# A proof of Sophie Germain primes conjecture

Marko V Jankovic

# One Proof of the Sophie Germain Primes Conjecture

Marko V. Jankovic

ARTORG Centre for Biomedical Engineering Research,
University of 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, the basic formula for prime numbers was analyzed with the intention of finding out when this formula would produce a Sophie Germain prime and when not. The 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 a 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 in the investigation of cases of Fermat's last theorem [3]. It has been conjectured that infinitely many Sophie Germain primes exist, but this was unproven (see for instance, [5]).

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

Remark 1: *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 6s- 1, where l, s ∈ N. In this paper most of the time we analyze 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.*

Remark 2: *In this paper any infinite number series in the form $c_1 * l \pm c_2$ is going to be called a thread, defined by number $c_1$. Here $c_1$ and $c_2$ are constants that belong to the set of natural numbers and l represents an infinite series of consecutive natural numbers in the form (1, 2, 3, …).*

# 2 Proof

Outline of the proof: *First, it is shown that Sophie Germain primes have to be in the form* $6s - 1$ ($s \in N$). *Then numbers that cannot be used for the generation of Sophie Germain primes are going to be analyzed. The number of those numbers is going to be compared with the number of numbers that are used for the generation of composite odd numbers. The comparison will lead to the conclusion that the number of numbers that can be used for the generation of Sophie Germain primes is infinite.*

*Note: The number of numbers n, that are used for the generation of odd numbers bigger than 1 (2n+1,* $n \in N$) *and the number of the numbers s, that are used to produce numbers in the form* $6s - 1$ ($s \in N$), *are the same.*

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 conclude that numbers in the form $6l + 1$, could never be Sophie Germain primes since their safe primes are in the form

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

and that is a composite number divisible by 3. Hence, the prime number that can potentially be a Sophie Germain prime must be in the form $6s - 1$. The safe prime will then 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 *CPN*5. A number in the form $6l + 1$ is marked with *mpl*, while a number in the form $6s - 1$ is marked with *mps* ($l, s \in N$). That means that any composite number *CPN*5 can be expressed in the form *mpl x mpl*, *mps x mps* or *mpl x mps*.

If we represent all composite numbers in *mps* form with $6k - 1$ ($k \in N$) it must hold

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

Since *CPN*5 should be in the *mps* form, *CPN*5 can be generally expressed as a product *mpl* x *mps*, or

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

which leads to

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

or, due to symmetry

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

which leads to

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

If (2.2) or (2.3) is replaced in (2.1) forms of $k$ that will not produce a Sophie Germain prime number will be obtained. Those forms are expressed by the following equations

$$k = (6x - 1)y + x \tag{2.4a}$$

$$k = (6x + 1)y - x \tag{2.4b}$$

where $x, y \in N$. These equations are equivalent (they will produce the same numbers) and can be used interchangeably.

Also, in order to have Sophie Germain pair (which consists of prime $p$ and safe prime $2p+1$), a safe prime which is generated by $2k$ cannot be composite. If the safe number is composite the following equation must hold

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

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

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

where $x, y \in N$. Alternatively, it is possible to use the equation (2.3) and replace it in (2.5) and then the following equation holds

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

A different equation that produces the same numbers as the equation (2.6a) is obtained.

Equations (2.4a) and (2.6a) (and other alternatives like (2.4b) and (2.6b)) give a sufficient and necessary condition for $k$, so that it cannot be used for the generation of prime pairs in the form $(p, 2p + 1)$. In order to prove that there are infinitely many prime pairs in the form $(p, 2p + 1)$ it is necessary to prove that infinitely many *natural numbers* $s$ ($s \in N$) exist, that cannot be expressed in the form (2.4a) or (2.6a).

First, the forms of (2.4a, 2.6a) for some values of $x$ are checked.

*Case x=1:* $k = 7y - 4, k = 5y + 1$

*Case x=2:* $k = 13y - 1, k = 11y + 2$

*Case x=3:* $k = 19y - 11, k = 17y + 3$

*Case x=4:* $k = 5(5y) - 2, k = 23y + 4$ *(here we should have the equivalence of the equations (2.6a) and (2.6b) in mind)*

*Case x=5:* $k = 31y - 18, k = 29y + 5$

*Case x=6:* $k = 37y - 3, k = 5(7y+1) + 1$

*Case x=7:* $k = 43y - 25, k = 41y + 7$

*Case x=8:* $k = 7(7y) - 4, k = 47y + 8$

It can be seen that $k$ is represented by the threads that are defined by prime numbers bigger than 3. From examples (cases $x$=4 and $x = 8$), it can be seen that if $(6x - 1)$ or $(6x + 1)$ represent a composite number, $k$ that is represented by the thread defined by that number, is also represented by the thread defined by one of the prime factors of that composite number. This can be easily proved by direct calculation in the general case, where

representations similar to (2.2) and (2.3) are used. Here only one case is going to be analyzed. All other cases can be analyzed analogously. In this case we assume

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

where ($l, s \in N$). Thus, the following equation holds

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

Considering that and using the following representation of $k$ that includes the form ($6x$ - 1)

$$k = (6x - 1)y + x,$$

the simple calculations leads to

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

or

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

which means

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

and these values of $k$ are also represented by the thread defined by ($6l$+1), where

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

Here the equivalency of the equations (2.4a) and (2.4b) is used. It can be seen that all patterns for $k$ are represented by the threads defined by prime numbers bigger than 3. Now, proof that the number of natural numbers $s$, that cannot be represented by the models (2.4a) and (2.6a) is infinite, will be provided.

In order to do that, the formula for odd numbers and the formula for the composite numbers are going to be analyzed. It is well known that odd numbers $a$ (bigger than 1) can be represented by the following formula

$$a = 2n + 1,$$

where $n \in N$. It is not difficult to prove that all composite odd numbers $a_c$ can be presented by the following formula

$$a_c = 2(2ij + i + j) + 1 = 2\big((2j + 1)i + j\big) + 1, \tag{2.7}$$

where $i, j \in N$. (<u>It is clear that odd prime numbers are odd numbers that cannot be represented by the previous formula and it is known that the number of the primes is infinite.)</u> It is simple to conclude that all composite numbers could be represented by the product ($i$+1) * ($j$+1), where $i, j \in N$. If it is checked how that formula looks like for odd numbers, after simple calculation, equation (2.7) is obtained. This calculation is presented here. The form $2m$+1, $m \in N$, will represent odd numbers that are composite. Then the following equation holds

$$2m + 1 = (i_1 + 1)(j_1 + 1),$$

where $i_1, j_1 \in N$. *Now it is easy to see that the following equation holds*

$$m = \frac{i_1 j_1 + i_1 + j_1}{2}.$$

In order to have $m \in N$, it is easy to check that $i_1$ and $j_1$ have to be in the forms

$$i_1 = 2{*}i \text{ and } j_1 = 2{*}j,$$

where $i, j \in N$. From that, it follows that $m$ must be in the form

$$m = 2ij + i + j.$$

Now, it is going to be checked which numbers can be presented by the formula $m = 2ij + i + j$ for some $j$.

*Case j=1:* $m = 3i + 1$

*Case j=2:* $m = 5i + 2$

*Case j=3:* $m = 7i + 3$

*Case j=4:* $m = 3(3i + 1) + 1$

*Case j=5:* $m = 11i + 5$

*Case j=6:* $m = 13i + 6$

*Case j=7:* $m = 5(3i + 1) + 2$

*Case j=8:* $m = 17i + 8$

It can be seen that *m* is represented by threads that are defined by odd prime numbers. Here, again, it can be seen from the examples that the threads that are defined by some composite number, can be generated by some of the threads defined by one of the prime factors of that composite number. This can be proved in the general case in the same manner as it was done for *k* that represents numbers that cannot be used for the generation of Sophie Germain primes.

Both the *k,* that represents numbers that cannot generate Sophie Germain primes, and *m,* that represents composite odd numbers, are represented by the threads defined by odd prime numbers. The differences are:

- in the case of composite odd numbers an additional thread that is defined by number 3, exists,
- in the case of numbers that cannot represent Sophie Germain primes, the threads defined by the primes in the *mpl* form, potentially represent one number more than

the thread defined by the same prime in the case of composite odd numbers, (if they are not already represented by some of the threads defined by the smaller prime).

It is not difficult to prove that the number of primes in the *mpl* form is much smaller than the number of numbers that are represented by the thread defined by number 3. However, there is also simpler proof.

Here, let us ignore the fact that an additional thread that is defined by the prime 3, in the case of composite odd numbers, exists. (Here, again, it is going to be stressed that the number of numbers, that are used for generation of odd numbers bigger than 1, is the same as the number of the numbers that are used to produce numbers in *mps* form.) It is known that infinitely many numbers, that define odd primes (infinite number of numbers $n$ ($n \in N$) that cannot be represented by (2.7)) exist. As already stated, in the case of numbers that cannot represent the Sophie Germain primes, every thread that is defined by the primes in the *mpl* form, potentially represents one number more than the thread defined by the same prime in the case of composite odd numbers. This results in the number of numbers that **can** produce Sophie Germain primes being equal or bigger than the number obtained by subtraction of the number of primes in the *mpl* form from the number of odd primes. This calculation gives us the number of numbers that can generate primes in the *mps* form (plus the number that generates number 3, but this has no impact on the final conclusion). As known from the Dirichlet's prime number theorem [7], infinitely many prime numbers in the *mps* form exist. That means that an infinite number of natural numbers $s$ ($s \in N$), that cannot be represented by the models (2.4a) and (2.6a) exist. And that completes the proof that the 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] Dirichlet, P. G. L. Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält, [Proof of the theorem that every unbounded arithmetic progression, whose first term and common difference are integers without common factors, contains infinitely many prime numbers], Abhandlungen der Königlichen Preußischen Akademie der Wissenschaften zu Berlin, 48 (1837), 45–71.