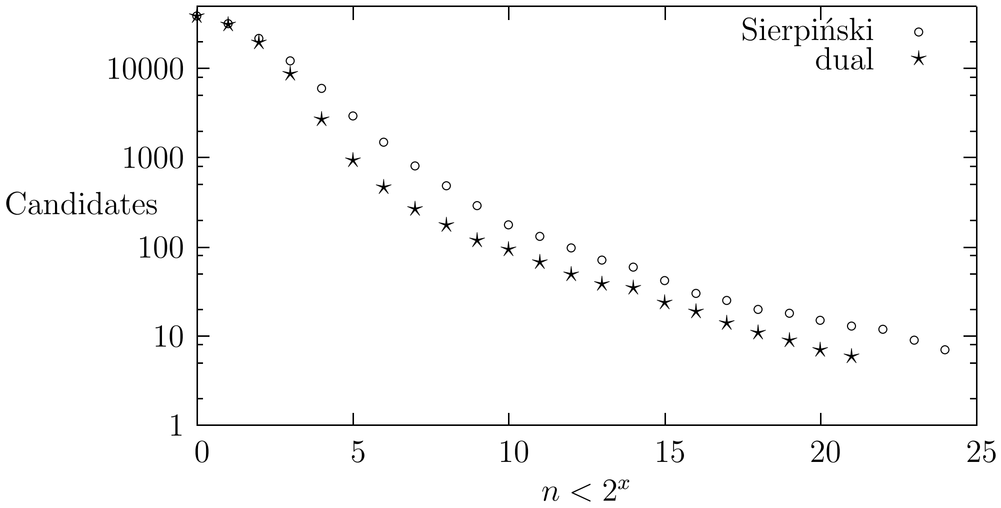

Figure 1: Remaining Sierpiński/Dual Candidates as n Increases

The accompanying figure shows the number of remaining candidates at each power of two for both the original and dual problems, illustrating the relative advantage of the dual problem in eliminating more candidates at low values of n. Interestingly, the dual problem seems to possess a strong advantage for n near 2^6 where the dual problem possesses only 32% of the candidates of the original problem, but this advantage seems to diminish at higher n, where this ratio rises to 63% for $n = 2^{16}$. The investigators do not have an explanation for this phenomenon. As the current dual problem search limits are considerably lower than those for the original problem, it would be interesting to see what discoveries might result from even a relatively modest further search effort.

Although full covering sets are unknown and in fact considered unlikely, the remaining k values in both problems possess partial covering sets which cause most values of n to result in composite values of the respective sequences. This tendency can be quantified by the concept of Proth weight, a measure of the asymptotic expected density of primes of the form $k \cdot 2^n + 1$ as compared to random numbers of the same magnitude, normalized so that the average of the Proth weight over all odd values of k equals 1 [4, 18]. Because of the close correspondence between covering sets in the two problems, this measure also applies to the sequences of the dual form. As most members of these sequences are eliminated by small prime divisors, a maximum prime cutoff value can be used in computing the Proth weight. The data computed by Brennan's Proth weight applet agree with the data of Seventeen or Bust on the remaining primality candidates for each sequence after sieving. Typical k values have Proth weights on the order of unity, while Sierpiński numbers have Proth weight 0. The Proth weights of the remaining k values in the original Sierpiński problem range from .03540 ($k = 67607$) to .14102 ($k = 55459$), with similar ranges for the dual problem, quantifying the extent to which we would expect primes in these sequences to be rare. Gallot has used this idea to compute the probabilities of solving the Sierpiński problem given certain bounds on the exponent n [10]. (Note that his weights differ by a factor of 2 from those given by Brennan because of a different normalization.) We have applied Gallot's methods to our current knowledge of these two problems to compute the expected n values to which we would have to search in order to estimate a 10% chance, a 50% chance, and a 90% chance of solving each problem. Our results are presented in the following table:

Probability	**10%**	**50%**	**90%**
Sierpiński Problem	4.1×10^9	3.4×10^{12}	1.2×10^{19}
Dual Problem	1.0×10^8	1.1×10^{10}	7.2×10^{13}

Table 5: Search Limits on n for Given Probabilities of Solution

The probabilities for the Sierpiński problem indicate a modest but distinct improvement over earlier estimates, due primarily to the fortunate elimination of several low-weight k values. Although the dual problem is expected to be significantly easier, it is still expected to be a difficult problem. The wide range of predicted maximum n values with the varying percentages indicates that the solution of either of these problems with our current level of technology and resources will require some considerable luck!

Acknowledgments

The authors would like to thank Michael Garrison for his dedicated support administering the Seventeen or Bust server for the past several years. We also thank Jean Penné, Phil Carmody, Paul Jobling, Mikael Klasson, Yves Gallot, and Kerry Jones for software development that improved the ability of Seventeen or Bust to discover the primes central to this investigation. Thanks also to Jim Fougeron and Geoff Reynolds for software work on the dual project, and a special thanks to the thousands of enthusiastic volunteers who have contributed their computer time to the primality testing that ultimately made these results possible.

References

[1] R. Baillie *New primes of the form $k \cdot 2^n + 1$*, Math. Comp., 33(1979) 1333-1336; MR 80h:10009

[2] R. Baillie, G. Cormack, H. C. Williams *The problem of Sierpiński concerning $k \cdot 2^n + 1$*, Math. Comp., 37(1981) 229-231 MR 83a:10006a; corrigendum, 39(1982) 308

[3] R. Ballinger, W. Keller *The Sierpiński Problem: Definition and Status*, http://www.prothsearch.net/sierp.html

[4] J. Brennen *Proth weight applet*, http://www.brennen.net/primes/ProthWeight.html

[5] D. A. Buell, J. Young *Some large primes and the Sierpiński problem*, SRC Technical Report 88-004, Supercomputing Research Center, Lanham MD, May 1988

[6] G. Cormack, H. C. Williams *Some very large primes of the form* $k \cdot 2^n + 1$, Math. Comp., 35(1980) 1419-1421; MR 81i:10011; corrigendum, W. Keller, 38(1982) 335; MR 82k:10011

[7] R. Crandall, B. Fagin *Discrete weighted transforms and large integer arithmetic*, Math. Comp., 62(1994) 305-324; MR 94c:11123

[8] P. Erdős *On integers of the form* $2^k + p$ *and some related problems*, Summa Brasiliensis Mathematica, 2(1950) 113-123; MR 0044558

[9] M. Filaseta, C. Finch, M. Kozek *On powers associated with Sierpiński numbers, Riesel numbers and Polignac's conjecture*, J. Number Theory 128(2008) 1916-1940

[10] Y. Gallot *On the number of primes in a sequence*, 2001, http://perso.wanadoo.fr/yves.gallot/papers/weight.pdf

[11] R. Guy *Unsolved Problems in Number Theory, 3rd ed.*, Springer, B21:119-121, F13:383-385 (2005) MR 96e:11002

[12] L. Helm, D. Norris *Seventeen or Bust: A Distributed Attack on the Sierpiński Problem*, http://www.seventeenorbust.com

[13] A. S. Izotov *A note on Sierpiński Numbers*, Fibonacci Quart. 33 (1995), 206-207; MR96f:11020

[14] G. Jaeschke *On the smallest* k *such that all* $k \cdot 2^n + 1$ *are composite*, Math. Comp., 40(1983) 381-384; MR 84k:10006; corrigendum, 45(1985) 637; MR 87b:11009

[15] W. Keller *Factors of Fermat numbers and Large Primes of the Form* $k \cdot 2^n + 1$, Math. Comput., 41(1983) 661-673; MR 85b:11117

[16] W. Keller *Factors of Fermat numbers and Large Primes of the Form* $k \cdot 2^n + 1$, II, preprint 1992

[17] M. Martin *Ellipsa > Primo*, http://www.ellipsa.net/primo/index.html

[18] *Primenumbers archive*, http://tech.groups.yahoo.com/group/primenumbers/

[19] C. Percival *Rapid multiplication modulo the sum and difference of highly composite numbers*, Math. Comp., 72(2003) 387-395; MR 2003i;11183

[20] P. Samidoost *The dual Sierpiński problem search*, http://sierpinski.insider.com/dual

[21] J. L. Selfridge *Solution of problem 4995*, Amer. Math Monthly, 70(1963) 101; MR 1532000

[22] W. Sierpiński *Sur un problèm concernant les nombres* $k \cdot 2^n + 1$, Elem. Math., 15(1960) 73-74; MR 22:7983; corrigendum, 17(1962) 85

[23] G. Woltman *The Great Internet Mersenne Prime Search*, IBDWT source code, http://www.mersenne.org/source.htm, (2004)