MATHEMATICS OF COMPUTATION
Volume 65, Number 213
January 1996, Pages 393–396

LARGE SOPHIE GERMAIN PRIMES

HARVEY DUBNER

ABSTRACT. If P is a prime and $2P+1$ is also prime, then P is a Sophie Germain prime. In this article several new Sophie Germain primes are reported, which are the largest known at this time. The search method and the expected search times are discussed.

1. INTRODUCTION

If P is a prime and $Q = 2P + 1$ is also a prime, then P is called a Sophie Germain prime. Sophie Germain, around 1825, showed that the first case of Fermat's Last Theorem is true for these primes, P, [5, p. 261]. Euler proved that if P is of the form $4k + 3$, then Q divides the Mersenne number, $M_p = 2^p - 1$ [4]. Thus, large Sophie Germain primes of the form $4k + 3$ lead to the largest known composite Mersenne numbers. It has not yet been shown that there are infinitely many Mersenne composites. It is generally believed that there are an infinite number of Sophie Germain primes (Germains) but this has not been proved. In fact there is good reason to believe that there are about the same number of Germains as twin primes [6, Ch. 3].

In this article we report several new Germains, which are the largest known at this time. We also discuss the search method and the expected search times.

2. METHOD

The numbers considered are,

$$P = c * 3003 * 10^b - 1, \qquad c = 1, 2, 3 \ldots, \tag{1}$$

$$Q = 2 * P + 1, \tag{2}$$

$$R = \frac{P-1}{2}. \tag{3}$$

These forms have the following advantages:

a. P, Q and R are always of the form $6k \pm 1$, so that it is easy to verify primality by using the methods of [1], even for very large numbers.

b. The sequence $6k - 1, 12k - 1, 24k - 1, \ldots, 6k * 2^n - 1$ has no theoretical limit to the number of possible consecutive primes, a useful consideration for possible future research.

c. P, Q, and R are never divisible by a prime less than 17. This means the bit array representing c is used efficiently in the sieving process.

Received by the editor March 7, 1994 and, in revised form, April 26, 1994.
1991 *Mathematics Subject Classification*. Primary 11A41.

d. If P and Q are prime, then P is a Germain. If P and R are prime, then R is a Germain. Because of this, the average time to find a Germain is halved.
e. Incidently, P, Q and R are always of the form $4k+3$.

First, the array representing c is sieved to eliminate any c for which P has a factor less than some maximum, pmax. Next, the same array is sieved to eliminate any c for which Q has a factor less than pmax. Then the same array is sieved to eliminate any c for which R has a factor less than pmax. For the remaining values of c, P is tested for probable primality using a Fermat test. If P is a probable prime (PRP), then Q and R are tested. If either is PRP, then a Germain has almost certainly been found.

To illustrate, consider searching for Germains near 2000 digits ($b = 2000$). Starting with a c-array of 4,000,000 bits, after sieving for P, Q and R with pmax of 30,000,000, approximately 18,000 c's remain. Each remaining c generates a P, Q and R all of which are prime candidates. The sieving process takes about three hours. A Fermat test on the surviving 18,000 P's takes about 100 hours. Note that by spending about 1% of the total test time to sieve R, the total test time to find a Germain is halved.

The primes are verified using methods from [1]. The number $(P+2)$ is also tested since by accident a twin prime might be found. Without sieving for a twin, and considering that the chance of finding a Germain has been doubled, finding a twin is about 14 times less likely than finding a Germain, but it costs almost nothing to try. Since the largest known twin prime has 4030 digits, the slight additional effort was certainly worthwhile, although no twins were discovered.

3. Test times

The Prime Number Theorem states that if N is a random number, $1/\log(N)$ is the probability that it is prime. Thus, on average about $\log(N)$ numbers in the vicinity of N must be tested to find a prime. By eliminating numbers that have "small" divisors (sieving) fewer numbers need to be tested. When sieved up to pmax, the average number that require testing is

$$(4) \qquad T(N,\text{pmax}) = \log(N) * \prod_{2}^{\text{pmax}} \left(\frac{p-1}{p}\right), \quad p = \text{prime}.$$

By Mertens's theorem [4],

$$(5) \qquad \prod_{2}^{\text{pmax}} \left(\frac{p-1}{p}\right) \approx \frac{e^{-\gamma}}{\log(\text{pmax})} \approx \frac{0.5616}{\log(\text{pmax})},$$

where γ is Euler's constant $(0.5772\ldots)$. Even for pmax as low as 101, (5) is accurate to about 2%. At pmax = 10007, the accuracy is about .1%. Combining (4) and (5), we get

$$(6) \qquad T(N,\text{pmax}) = 0.5615 * \frac{\log(N)}{\log(\text{pmax})}.$$

Equation (4) holds even when N is not random but cannot possibly be divisible by a particular prime, $p1$, because the change in the probability of N being prime is exactly compensated by the necessary omission of the term, $(p1-1)/p1$, in the product associated with sieving. Thus, (6) is a useful general equation, which often gives the average number of tests required to find a prime when sieving is used.

TABLE 1. Expected time to find Sophie Germain Primes Sieving up to pmax $= 30,000,000$

size of N	Average tests	test time	total time
digits	required	seconds	days
1000	2,805	5.0	.17
1500	6,310	11.4	.83
2000	11,219	23.1	3.39
2500	17,529	37.4	7.59
3000	25,242	56.9	16.6
3500	34,357	83.5	33.2
4000	44,874	114	59.1
4500	56,794	160	105
5000	70,116	203	165

It is applicable to many forms of N, but the ratio of the logs must be large. In fact, the study of (6) is quite interesting and important, but its subtleties are not numerically significant for this paper because of the large primes involved.

After a P is found to be PRP, Q and R are also tested. If either is PRP, then a Germain has almost certainly been found. Thus, the average number of P's that must be tested to find a Germain is approximately

$$T_G(N) = \frac{T(N, \text{pmax}) * T(N, \text{pmax})}{2}. \tag{7}$$

Table 1 shows the expected times to find a Germain after sieving up to pmax $=$ $30,000,000$, a reasonably typical value. Column 2, Average tests required, is independent of the computing hardware. Column 3, the test times, are for a PC 486/33 with a special-purpose Cruncher plug-in board [2], the hardware which was used to find the new large Germains.

4. RESULTS

Prior to this study the size of the largest known Sophie Germain prime was 2038 digits [3]. The availability of six PC Cruncher systems made it feasible to search for considerably larger Germains based on the time estimates from Table 1. The results of this search are shown in Table 2.

TABLE 2. New large Sophie Germain primes $P = c * 3003 * 10^b - 1$

		number of	discovery
c	b	digits	date
7014	2110	2118	Nov. 1993
581436	2581	2591	Dec. 1993
15655515	2999	3010	Dec. 1993
5199545	3529	3540	Jan. 1994
488964	4003	4013	Jan. 1994
1803301	4526	4536	Jan. 1994

The actual search times were reasonably close to the estimated times. For example, the time to find the 4526-digit Germain took about 110 Cruncher-days as compared to the expected time of 105 days. Although we tried for several weeks, we did not find a 5000-digit Germain, and the search was terminated.

The largest known composite Mersenne number is now a number with more than 1.63×10^{4535} digits.

References

1. J. Brillhart, D. H. Lehmer, and J. L. Selfridge, *New primality criteria and factorizations of* $2^m \pm 1$, Math. Comp. **29** (1975), 620–647. MR **52:**5546
2. C. Caldwell, *Review of the Cruncher PC plug-in board*, J. Recreational Math. **25** (1993), 56–57.
3. ———, *The largest known primes*, updated regularly and is available on request. Tel. (901) 587–7360, E-mail: caldwell@UTmartn.bitnet
4. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford Univ. Press, New York, 1979. MR **81i:**10002
5. P. Ribenboim, *The book of prime number records*, 2nd ed., Springer-Verlag, New York, 1989. MR **90g:**11127
6. H. Riesel, *Prime numbers and computer methods for factorization*, Birkhäuser, Boston, 1958.

449 Beverly Road, Ridgewood, New Jersey 07450
E-mail address: 70372.1170@compuserve.com