A proof of Sophie Germain primes conjecture

Marko V Jankovic

A proof of Sophie Germain primes conjecture

Marko V. Jankovic

ARTORG Centre for Biomedical Engineering Research,

University of Bern,

Murtenstrasse 50, 3008 Bern, Switzerland

Abstract

In this paper a proof of the existence of an infinite number of Sophie Germain primes, is going to be presented. In order to do that, we analyse the basic formula for prime numbers and decide when this formula would produce a Sophie Germain prime, and when not. Originally very difficult problem (in observational space) has been transformed into a simpler one (in generative space) that can be solved.

1 Introduction

A prime p is a Sophie Germain prime if $2p + 1$ is prime, too [1]. In that case the prime number $2p + 1$ is called safe prime. These special primes have applications in public key cryptography, pseudorandom number generation, and primality testing; see, for example,

[2, 4, 6]. Originally, they have been used also in the investigation of cases of Fermat's last theorem [3]. It has been conjectured that there exist infinitely many Sophie Germain primes, but this was unproven (see for instance, [5]).

In this paper it is going to be proved that exists an infinite number of Sophie Germain primes. The problem is addressed in generative space, which means that prime numbers are not going to be analysed directly, but rather their representatives, in the other space, that can be used to produce them.

Remark: Prime numbers 2 and 3 are in a sense special primes, since they do not share some of the common features of all other prime numbers. For instance, every prime number, apart from 2 and 3, can be expressed in the form $6l + 1$ *or* $6l - 1$*, where* $l \in N$*. So, in this paper most of the time we analyse prime numbers bigger than 3. It has to be said that both 2 and 3 are Sophie Germain primes, but that has no impact on the conclusion of this paper.*

2 Proof

It is easy to check that any prime number (apart form 2 and 3) can be expressed in the form $6l+1$ or $6s-1$ $(l, s \in N)$. Having that in mind, it is easy to check that numbers in the form $6l + 1$, could never be Sophie Germain primes since the safe prime is in the form

$$2(6l + 1) + 1 = 12l + 3 = 3(4l + 1)$$

and that is composite number divisible by 3. So, the prime number that can potentially be Sophie Germain prime must be in the form $6s - 1$ and then the safe prime is going to be in

the form $6(2s) - 1$.

We denote any composite number (that is represented as a product of prime numbers bigger than 3) with $CPN5$. Also, we mark with mpl a number in the form $6l + 1$, and with mps a number in the form $6s - 1$ $(l, s \in N)$. In that case, it is easy to check that any composite number $CPN5$ can be expressed in the form $mpl \times mpl$, $mps \times mps$ or $mpl \times mps$. So, if we have a number in the form $6k - 1$ that is composite number it must hold

$$k = \frac{CPN5 + 1}{6}. \tag{2.1}$$

Since $CPN5$ should be in the mps form, $CPN5$ can be generally expressed as a product $mpl \times mps$, or

$$mpl = 6x + 1 \text{ and } mps = 6y - 1 (x, y \in N),$$

which leads to

$$CPN5 = mpl \times mps = 6(6xy - x + y) - 1, \tag{2.2}$$

or, due to symmetry

$$mpl = 6y + 1 \text{ and } mps = 6x - 1,$$

which leads to

$$CPN5 = mpl \times mps = 6(6xy + x - y) - 1. \tag{2.3}$$

So, if we replace (2.2, 2.3) in (2.1) we obtain forms of k that potentially cannot produce a Sophie Germain prime number. Those forms are expressed by the following equation

$$k = \begin{cases} (6x - 1)y + x \\ \\ (6x + 1)y - x \end{cases}, \tag{2.4}$$

where $x, y \in N$.

Also, we know that safe prime which is generated by $2k$ cannot be composite. If the safe number is composite the following equation must hold

$$k = \frac{CPN5 + 1}{6 \times 2}, \tag{2.5}$$

where $CPN5$ is composite number in the *mps* form. Using the same analysis as in the previous case, and replacing (2.2) and (2.3) in (2.5), we obtain additional cases in which k cannot be used to produce Sophie Germain prime pairs, and they are defined by the following equation

$$k = \begin{cases} (6x-1)y + \frac{x}{2}, \ x \text{ is even} \\ (6x+1)y - \frac{x}{2}, \ x \text{ is even} \\ (6x-1)y + 3x + \frac{x-1}{2}, \ x \text{ is odd} \\ (6x+1)y - 3x - \frac{x+1}{2}, \ x \text{ is odd} \end{cases}, \tag{2.6}$$

where $x, y \in N$. Equations (2.4) and (2.6) give a sufficient and necessary condition for k, so that it cannot be used for generation of the prime pairs in the form $(p, 2p+1)$. In order to prove that there are infinitely many prime pairs in the form $(p, 2p+1)$ we need to prove that exists infinitely many k that cannot be expressed in the form (2.4) or (2.6). First, we will check the form of (2.4, 2.6) for some values of x.

Case $x = 1$	Case $x = 2$	Case $x = 3$	Case $x = 4$
$k = 5y + 1$	$k = 11y + 2$	$k = 17y + 3$	$k = 23y + 4$
$k = 5y + 3$	$k = 11y + 1$	$k = 17y + 10$	$k = 23y + 2$
$k = 7y - 1$	$k = 13y - 2$	$k = 19y - 3$	$k = 5(5y - 1) + 1$
$k = 7y - 4$	$k = 13y - 1$	$k = 19y - 11$	$k = 5(5y - 1) + 3$
Case $x = 5$	Case $x = 6$	Case $x = 7$	Case $x = 8$
$k = 29y + 5$	$k = 7(5y + 1) - 1$	$k = 41y + 7$	$k = 47y + 8$
$k = 29y + 17$	$k = 5(7y) + 3$	$k = 41y + 24$	$k = 47y + 4$
$k = 31y - 5$	$k = 37y - 6$	$k = 43y - 7$	$k = 7(7y - 1) - 1$
$k = 31y - 18$	$k = 37y - 3$	$k = 43y - 25$	$k = 7(7y - 1) + 3$

From examples, we can see that if $(6x - 1)$ or $(6x + 1)$ represent a composite number, k that is represented by that number has also representation by one of the prime factors of that composite number. This can be easily proved in the general case, by direct calculation, using representations similar to (2.2, 2.3). Here only one case is going to be analysed. All other cases can be analysed analogously. In this case we assume

$$(6x - 1) = (6l + 1)(6s - 1),$$

where $(l, s \in N)$. From previous equation x can be expressed as

$$x = 6ls - l + s.$$

Having that in mind, and selecting one representation of k that includes form $(6x - 1)$, we

have

$$k = (6x-1)y - x = (6l+1)(6s-1)y - 6ls + l - s$$

or

$$k = (6l+1)(6s-1)y - s(6l+1) + l = (6l+1)((6s-1)y - s) + l,$$

which means

$$k = (6l+1)f + l$$

and that represents already existing form of the representation of k for factor $(6l+1)$, where

$$f = (6s-1)y - s.$$

It can be seen that all patterns for k that potentially result in composite number, include prime numbers. It is going to be checked is the number of k that cannot be represented by the models (2.4, 2.6), finite or infinite. In order to do it, a method similar to the sieve of Eratosthenes [7] is going to be used.

When all numbers that can be represented in forms

$$5y + 1 \text{ and } 5y + 3,$$

are removed from natural numbers, it can be seen that ratio $r_1 = 2/5$ of all natural numbers are removed. So, ratio $c_1 = 1 - 2/5 = 3/5$ of all natural numbers cannot be represented by those two patterns and they still contain some k that could be potentially used for representation of Sophie Germain primes. If, now, in addition, the natural numbers in the form

$$7y-1 \text{ and } 7y-4,$$

are removed, then the ratio of removed numbers can be calculated by the following equation (taking care that every removed number is calculated only once; basically, we apply the formula for calculation of the probability of occurring of two events that are not mutually exclusive $P(A \cup B) = P(A) + P(B) - P(A \cap B)$)

$$r_2 = r_1 + \frac{2}{7} - \frac{2}{7} \times r_1 = r_1 + \frac{2}{7}(1 - r_1) = \frac{2}{5} + \frac{2}{7}\left(1 - \frac{2}{5}\right) = \frac{20}{5 \times 7}.$$

The ratio of all natural numbers that, still, potentially can be used for "generation" of Sophie Germain primes, is given by the following equation

$$c_2 = 1 - r_2 = 1 - r_1 - \frac{2}{7}(1 - r_1) = \left(1 - \frac{2}{7}\right) \times c_1 = \frac{3 \times 5}{5 \times 7}.$$

Now, we denote prime numbers bigger than 3 as $p5$, where $p5(1) = 5$, $p5(2) = 7$ and so on.

Suppose that after step n we have ratio r_n of all numbers removed and ratio c_n of all numbers still potentially available for generation of Sophie Germain primes. In the step $n+1$ we have

$$r_{n+1} = r_n + \frac{2}{p5(n+1)} - \frac{2}{p5(n+1)} \times r_n.$$

After a few elementary calculations, the following equation is obtained

$$r_{n+1} = r_n + \frac{2}{p5(n+1)}(1 - r_n).$$

Now, the following equation holds

$$c_{n+1} = 1 - r_{n+1} = 1 - r_n - \frac{2}{p5(n+1)}(1 - r_n) = c_n - \frac{2}{p5(n+1)} \times c_n, \tag{2.7}$$

or

$$c_{n+1} = \left(1 - \frac{2}{p5(n+1)}\right) \times c_n. \tag{2.8}$$

That can be interpreted in the following way: in every step we remove $2/p5(n+1)$ numbers of what is left for potential representation of Sophie Germain primes.

Now, one additional sieve elimination process (SEP), denoted as R25, is considered. R25 is defined by the following rule - start with all natural numbers, and in every step remove $2/5$ of the numbers that is left. In this case, ratio of the numbers that are still not removed after step $n+1$, ca_{n+1}, is given by the following equation (ca_n denotes the ratio of the numbers that are still available after removal in step n)

$$ca_{n+1} = \left(1 - \frac{2}{5}\right) \times ca_n. \tag{2.9}$$

From (2.8) and (2.9) we can easily conclude that the following equation holds (since $p5(n) > 5, n > 1$)

$$ca_{n+1} < c_{n+1}, n \in N, \tag{2.10}$$

that can be interpreted that starting from step 2, R25 removes more balls than SEP defined by (7). However, it is not difficult to be seen that removal of the $2/5$ of the numbers from those that are left, will result in infinite number of numbers that cannot be removed, at the end of the process. In order to show this in an elementary way, we create an associated experiment with boxes and balls (BB experiment).

It is going to be assumed that an infinite number of numbered balls (with all natural numbers written on them only once), as well as, infinite number of boxes of proper (finite

or infinite) size, are available. At the beginning of the experiment, all balls are moved from the source box (SB) to the infinite number of experimental boxes (EB) of size 1. Our SEP process, R25, has corresponding BB experiment that fuses 5 EB's in every step (this step insures EB of proper size and proper number of balls, so that it enables removal of natural number of balls, and in the case of interest, there is no other way that insures that and create smaller size of EB). After that, in every step, 2/5 of the balls is removed from every EBs. We are going to check what is the number of the balls in the EB, at the end of the experiment, finite or infinite.

Experiment:

STEP 1 - (1 minute before midnight). Move all balls from SB to EBs of size 1, and fuse every 5 EBs to obtain the EBs of size 5 with 5 balls inside. Then, remove the 2 balls from every EB. So, in this moment nominator of ca_1 is $5-2=3$, and that is equal to the number of the balls in EB.

STEP 2 - (1/2 minute before midnight). Again fuse every 5 EBs to obtain the EBs of size 25 with $5 \times 3 = 15$ balls inside. Then, remove the 2/5 of the balls from every EB. In this moment number of the balls in each EB, that is equal to the nominator of ca_2, is $(1-2/5) \times 5 \times (5-2) = 3^2$.

...

STEP N-($1/2^{N-1}$ minute before midnight). Fuse every 5 EBs. Then, remove the 2/5 of the balls from every EB. In this moment $nominator(ca_N) = (1-2/5) \times 5 \times nominator(ca_{N-1}) = 3^N$, and that equals the number of the balls in every EB.

Now, we can conclude that nominator of ca_N is increasing function of time and at midnight, without suffering from collapse of elementary reasoning (CER), we can conclude that the number of the balls in the EB (which at midnight has the size of number of natural numbers) is going to be infinite ($\lim_{n\to+\infty} 3^n$ is $+\infty$). For instance, if we can conclude that

$$1+2+3+4+...=-\frac{1}{12},$$

we obviously have a CER problem. Standard summation of infinite series of natural numbers give us as a solution plus infinity, and in that case we have no CER problem.

Since the number of the balls in EB tends toward infinity at the end of BB experiment, we can conclude that R25 will leave infinitely many numbers that cannot be removed at the end of that SEP. Having that in mind, and equation (2.10), we can also conclude that the SEP proposed by (2.4, 2.6), will leave an infinite number of the numbers that cannot be represented by it. And that completes the proof that number of Sophie Germain primes is infinite.

References

[1] T. Agoh. On Sophie Germain primes, *Tatra Mt. Math. Publ* **20**(65) (2000), 65–73.

[2] M. Agrawal, N. Kayal, N. Saxena. PRIMES is in P, *Ann. of Math.* **160**(2) (2004), 781–793.

[3] H. M. Edwards. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer, 2000.

[4] R.A.J. Matthews. Maximally periodic reciprocals, *Bull. Inst. Math. Appl.* **28** (1992), 147–148.

[5] V. Shoup. *A Computational Introduction to Number Theory and Algebra*, Cambridge University Press, 2009.

[6] W.-S. Yap, S.L. Yeo, S.-H. Heng, M. Henricksen. Security analysis of GCM for communication, *Security and Communication Networks* **7**(5) (2014), 854–864.

[7] D. Wells. *Prime Numbers: The Most Mysterious Figures in Math.*, Hoboken, NJ: John Wiley & Sons, Inc., 2005, pp. 58–59.