A proof that exists an infinite number of Sophie Germain primes

Marko V. Jankovic

ARTORG Centre for Biomedical Engineering Research,

University of Bern,

Murtenstrasse 50, 3008 Bern, Switzerland

email: marko.jankovic@artorg.unibe.ch

fax: +41 31 632 75 76

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 by elementary math.

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 when we address prime numbers, we talk about the 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

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

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

$$CPN5 = mpl \times mps = 6(6xy - x + y) - 1, \tag{2}$$

or, due to symmetry

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

which leads to

$$CPN5 = mpl \times mps = 6(6xy + x - y) - 1. \tag{3}$$

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

$$k = \begin{cases} (6x-1)y + x \\ (6x+1)y - x \end{cases}, \tag{4}$$

where $x, y \in N$. In order to prove the conjecture that exists infinitely many prime pairs in the form $(p, 2p+1)$, we need to check when $(2p+1)$ is going to be a composite number although p is prime number. First, we are going to consider case when k is in the form $(6x-1)y + x$. In that case we know that $k1$ that generates $2p+1$ is

$$k1 = 2k.$$

So, we are going to check situations in which k cannot be expressed in the form $(6x-1)y+x$ and $2k$ can be expressed in the form $(6x-1)y+x$. (Simple example is in the case $x = 1$, than $k = 5y+1$ produces composite number, while the forms $k = 5y, k = 5y+2, k = 5y+3, k = 5y+4$, still have potential to produce prime numbers. However if $k = 5y+3$, than $k1 = 2k$ is actually in the form $k1 = 5y+1$. So, the forms of $k = 5y+1 \wedge k = 5y+3$ cannot produce Sophie Germain prime.) In order to check it, $k = (6x-1)y + a$, where $a \in \{0, 1, ..., 6x-2\} \wedge a \neq x$ are going to be analysed. In those cases k has potential to produce prime numbers. In those cases $k1 = 2k = 2((6x-1)y+a)$ is going to produce a composite number if $k1$ can be expressed as

$$k1 = 2((6x-1)y+a) = (6x-1)z+x,$$

where $z \in N$. If we solve this equation for a we get

$$a = \frac{(6x-1)(z-2y)+x}{2}. \tag{5}$$

Now, we analyse 2 cases, when x is even and odd, separately. In the case x is even x can be expressed as $x = 2m, m \in N$. In that case equation (2.5) can be written as

$$a = \frac{(12m-1)(z-2y)+2m}{2}. \tag{6}$$

Equation (2.6) has a solution that fulfils requirements for a only in the case $z-2y=0$, and in that case

$$a = m = \frac{x}{2}. \tag{7}$$

In the case x is odd, x can be expressed as $x = 2m-1, m \in N$. In that case equation (2.5) can be written as

$$a = \frac{(12m-7)(z-2y)+2m-1}{2}. \tag{8}$$

Equation (2.8) has a solution that fulfils requirements for a only in the case $z-2y=1$, and in that case

$$a = 7m-4 = 3x+\frac{x-1}{2}. \tag{9}$$

Using similar analysis for the case when k is in the form $(6x+1)y-x$, we obtain that in the case k is even, a is defined as

$$a = -\frac{x}{2}, \tag{10}$$

while in the case x is odd, a is defined as

$$a = -7m + 3 = -3x - \frac{x+1}{2}. \tag{11}$$

Now we can say, that in addition to equation (2.4), k cannot produce prime pairs in the form $(p, 2p+1)$ in the following cases

$$k = \begin{cases} (6x-1)y + \frac{x}{2},\ x \text{ is even} \\ (6x+1)y - \frac{x}{2},\ x \text{ is even} \\ (6x-1)y + 3x + \frac{x-1}{2},\ x \text{ is odd} \\ (6x+1)y - 3x - \frac{x+1}{2},\ x \text{ is odd} \end{cases}, \tag{12}$$

where $x, y \in N$. Equations (2.4) and (2.12) 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 (2.12). In order to do it, the number of k that cannot be expressed in the forms (2.4, 2.12) is going to be calculated. First, we will check the form of (2.4, 2.12) 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 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

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

So, we can see that all patterns for k that potentially result in composite number, include prime numbers and we can calculate how many k cannot be represented by the models (2.4, 2.12). 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 form

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

are removed from natural numbers, it can be seen that $r_1 = 2/5$ of all natural numbers are removed. So, $c_1 = 1 - 2/5 = 3/5$ of all natural numbers cannot be represented by those two patterns and they still contain some k that could be 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 (together with previously removed numbers, and 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 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$)

$$r_2 = r_1 + \frac{2}{7} - \frac{2}{7} \times r_1 = \frac{2}{5} + \frac{2}{7} - \frac{4}{5 \times 7} = \frac{20}{5 \times 7}.$$

So, we have

$$c_2 = 1 - r_2 = \frac{3 \times 5}{5 \times 7},$$

of all natural numbers that potentially can be used for "generation" of Sophie Germain primes.

Now, we denote prime numbers bigger than 3 as $p5$, and define set S as

$$S = \{p5(1), p5(2), ..., p5(i)\},$$

where i is a natural number (in this moment we consider i bigger than 2), and $p5(1) = 5$, $p5(2) = 7$ and so on. So, after removal of i^{th} pair of numbers related to the i^{th} $p5$, we will have

$$r_i = 2^1 \sum_{p \in S} \frac{1}{p} - 2^2 \sum_{\substack{p,q \in S, \\ p \neq q}} \frac{1}{p \times q} + 2^3 \sum_{\substack{p,q,r \in S, \\ p \neq q \neq r}} \frac{1}{p \times q \times r} + ... + (-1)^{i+1} 2^i \prod_{p \in S} \frac{1}{p},$$

and

$$c_i = 1 - r_i = \frac{\prod_{p \in S}(p-2)}{\prod_{p \in S} p}.$$

This formula can be proved by using mathematical induction method (but also by some other methods). Here we briefly present the induction method.

We have already seen that basis of the induction method for $i = 2$ is already satisfied. We suppose that formula holds for $i = n$. Now, we are going to prove that it holds for $i = n + 1$, too. So, since it holds

$$c_n = 1 - r_n = \frac{\prod_{j=1}^{n}(p5(j) - 2)}{\prod_{j=1}^{n} p5(j)},$$

also holds

$$r_n = 1 - c_n = \frac{\prod_{j=1}^{n} p5(j) - \prod_{j=1}^{n}(p5(j) - 2)}{\prod_{j=1}^{n} p5(j)}.$$

When we remove the next pair of numbers that correspond to $p5(n + 1)$ 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, we obtain the following equation

$$r_{n+1} = \frac{\prod_{j=1}^{n+1} p5(j) - \prod_{j=1}^{n+1}(p5(j) - 2)}{\prod_{j=1}^{n+1} p5(j)} = 1 - \frac{\prod_{j=1}^{n+1}(p5(j) - 2)}{\prod_{j=1}^{n+1} p5(j)},$$

which means that c_{n+1} is defined by the following equation

$$c_{n+1} = \frac{\prod_{j=1}^{n+1}(p5(j) - 2)}{\prod_{j=1}^{n+1} p5(j)},$$

and that concludes the proof.

If we continue process until we remove all possible patterns (defined by (2.4, 2.12)) related to all prime numbers bigger than 3 (and that is an infinite number), the number of k that cannot be represented by any of the combinations in the form (2.4, 2.12), is given by

$$C = \lim_{n \to +\infty} \left(c_n \times \prod_{j=1}^{n} p5(j) \right) = \prod_{j=1}^{+\infty} (p5(j) - 2).$$

It can be easily concluded that C is an infinite number. That concludes 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.