Notes on Number Theory and Discrete Mathematics
Print ISSN 1310–5132, Online ISSN 2367–8275
2023, Volume 29, Number 1, 62–77
DOI: 10.7546/nntdm.2023.29.1.62-77

A primality test for $Kp^n + 1$ numbers and a generalization of Safe primes and Sophie Germain primes

Abdelrahman Ramzy

Department of Mathematics, Faculty of Education,
Al-Azhar University, Cairo, Egypt
e-mail: hosam7101996@gmail.com

Received: 4 August 2022 **Revised:** 6 January 2023
Accepted: 19 February 2023 **Online First:** 22 February 2023

Abstract: In this paper, we provide a generalization of Proth's theorem for integers of the form $Kp^n + 1$. In particular, a primality test that requires a modular exponentiation (with a proper base a) similar to that of Fermat's test without the computation of any GCD's. We also provide two tests to increase the chances of proving the primality of $Kp^n + 1$ primes. As corollaries, we provide three families of integers N whose primality can be certified only by proving that $a^{N-1} \equiv 1 \pmod{N}$ (Fermat's test). One of these families is identical to Safe primes (since $N-1$ for these integers has large prime factor the same as Safe primes). Therefore, we considered them as a generalization of Safe primes and defined them as a-Safe primes. We address some questions regarding the distribution of those numbers and provide a conjecture about the distribution of their generative numbers a-Sophie Germain primes which seems to be true even if we are dealing with 100, 1000, or 10000 digits primes.

Keywords: Primality test, Safe prime, Sophie Germain prime.
2020 Mathematics Subject Classification: 11Y11, 11N80, 11N05.

1 Introduction

One of the fundamental facts about prime numbers is Fermat's result that if N is prime then for every integer a we have

$$a^N \equiv a \pmod{N}. \tag{1.1}$$

This leads to Fermat's primality test which is a probabilistic test. But since verifying (1.1) for a given a and N is computationally inexpensive, there has been many results combining additional conditions with (1.1) to conclude primality. For example, in 1878, Proth (see [6]) presented the following theorem to determine whether N is prime or not when N is of the form $N = K2^n + 1$ (Proth numbers).

Theorem 1.1 (Proth, 1878)**.** *Let $N = K2^n + 1$, where K is odd and $K < 2^n$. If there exists an integer $a > 1$, such that $a^{\frac{N-1}{2}} \equiv -1 \pmod{N}$, then N is prime.*

In 1914, Pocklington gave the first generalization of Proth's theorem suitable for numbers of the form $N = Kp^n + 1$, which are called *Generalized Proth numbers* (see [6]):

Theorem 1.2 (Pocklington, 1914)**.** *Let $N = Kp^n + 1$ with $K < p^n$, and p is prime. If there exists an integer $a > 1$ such that:*

(i) $a^{N-1} \equiv 1 \pmod{N}$*, and*

(ii) $\gcd(a^{\frac{N-1}{p}} - 1, N) = 1$*,*

then N is prime.

In [2], the following simpler generalization was presented by Grau, Oller-Marcén, and Sadornil.

Theorem 1.3 (Grau et al., 2015)**.** *Let $N = Kp^n + 1$, where p is prime and $K < p^n$. Assume that $a \in \mathbb{Z}$ is a p-th power non-residue, then N is a prime if and only if $\Phi_p\left(a^{\frac{N-1}{p}}\right) \equiv 0 \pmod{N}$.*

Then they gave a more useful test which increases the chances of proving the primality of N (if N is prime indeed). Note that we replaced (J) in the original theorem by $(n - j)$ in order to compare it to Theorem 2.2 in the following section. Thus, the conditions $n - 1 \geq j \geq 0$ and $2(n - j) > \log_p(K) + n$ in the following theorem are equivalent to the conditions $1 \leq J \leq n$ and $2J > \log_p(K) + n$ in the original theorem in [2].

Theorem 1.4 (Grau et al., 2015)**.** *Let $N = Kp^n + 1$, where p is prime. If there exists $n - 1 \geq j \geq 0$ such that:*

(i) $\Phi_p(a^{Kp^{n-j-1}}) \equiv 0 \pmod{N}$

(ii) $2(n - j) > \log_p(K) + n$*.*

Then N is prime.

Theorem 1.3 states that if $\Phi_p\left(a^{\frac{N-1}{p}}\right) \equiv 0 \pmod{N}$ such that $N = Kp^n + 1$, then N is prime. In practice, $(\Phi_p(x))$ is easily computed as $\dfrac{x^p - 1}{x - 1}$. Thus, to use Theorem 1.3 we would need to verify that $a^{N-1} \equiv 1 \pmod{N}$ and that $a^{\frac{N-1}{p}} - 1$ is invertible $\pmod{N}$. The latter condition

can be verified by checking that $\gcd(a^{\frac{N-1}{p}}-1,N)=1$, which is well known to be an $O(\log N)$ computation. Similarly, computing $\Phi_p\left(a^{Kp^{n-j-1}}\right)$ in Theorem 1.4 will require the $\gcd$ step. Note that we can compute $\Phi_p(x)$ as $1+x+x^2+\cdots+x^{p-1}$ but only when p is small, since computing $\Phi_p(x)$ that way would be time consuming if p is large.

We organized the paper as follows. In Section 2, we prove the required lemmas that we use to prove (Theorems 2.1, 2.2, 2.3 and 2.4). Theorems 2.1 and 2.2 generalize and simplify Proth's theorem, and Theorems 2.3 and 2.4 increase the chances of proving the primality of Kp^n+1 primes such that $p^j \geq p^{n-j} \geq (N-1)^{1/3}$ or $p^j \geq p^{2(n-j)}$ and $p^{n-j} \geq (N-1)^{2/7}$. Note that the conditions on p^{n-j} will be satisfied mostly, since in most cases p^n is much larger than K. As corollaries, we give three families of integers N whose primality can be certified only by proving that $a^{N-1} \equiv 1 \pmod N$ (Fermat's test). The family obtained from Corollary 2.1 is identical to Safe primes (since $N-1$ for it's integers has large prime factor the same as Safe primes). Therefore, we considered them as a generalization of Safe primes and defined them as a-Safe primes, and defined their generative primes as a-Sophie Germain primes.

In Section 3, we address some questions regarding the distribution of a-Safe primes and provide some computations for 2-Safe primes, and the distribution of 2-Sophie Germain primes will be addressed in Section 4.

In Section 4, we find the probability that a given prime is a-Sophie Germain prime. Then we give some computations for 2-Sophie Germain primes which shows that the accuracy of the estimates is very acceptable, even if we are dealing with 100, 1000 or 10000 digits primes (random or consecutive). We also give a conjecture about the distribution of 2-Sophie Germain primes (and a-Sophie Germain primes in general) which helped us to discover many primes larger than 10^{999} and some of them were larger than 10^{9999}.

2 Generalizing Proth's theorem and increasing the chances of proving the primality of Kp^n+1 primes

In this section we shall state and prove Theorems 2.1, 2.2, 2.3 and 2.4. They provide a simple primality test for generalized Proth's numbers $N=Kp^n+1$ and increase the chances of proving the primality of some Kp^n+1 primes. Then we prove three corollaries whose provide three families of integers N whose primality can be certified only by proving that $a^{N-1} \equiv 1 \pmod N$. To prove the theorems we require the following lemmas.

2.1 Proving required lemmas

Recall that the *order* of $a \pmod N$ is the least positive integer m such that $a^m \equiv 1 \pmod N$. We shall denote the order of $a \pmod N$ with $\mathrm{ord}_N(a)$.

Lemma 2.1. *Assume that $p^n \mid \mathrm{ord}_N(a)$ where p is prime and $\gcd(p,N)=1$. Then there exist a prime divisor q of N such that $p^n \mid \mathrm{ord}_q(a)$, therefore $p^n \mid q-1$.*

Proof. Assume that the prime factorization of $N = q_1^{e_1} \cdots q_s^{e_s}$. Then $\mathrm{ord}_N(a) = \mathrm{lcm}(\mathrm{ord}_{q_1^{e_1}}(a), \ldots, \mathrm{ord}_{q_s^{e_s}}(a))$. But since $p^n \mid \mathrm{ord}_N(a)$, then there exist a prime divisor q_c of N such that $p^n \mid \mathrm{ord}_{q_c^{e_c}}(a)$. But since $\mathrm{ord}_{q_c^{e_c}}(a) = q_c^k \times \mathrm{ord}_{q_c}(a)$, where $k \leq e_c - 1$, and since $\gcd(p, N) = 1$. Then p^n divide $\mathrm{ord}_{q_c}(a)$, therefore, $p^n \mid q_c - 1$. □

Lemma 2.2. *Assume that A, P are integers with $1 \leq A \leq P$. If there is an integer $D > 0$ such that*

$$\frac{AP+1}{DP+1} \in \mathbb{Z},$$

then we must have $D = A$.

Proof. If $DP + 1 \mid AP + 1$ we must have $D \leq A$. Write $A = cD + r$ with $c > 0$ and $0 \leq r < D$. Hence

$$\frac{(cD+r)P+1}{DP+1} = \frac{cDP + rP + 1 + c - c}{DP+1} = c + \frac{rP + 1 - c}{DP+1}.$$

Note that $c \leq A \leq P$, so if $r \geq 1$ we will have $rP - c + 1 \geq 1$, hence the numerator of the last fraction is a positive number which is strictly less than its denominator (since $r < D$). Thus the fraction can only be an integer when $r = 0$ and $c = 1$, proving the result. □

Lemma 2.3. *If $X > 11 + 4\sqrt{7}$, $a \leq \frac{X}{2}$, $a \neq 1$, $a \neq \sqrt{X+1}$ and $\frac{(X-a)(a)+1}{X} = M \in \mathbb{Z}$. Then $M > \sqrt{X}$.*

Proof. If $\frac{(X-a)(a)+1}{X} \in \mathbb{Z}$, then $\frac{-a^2+1}{X} \in \mathbb{Z}$, as well. Therefore, we could have:

(i) $\frac{-a^2+1}{X} = 0 \Longrightarrow a = 1 \Longrightarrow M = 1$

(ii) $\frac{-a^2+1}{X} = -1 \Longrightarrow a = \sqrt{X+1} \Longrightarrow M = \sqrt{X+1} - 1$

(iii) $\frac{-a^2+1}{X} \leq -2 \Longrightarrow \frac{a^2-1}{X} \geq 2$.

The cases (i) and (ii) are excluded by the assumptions, since in case (i) $a = 1$ and in case (ii) $a = \sqrt{X+1}$.

In case (iii) we have $\frac{a^2-1}{X} \geq 2$. Also $a > \sqrt{X} + 2$ (since $X > 11 + 4\sqrt{7}$). But for which $X > 11 + 4\sqrt{7}$ and $\frac{X}{2} \geq a > \sqrt{X} + 2$ we will have $\frac{(X-a)(a)+1}{X} > \sqrt{X}$, which proves our result. □

2.2 Generalizing Proth's theorem

Theorem 2.1. *Let $N = Kp^n + 1$, where p is prime, $p^n \geq K$. Assume that there exists an integer a such that:*

(i) $a^{Kp^{n-1}} \equiv L \neq 1 \pmod{N}$, *and*

(ii) $L^p \equiv 1 \pmod{N}$.

Then N is prime.

Proof. Assume both of (i) and (ii) hold, then $p^n \mid \mathrm{ord}_N(a)$. But since $\gcd(p, N) = 1$, then there exist a prime divisor q of N such that $p^n \mid q - 1$, due to Lemma 2.1. Hence, $q = hp^n + 1$ divides N implying $\frac{Kp^n+1}{hp^n+1} \in \mathbb{Z}$. But due to Lemma 2.2 we must have $h = K$, therefore, $q = N$ and N is prime. □

Theorem 2.1 and the results of Pocklington and Grau–Oller-Marcén–Sadornil (Theorems 1.2 and 1.3) are normal generalizations of Proth's theorem. But with all respect to their results we can replace the $\gcd$ step by a simple non-equality (namely $a^{\frac{N-1}{p}} \not\equiv 1 \pmod N$). Computationally and theoretically, this provides a simpler generalization of Proth's theorem.

As an example to illustrate Theorem 2.1, consider $N = 2 \cdot 107^3 + 1 = 2450087$. We can verify that (i) $2^{2 \cdot 107^2} \equiv 1302367 \pmod{2450087}$ and (ii) $1302367^{107} \equiv 1 \pmod{2450087}$. Thus, the hypotheses of Theorem 2.1 are satisfied and we deduce that N is prime. Of course, we can determine the primality of N by Theorems 1.2 and 1.3, but both of them require the additional $\gcd$ step.

2.3 Improving Pocklington's theorem

The following theorem and the second result of Grau–Oller-Marcén–Sadornil (Theorem 1.4) both improve Pocklington's theorem. Because their conditions of primality can be satisfied by many bases more than the satisfying bases of Pocklington's theorem. Therefore, they increase the chances of proving the primality of $Kp^n + 1$ primes.

Theorem 2.2. *Let $N = Kp^n + 1$, where p is prime, $p^{n-j} \geq Kp^j \implies p^{2(n-j)} \geq Kp^n \implies 2(n-j) \geq \log_p(K) + n$ where $(0 \leq j \leq n-1)$. Assume that there exists an integer $a > 1$ such that:*

(i) $a^{Kp^{n-j-1}} \equiv L \neq 1 \pmod N$, *and*

(ii) $L^{p^{j+1}} \equiv 1 \pmod N$.

Then N is prime.

Proof. As in the proof of Theorem 2.1 we can deduce that there exist a prime divisor q of N such that $p^{n-j} \mid q-1$. Thus, $\frac{Kp^j.p^{n-j}+1}{hp^{n-j}+1} \in \mathbb{Z}$, but due to Lemma 2.2 we must have $h = Kp^j$, therefore, $q = N$ and N is prime. □

Similarly here and with all respect to the result of Grau–Oller-Marcén–Sadornil, we can replace the $\gcd$ step by that simple non-equality $a^{Kp^{n-j-1}} \not\equiv 1 \pmod N$. Note that if N is indeed prime, then the chances of proving the primality of N are even whether we used Theorem 1.4 or Theorem 2.2. Since if $N = Kp^n + 1$ then the base (a) will fail to satisfy condition (i) in Theorems 2.2 and 1.4 only when (a) itself is a p^{j+1}-th power residue modulo N. That happens exactly for $\frac{1}{p^{j+1}}$ of the possible choices for (a), which means that condition (i) in Theorems 2.2 and 1.4 will be satisfied by a number of bases that equals to $Kp^n - Kp^{n-j-1}$. But since (j) in Theorem 2.2 can be as large as (j) in Theorem 1.4 then the number of the satisfying bases of (i) in both of them is the same.

2.4 Increasing the chances of proving the primality of $Kp^n + 1$ primes

The following two theorems can increase the chances of proving the primality of some $Kp^n + 1$ primes. Because their conditions of primality can be satisfied by many bases more than the satisfying bases of Theorems 2.2 and 1.4. We shall give an example after each theorem to show their main contributions.

Theorem 2.3. *Let $N = Kp^n + 1$, where p is prime, $p^j \geq p^{n-j} \geq (N-1)^{1/3}$ and $N \neq p^{3(n-j)} + 1$. If there exists an integer $a > 1$ such that:*

(i) $a^{Kp^{n-j-1}} \equiv L \neq 1 \pmod{N}$, *and*

(ii) $L^{p^{j+1}} \equiv 1 \pmod{N}$.

Then N is prime.

Proof. Assume both of (i) and (ii) hold, then $p^{n-j} \mid \mathrm{ord}_N(a)$. But since $\gcd(p, N) = 1$, then there exist a prime divisor q of N such that $p^{n-j} \mid q - 1$, due to Lemma 2.1. Therefore, we have two cases:

(i) $q = N$ and N is prime, or

(ii) $q < N$ and N is composite.

Now, we will try to exclude case (ii) by contradiction. In (ii), we have $q \mid N$, hence, $\frac{N}{q} = M \in \mathbb{Z}$. But since $q = ap^{n-j} + 1$ then $M = bp^{n-j} + 1$ as well. Now, if we multiply q by M we get $N = Kp^n + 1 = abp^{2(n-j)} + (a+b)p^{n-j} + 1 \implies Kp^j = abp^{n-j} + a + b$. Therefore, $p^{n-j} \mid (a+b) \implies a + b \geq p^{n-j} \implies a + b \geq (N-1)^{1/3}$. But since $ab < (N-1)^{1/3}$, and since $ab \geq a + b - 1$. Then the only chance to have $ab < (N-1)^{1/3}$ and $p^{n-j} \mid (a+b)$ is when $p^{n-j} = (N-1)^{1/3}$, $ab = p^{n-j} - 1$ and $a + b = p^{n-j} \implies N = p^{3(n-j)} + 1$. But the latter case is excluded by the assumption on N. Thus, $q = N$ and N is prime. □

To see the utility of Theorem 2.3 recall that if $N = Kp^n + 1$ is prime, then a will fail to satisfy condition (i) of Theorems 1.4, 2.2 and 2.3 only when a itself is a p^{j+1}-th power residue modulo N. That happens exactly for $\frac{1}{p^{j+1}}$ of the possible choices for a. But (j) in Theorem 2.3 can be larger than (j) in Theorems 2.2 and 1.4. Thus, (i) in Theorem 2.3 can be satisfied by many bases more than the satisfying bases of Theorems 2.2 and 1.4.

To give a concrete example, consider $N = 2 \cdot 3^{17} + 1$. Then, according to Theorems 1.4 and 2.2, we can only have $j \leq 8$. However, taking $1 \leq j \leq 8$ and $a = 136837116$ will not satisfy (i) in neither Theorem 1.4 nor Theorem 2.2. But in Theorem 2.3 we can have $j \leq 11$, and by taking $j = 9$ for example, we obtain $136837116^{2 \cdot 3^7} \equiv 216758952 \pmod{N}$ and the result follows from Theorem 2.3.

Theorem 2.4. *Let $N = Kp^n + 1 \neq (\sqrt{p^{n-j} + 1} - 1)p^{3(n-j)} + 1$ and $N \neq H^3 p^{3(n-j)} + 1$, where p is prime, $p^j \geq p^{2(n-j)}$, $p^{n-j} \geq (N-1)^{2/7}$ and $p^{n-j} > 11 + 4\sqrt{7}$. If there exists an integer $a > 1$ such that:*

(i) $a^{Kp^{n-j-1}} \equiv L \neq 1 \pmod{N}$, *and*

(ii) $L^{p^{j+1}} \equiv 1 \pmod{N}$.

Then N is prime.

Proof. As in Theorem 2.3, we can deduce that there exist a prime divisor q of N such that $p^{n-j} \mid q - 1$. Thus, we have two cases:

(i) $q = N$ and N is prime, or

(ii) $q < N$ and N is composite.

Now, we will try to exclude case (ii) by contradiction. In (ii), we have $q \mid N$, hence, $\frac{N}{q} = M \in \mathbb{Z}$, but since $q = ap^{n-j} + 1$ then $M = bp^{n-j} + 1$ as well. Now, if we multiply q by M we will get $N = Kp^n + 1 = abp^{2(n-j)} + (a+b)p^{n-j} + 1 \implies Kp^j = abp^{n-j} + a + b$. Therefore, $p^{n-j} \mid a + b$ and $p^{n-j} \mid (ab + \frac{a+b}{p^{n-j}})$, which leads to the following two cases (let $p^{n-j} = X$ for abbreviation):

(1) $a, b < X \implies a + b = X \implies X \mid (ab + 1) \implies \frac{ab+1}{X} = \frac{N-1}{X^3}$

(2) $a = AX + c, b = X - c \implies X \mid (AX^2 - AcX + cX - c^2 + A + 1) \implies AX - Ac + c + \frac{A+1-c^2}{X} = \frac{N-1}{X^3}$

In case (1) we have $\frac{ab+1}{X} \in \mathbb{Z}$, and by assuming that $a \leq b \implies \frac{(X-a)(a)+1}{X} \in \mathbb{Z}$. Then due to Lemma 2.3 (since $X > 11 + 4\sqrt{7}$) we can deduce that $a = 1 \implies \frac{N-1}{X^3} = 1$ or $a = \sqrt{X+1} \implies \frac{N-1}{X^3} = \sqrt{X+1} - 1$ or $a > \sqrt{X} + 2 \implies \frac{(X-a)(a)+1}{X} = \frac{N-1}{X^3} > \sqrt{X}$ and the latter is impossible since $X \geq (N-1)^{2/7}$. Hence, we have excluded all the cases of (1) (since we assumed that $N \neq H^3 p^{3(n-j)} + 1$ and $N \neq (\sqrt{p^{n-j}+1} - 1)p^{3(n-j)} + 1$).
In case (2) we must have $\frac{A+1-c^2}{X} \in \mathbb{Z}$ which leads to the following three cases:

(α) $A + 1 - c^2 \geq X$, which is impossible since $A < \sqrt{X}$.

(β) $A + 1 - c^2 = 0 \implies c = \sqrt{A+1}$. In this case we will have $ab = (AX + \sqrt{A+1})(X - \sqrt{A+1}) = AX^2 - A\sqrt{A+1}X + \sqrt{A+1}X - (A+1) = AX(X - \sqrt{A+1}) + \sqrt{A+1}X - (A+1)$ which is clearly larger than $(X^{3/2})$. Because $\sqrt{A+1}X - (A+1) > 0$ and $AX(X - \sqrt{A+1}) > X^{3/2}$ (since $X - \sqrt{A+1} > \sqrt{X}$). Hence, we have excluded this case as well.

(γ) $c^2 - A - 1 \geq X$. In this case we should have $A + 1 = c^2 \pmod X = (X - c)^2 \pmod X$. But since $(X - c) < \sqrt{X}$ then $(X - c)^2 < X$ and we can deduce that $A = (X - c)^2 - 1 \implies \frac{N-1}{X^3} = ((X-c)^2 - 1)(X - c) + c - \frac{c^2 - 1 - ((X-c)^2 - 1)}{X} = (X - c)^3 - X + 2c - \frac{c^2 - 1 - X^2 + 2cX - c^2 + 1}{X} = (X - c)^3$, and we have excluded that case by the assumption $N \neq H^3 p^{3(n-j)} + 1$.

Hence, we have excluded all the cases of (ii) and then we can deduce that $q = N$ and N is prime. □

As an example to illustrate Theorem 2.4, consider $N = 14 \cdot 3^{18} + 1 = 5423886847$. Then according to Theorems 1.4 and 2.2 we can only have $j \leq 7$. However, taking $1 \leq j \leq 7$ and $a = 1481700844$ will not satisfy (i) in neither Theorem 1.4 nor Theorem 2.2. Also in Theorem 2.3 we can only have $j \leq 11$, and taking $1 \leq j \leq 11$ with $a = 1481700844$ will not satisfy (i) in Theorem 2.3 as well. But in Theorem 2.4 we can have $j \leq 12$, and by taking $j = 12$, we obtain $1481700844^{14 \cdot 3^{18-12-1}} \equiv 3256260648 \pmod N$ and the result follows from Theorem 2.4. Note that the condition $N = Kp^n + 1 \neq (\sqrt{p^{n-j}+1} - 1)p^{3(n-j)} + 1$ in Theorem 2.4 is not necessary (it is not necessary in Corollary 2.3 as well), since the only cases we can have $p^{n-j} + 1$ being square is when $p^{n-j} = 2^3$ or 3 which are smaller than $11 + 4\sqrt{7}$.

Remark 2.1. *There are some composite numbers $N = Kp^n + 1$ that cannot be proved composite by any of the above results or their corollaries (unless $(a, N) > 1$ which is the trivial case). Since if $(a, N) = 1$, then $a^{(N-1)/p} \equiv 1 \pmod{N}$ always. Of course there are many p-th power non-residue bases in $[1, N-1]$. But to find such bases we cannot use the easy way (namely by proving that $a^{(N-1)/p} \not\equiv 1 \pmod{N}$ and $a^{N-1} \equiv 1 \pmod{N}$) since $a^{(N-1)/p} \equiv 1 \pmod{N}$ always. There are many results that can deal with such numbers but not the results in this paper, and we still do not know if such numbers are finite or infinite. $N = 228842209 = 2016 \times 113513 + 1$ is an example.*

2.5 Making Fermat's primality test deterministic for some families of integers

The following three corollaries show that we can certify the primality of some families of integers only by proving that $a^{N-1} \equiv 1 \pmod{N}$ (Fermat's test). Note that the primality of these integers can be certified by the results of Pocklington and Grau–Oller-Marcén–Sadornil as well. But both of them will require the additional $\gcd$ step. The first family of these integers is of the form $N = 2Kp + 1$ with $2K < \log_a(2Kp + 1)$ (namely, $\frac{\log 2Kp+1}{\log a}$ such that $2 \leq a \leq N - 1$). But taking $a = 2$ is the best choice, since then we will have many numbers K satisfying the condition $2K < \log_a(2Kp^2 + 1)$. Similarly, taking $a = 2$ is the best choice for the other two families, since they are of the form $N = 2Kp^2 + 1$ (with $2K < \log_a(2Kp^2 + 1)$) or $N = 2Kp^3 + 1$ (with $2K < \log_a(2Kp^3 + 1)$).

Corollary 2.1. *Let $N = 2Kp + 1$, where p is prime, $2K \leq p$, and $2K < \log_a(2Kp + 1)$. Then N is prime if and only if*

$$a^{N-1} \equiv 1 \pmod{N} \tag{2.1}$$

Proof. By the hypothesis we have $2K < \log_a(2Kp + 1)$, hence $a^{2K} \equiv L \not\equiv 1 \pmod{N}$, but $a^{N-1} \equiv 1 \pmod{N}$. Thus, all the conditions of Theorem 2.1 are satisfied and N is prime. The converse is simply Fermat's primality test. □

Corollary 2.2. *Let $N = 2Kp^2 + 1$, where p is prime, $2K < p$ and $2K < \log_a(2Kp^2 + 1)$. If there exists an integer $a > 1$ such that:*

$$a^{N-1} \equiv 1 \pmod{N}. \tag{2.2}$$

Then N is prime.

Proof. Since $2K < p$ then the condition $p^j \geq p^{2-j} \geq (N-1)^{1/3}$ is satisfied at$j = 1$, and $a^{2Kp^{n-j-1}} = a^{2Kp^{2-1-1}} = a^{2K} \equiv L \not\equiv 1 \pmod{N}$ (since $2K < \log_a(2Kp^2 + 1)$). Also $a^{N-1} \equiv 1 \pmod{N}$. Thus, all the conditions of Theorem 2.3 are satisfied and N is prime. □

Corollary 2.3. *Let $N = 2Kp^3 + 1$, where p is prime, $p > 11 + 4\sqrt{7}$, $2K < \log_a(2Kp^3 + 1)$, $2K < \sqrt{p}$, and $2K$ is not a perfect cube. If there exists an integer $a > 1$ such that:*

$$a^{N-1} \equiv 1 \pmod{N}. \tag{2.3}$$

Then N is prime.

Proof. Since $2K < \sqrt{p}$, then the conditions $p^j \geq p^{2(3-j)}$ and $p^{3-j} \geq (N-1)^{2/7}$ are satisfied at $j = 2$, and $a^{2Kp^{n-j-1}} = a^{2Kp^{3-2-1}} = a^{2K} \equiv L \neq 1 \pmod{N}$ (since $2K < \log_a(2Kp^3+1)$). Also $a^{N-1} \equiv 1 \pmod{N}$ and $2K$ is not a perfect cube. Thus, all the conditions of Theorem 2.4 are satisfied and N is prime. □

The integers of the first family (namely, $2Kp+1$ numbers with $2K < \log_a(2Kp+1)$) are identical to Safe primes (since $N-1$ for these integers has large prime factor the same as Safe primes). This motivated us to consider these integers as a generalization of Safe primes and to address some questions regarding their density and distribution. We defined those numbers as a-Safe primes and they will be addressed in Section 3. Their generative numbers are defined as a-Sophie Germain primes and they will be addressed in Section 4.

3 A generalization of Sophie Germain and Safe primes

Recall that a prime N is called a *Safe prime* if $\frac{N-1}{2}$ is also prime. A prime q is called a *Sophie Germain prime* if $2q+1$ is prime. Thus, there is a one-to-one correspondence between Safe primes and Sophie Germain primes. Safe primes are useful cryptographic parameters as they are resilient against certain attacks, see [5, 6] for more details. In this section we provide certain classes of primes which generalize those notions and define them as a-Safe primes. Our motivation comes from the observation that the key feature of Safe primes is that $N-1$ has a "large" prime factor (namely $\frac{p-1}{2}$). We thus propose the following extension.

Definition 3.1. *a-Safe prime is a prime number of the form $N = 2Kp+1$, where p is prime and $2K < \log_a(2Kp+1)$.*

In the classical definition of Safe primes we simply have $K = 1$ and the above condition is clearly satisfied. Now we will define a-Sophie Germain primes.

Definition 3.2. *a-Sophie Germain prime is a prime p for which $2Kp+1$ is also prime for any K with $2K < \log_a(2Kp+1)$.*

The condition $2K < \log_a(2Kp+1)$ implies that $2K \leq \log_a(2Kp)$. Therefore, $a^{2K} \leq 2Kp \implies \frac{a^{2K}}{2K} \leq p$. Hence, we can replace the condition $2K < \log_a(2Kp+1)$ by $\frac{a^{2K}}{2K} \leq p$ which will help us to study the distribution of a-Safe primes and a-Sophie Germain primes.

Questions regarding the infinitude and density of a-Sophie Germain primes seem to be more or less of the same order of difficulty as the corresponding questions for the classical case, where one mostly relies on conjectures and heuristics (see [1, 4] for instance). The density of a-Sophie Germain primes will be addressed in the following section, and we will give an important conjecture about the distribution of 2-Sophie Germain primes (and a-Sophie Germain primes in general). As for the density of a-Safe primes, we recall Hardy and Littlewood's conjecture B [3] which states that for $K \geq 1$, the number of the prime pairs $(p, p+2K)$, where $p \leq x$, is asymptotically given by

$$2C_K \frac{x}{\log^2(x)} \sim 2C_K \int_2^x \frac{dt}{\log^2 t}, \tag{3.1}$$

where

$$C_K = c_2 \prod_{2<q,q|2K} \frac{q-1}{q-2},$$

and c_2 (the so-called *twin prime constant*) is given by

$$c_2 = \prod_{p\geq 3} \frac{p(p-2)}{(p-1)^2} \approx 0.66016181.$$

The exactly same asymptotics can be used to estimate the number of the primes of the form $2Kp+1$, where p is prime and $p \leq x$. Furthermore (as discussed in [1] for instance), the following function, which is asymptotic to the right-hand side of (3.1) gives a more accurate estimate, especially for smaller values of x

$$2C_K \int_2^x \frac{dt}{\log t \log 2Kt}. \tag{3.2}$$

If we assume that the number of $2Kp+1$ primes with K ranging between 1 and n and $p \leq x$ is simply asymptotic to the sum of the individual estimates (as given in (3.1) or (3.2)), then we get the estimate:

$$2\frac{x}{\log^2(x)}(C_1 + C_2 + \cdots + C_n). \tag{3.3}$$

If p is prime and $\frac{a^{2n}}{2n} \leq p < \frac{a^{2(n+1)}}{2(n+1)}$, then K can take values up to n with the condition $\frac{a^{2K}}{2K} \leq p$ satisfied.

For typographical convenience we set $f(a,n) := \frac{a^{2n}}{2n}$. Utilizing (3.3) the asymptotic behavior for the number of a-Safe primes for which $f(a,n) \leq p < f(a,n+1)$ is:

$$\alpha_1(a,n) := 2c_2 \left(\frac{f(a,n+1)}{\log^2(f(a,n+1))} - \frac{f(a,n)}{\log^2(f(a,n))} \right) \sum_{K=1}^{n} \prod_{2<q,q|2K} \frac{q-1}{q-2}.$$

Similarly, we can use the integral estimate in (3.2) to obtain the estimate:

$$\alpha_2(a,n) := 2c_2 \sum_{K=1}^{n} \prod_{2<q,q|2K} \frac{q-1}{q-2} \int_{f(a,n)}^{f(a,n+1)} \frac{dt}{\log t \log 2Kt}.$$

Table 1 lists samples of the actual count of 2-Safe primes for which p is in the interval $[f(2,n), f(2,n+1)]$ (labeled 2-Sp for brevity), as well as the corresponding estimates $\alpha_1(2,n)$, $\alpha_2(2,n)$ for n up to 14. We also list the number of 2-Sophie Germain primes in those intervals (labeled 2-SGp). We can not say that there is a one-to-one correspondence between 2-Sp and 2-SGp due to the simple fact that the same 2-SGp might give rise to more than one 2-Sp for various values of K (the same can be said as for correspondence between a-Sp and a-SGp). For example when $n = 3$ in the table below, the interval $[f(2,3), f(2,4)] = [2^6/6, 32]$ contains exactly 7 primes; namely 11 through 31. Two of them (19 and 31) are not 2-SGp, but three others give rise to two 2-Sp each (namely 11, 13 and 23).

Table 1. Some counts and asymptotics of 2-Safe primes

n	2-SGp in $[f(2,n), f(2,n+1)]$	2-Sp $2Kp+1$, p in $[f(2,n), f(2,n+1)]$	$\alpha_1(2,n)$	$\alpha_2(2,n)$
1	2	2	-2.74	1.5
2	2	2	-0.469	3
3	5	8	4	9
4	14	18	14	20
5	36	52	44	54
6	104	168	148	165
7	295	463	450	483
8	895	1414	1380	1447
9	2970	4854	4724	4836
10	9496	15783	15484	15672
11	30788	50832	50827	51030
12	104997	177808	178920	178090
13	353357	596973	602484	597141
14	1211233	2041459	2066125	2040547

4 The probability that a given prime is a-SGp

From Definition 3.2, it is obvious that for which prime $f(a,n) \leq p \leq f(a,n+1)$, there is a set of numbers of the form $2Kp+1$, where $1 \leq K \leq n$, that if there is at least one prime of them, then we call p an a-SGp. That set of numbers will be denoted by $S_p(n)$, namely, $S_p(n) = \{2p+1, 4p+1, 6p+1, \ldots, 2np+1\}$.

We cannot declare that the elements of $S_p(n)$ are independent. Since if $2p+1$ is divisible by 3 (for example), then all elements in $\{8p+1, 14p+1, 20p+1, \ldots\}$ will be divisible by 3, as well. Namely, we cannot declare an element of $S_p(n)$ is independent of the remained elements of $S_p(n)$ unless it is not divisible by any prime less than n. Therefore, to approximate the number of independent elements in $S_p(n)$ we should approximate the number of elements in $S_p(n)$ which are coprime to all primes less than n (except for 2, since $S_p(n)$ elements are coprime to 2 already). But since for any $N \in S_p(n)$ the probability that $(N,q) = 1$ (such that $3 \leq q$, and q is prime) is $\frac{q-1}{q}$. Then the probability that $(N,q) = 1$ for all primes $3 \leq q \leq n$ is $\prod_{3 \leq q \leq n} \frac{q-1}{q}$. Thus, the number of elements in $S_p(n)$ which are coprime to all primes less than n (it is also the number of independent elements in $S_p(n)$) is approximately

$$|S_p(n)| \prod_{3 \leq q \leq n} \frac{q-1}{q} = n \prod_{3 \leq q \leq n} \frac{q-1}{q} \approx n \frac{4e^{-\gamma}}{\log n^2}.$$

That set of independent numbers will be denoted by $ind\text{-}S_p(n)$. Now assume that $N = 2Kp + 1 \in ind\text{-}S_p(n)$. Then the probability that N is prime (since N is in the interval $[2f(a,n), 2nf(a,n+1)]$ and not divisible by any prime less than n) is about:

$$\frac{\int_{2f(a,n)}^{2nf(a,n+1)} \frac{dt}{\log t}}{(2nf(a,n+1) - 2f(a,n)) \prod_{2 \le q \le n} \frac{q-1}{q}}.$$

Hence, the probability that N is composite is about:

$$1 - \frac{\int_{2f(a,n)}^{2nf(a,n+1)} \frac{dt}{\log t}}{\frac{2e^{-\gamma}}{\log n^2}(2nf(a,n+1) - 2f(a,n))}.$$

Thus, the probability that all the elements of $ind\text{-}S_p(n)$ are composite is about:

$$\left(1 - \frac{\int_{2f(a,n)}^{2nf(a,n+1)} \frac{dt}{\log t}}{\frac{2e^{-\gamma}}{\log n^2}(2nf(a,n+1) - 2f(a,n))}\right)^{n\frac{4e^{-\gamma}}{\log n^2}}.$$

Therefore, the probability that not all the elements of $ind\text{-}S_p(n)$ are composite is about:

$$\alpha_3(a,n) := 1 - \left(1 - \frac{\int_{2f(a,n)}^{2nf(a,n+1)} \frac{dt}{\log t}}{\frac{2e^{-\gamma}}{\log n^2}(2nf(a,n+1) - 2f(a,n))}\right)^{n\frac{4e^{-\gamma}}{\log n^2}}.$$

Consequently, $\alpha_3(a,n)$ finds the probability that a given prime $f(a,n) \le p \le f(a,n+1)$ is a-SGp. In correspondence to $\alpha_3(a,n)$ we obtained a similar estimate by assuming that the elements of $S_p(n)$ are independent (in fact they are not). Surprisingly, this obtained estimate behaved similarly to $\alpha_3(a,n)$:

$$\alpha_4(a,n) := 1 - \left(1 - \frac{\int_{2f(a,n)}^{2nf(a,n+1)} \frac{dt}{\log t}}{\frac{1}{2}(2nf(a,n+1) - 2f(a,n))}\right)^{n}.$$

But we should acknowledge that both of $\alpha_3(a,n)$ and $\alpha_4(a,n)$ failed to expect the number of a-SGp for large values of a with small n. That compelled us to obtain the following estimate which depends on Hardy and Littlewood's Conjecture:

$$\alpha_5(a,n) := 1 - \prod_{K=1}^{n} \left(1 - \frac{2C_K \int_{f(a,n)}^{f(a,n+1)} \frac{dt}{\log t \log 2Kt}}{\int_{f(a,n)}^{f(a,n+1)} \frac{dt}{\log t}}\right).$$

Note that if we multiply either $\alpha_3(a,n)$, $\alpha_4(a,n)$ or $\alpha_5(a,n)$ by the expected number of primes in the interval $[f(a,n), f(a,n+1)]$ (namely, $\int_{f(a,n)}^{f(a,n+1)} \frac{dt}{\log t}$), then we will get an approximation to the number of a-SGp primes in the interval $[f(a,n), f(a,n+1)]$.

Table 2 lists some counts and asymptotics of 2-SGp primes in many different intervals which shows that the accuracy of either $\alpha_3(2,n)$, $\alpha_4(2,n)$ or $\alpha_5(2,n)$ is very acceptable.

Table 2. Some counts and asymptotics of 2-SGp

n	$\pi(f(2,n+1))-\pi(f(2,n))$	2-SGp in the interval $[f(2,n), f(2,n+1)]$	$\alpha_3(2,n)$ $\int_{f(2,n)}^{f(2,n+1)} \frac{dt}{\log t}$	$\alpha_4(2,n)$ $\int_{f(2,n)}^{f(2,n+1)} \frac{dt}{\log t}$	$\alpha_5(2,n)$ $\int_{f(2,n)}^{f(2,n+1)} \frac{dt}{\log t}$
1	2	2	–	2.2	2.5
2	2	2	2.8	3	2.9
3	7	5	5.9	5.9	6.3
4	15	14	14.1	13.8	13.9
5	42	36	37	36	36
6	124	104	104	100	102
7	372	295	307	296	297
8	1144	895	941	906	892
9	3647	2970	2974	2864	2874
10	11861	9496	9621	9272	9284
11	39258	30788	31741	30617	30197
12	132119	104997	106446	102770	102823
13	450453	353357	361946	349778	346471
14	1553274	1211233	1245390	1204660	1185604
15	5411233	4304679	4329368	4191628	4233170
16	19015798	14953724	15185847	14715731	14712906
17	67343702	52508562	53689079	52071171	51643313
18	240139092	188600098	191152371	185542307	185655903
19	861585192	671209186	684850472	665264341	660813007

But we should acknowledge that in some similar problems the data agrees well with heuristics only when the numbers are small. Therefore, we tried to test the accuracy of $\alpha_3(2,n)$, $\alpha_4(2,n)$ and $\alpha_5(2,n)$ when dealing with larger numbers. Surprisingly, their accuracy was very acceptable even if we are dealing with 100, 1000 or 10000 digits numbers, check Table 3.

Table 3. Some counts and asymptotics of 2-SGp out of some 100, 1000, and 10000 digits consecutive primes

n	L consecutive primes in $[f(2,n), f(2,n+1)]$	2-SGp out of L consecutive primes in $[f(2,n), f(2,n+1)]$	$L\alpha_3(2,n)$	$L\alpha_4(2,n)$	$L\alpha_5(2,n)$
169	100000	76832	76986	76460	76346
1666	10000	7669	7646	7638	7636
16617	100	77	76	76	76

The computations in Tables 2 and 3 show that the ratio of 2-SGp to primes slowly converges to 0.76. Also our computations of $\alpha_3(2,n)$, $\alpha_4(2,n)$ and $\alpha_5(2,n)$ for some large values of n show that they converge to 0.76. Which means that the probability that a given prime p is 2-SGp converges to 0.76 as well (as $p \to \infty$). Therefore, according to our computations of $\alpha_3(2,n)$, $\alpha_4(2,n)$ and $\alpha_5(2,n)$ and supported by the computations in Tables 2 and 3, we believe that it is reasonable to give the following conjecture.

Conjecture 4.1. *The probability that a given prime p is 2-SGp converges to 0.76 as $p \to \infty$.*

We also computed $\alpha_3(a,n)$, $\alpha_4(a,n)$ and $\alpha_5(a,n)$ for some large values of n with many values of a (other than $a=2$), and according to these computations and supported by the computations in Table 4 we give a general conjecture about the probability that a given prime p is a-SGp.

Conjecture 4.2. *The probability that a given prime p is a-SGp converges to*

$$\lim_{n\to\infty} \alpha_3(a,n) \approx \lim_{n\to\infty} \alpha_4(a,n) \approx \lim_{n\to\infty} \alpha_5(a,n)$$

as $p \to \infty$.

Table 4. Some counts and asymptotics of a-SGp out of 10^5 100 digits consecutive primes

(a,n)	L consecutive primes in $[f(a,n),f(a,n+1)]$	a-SGp out of L consecutive primes in $[f(a,n), f(a,n+1)]$	$L\alpha_3(a,n)$	$L\alpha_4(a,n)$	$L\alpha_5(a,n)$
$(2,169)$	100000	76832	76986	76460	76346
$(3,106)$	100000	59943	60230	59727	59555
$(4,84)$	100000	51493	51724	51276	50998
$(5,72)$	100000	46186	46512	46104	45798
$(6,65)$	100000	42743	42932	42557	42077
$(7,59)$	100000	39449	40288	39934	39239
$(8,56)$	100000	37924	38236	37906	37316
$(9,53)$	100000	36224	36585	36271	35606
$(10,50)$	100000	34702	35217	34915	34332

To show the utility of Conjecture 4.1 we considered the question of which of the Mersenne primes (i.e., those of the form 2^p-1 where p itself is prime) are also 2-SGp. Either $\alpha_3(2,n)$, $\alpha_4(2,n)$ or $\alpha_5(2,n)$ (or according to Conjecture 4.1) can tell us that there are about 23 2-SGp out of the first 30 Mersenne primes. If that were true, then we would be able to generate 23 2-Sp at least. But out of the first 30 Mersenne primes there are exactly 21 2-SGp. That shows again that the accuracy of either $\alpha_3(2,n)$, $\alpha_4(2,n)$ or $\alpha_5(2,n)$ (or Conjecture 4.1) is acceptable (even if we are dealing with completely random numbers such as Mersenne primes). We also can use them as a method to search for huge 2-Sp or a-Sp in general. The following table lists all of the 21 Mersenne primes which are also 2-SGp as well as the corresponding values of $2K$ for each one (for some of them, more than one value of K works).

Table 5. Mersenne primes which are also 2-Sophie Germain primes, and the corresponding values of K

$N = 2K \times (2^p - 1) + 1$	digits
$2 \times (2^2 - 1) + 1$	1
$4 \times (2^3 - 1) + 1$	2
$4 \times (2^7 - 1) + 1$	3
$12 \times (2^{17} - 1) + 1$	7
$16 \times (2^{19} - 1) + 1$	7
$52 \times (2^{61} - 1) + 1$	21
$66 \times (2^{61} - 1) + 1$	21
$114 \times (2^{127} - 1) + 1$	41
$124 \times (2^{127} - 1) + 1$	41
$336 \times (2^{521} - 1) + 1$	160
$154 \times (2^{607} - 1) + 1$	185
$550 \times (2^{607} - 1) + 1$	186
$156 \times (2^{2203} - 1) + 1$	666
$546 \times (2^{2203} - 1) + 1$	666
$1110 \times (2^{2203} - 1) + 1$	667
$1144 \times (2^{2203} - 1) + 1$	667
$1086 \times (2^{2281} - 1) + 1$	690
$1656 \times (2^{2281} - 1) + 1$	690
$1816 \times (2^{3217} - 1) + 1$	972

$N = 2K \times (2^p - 1) + 1$	digits
$2010 \times (2^{4253} - 1) + 1$	1284
$3708 \times (2^{4253} - 1) + 1$	1284
$1746 \times (2^{9689} - 1) + 1$	2920
$3426 \times (2^{9941} - 1) + 1$	2997
$3696 \times (2^{9941} - 1) + 1$	2997
$6238 \times (2^{11213} - 1) + 1$	3380
$9048 \times (2^{11213} - 1) + 1$	3380
$858 \times (2^{21701} - 1) + 1$	6536
$14712 \times (2^{21701} - 1) + 1$	6537
$4018 \times (2^{23209} - 1) + 1$	6991
$20808 \times (2^{23209} - 1) + 1$	6991
$17262 \times (2^{44497} - 1) + 1$	13400
$15418 \times (2^{86243} - 1) + 1$	25966
$42844 \times (2^{86243} - 1) + 1$	25967
$58818 \times (2^{86243} - 1) + 1$	25967
$6526 \times (2^{132049} - 1) + 1$	39755
$30690 \times (2^{132049} - 1) + 1$	39756
$47142 \times (2^{132049} - 1) + 1$	39756
$110086 \times (2^{132049} - 1) + 1$	39756

We also considered the question of what is the probability that the largest known prime $p = 2^{82589933} - 1$ is 2-SGp. This probability can be found by computing either $\alpha_3(2, n)$, $\alpha_4(2, n)$ or $\alpha_5(2, n)$ for $n = 41294979$. Hence, the probability that $p = 2^{82589933} - 1$ is 2-SGp is about $\alpha_3(2, 41294979) \approx \alpha_4(2, 41294979) \approx \alpha_5(2, 41294979) \approx 0.7637$. Which means that the probability to find a prime number of the form $N = 2K(2^{82589933} - 1) + 1$, where $1 \leq K \leq 41294979$ is about 0.7637 .

Acknowledgements

I am grateful to my supervisors Prof. Ahmed El-Guindy of Cairo University and Prof. Hatem M. Bahig of Ain Shams University for all their help and guidance that they have given me. Also I would like to thank Prof. Wafik Lotfallah of the American University in Cairo and Ms. Hagar Gamal of Cairo University for several useful discussions.

Also I would like to express my thanks to the editors and anonymous reviewers for their feedback and helpful comments that led to the improvement of this manuscript.

References

[1] Caldwell, C. K. (1997). *An amazing prime heuristic*. Preprint. Available online at: `https://arxiv.org/abs/2103.04483v1`.

[2] Grau, J. M., Oller-Marcén, A. M., & Sadornil, D. (2015). A primality test for $Kp^n + 1$ numbers, *Mathematics of Computation*, 84, 505–512.

[3] Hardy, G. H., & Littlewood J. E. (1923). Some problems of 'partitio numerorum' III. On the expression of a number as a sum of primes. *Acta Mathematica*, 44, 1–70.

[4] Korevaar, J. (2012). The prime pair conjecture of Hardy and Littlewood. *Indagationes Mathematicae*, 23, 269–299.

[5] Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1997). *Handbook of Applied Cryptography*. CRC Press, Boca Raton, FL.

[6] Ribenboim, P. (1996). *The New Book Of Prime Number Records*, 3rd ed. Springer, New York.