Prime Numbers with a Fixed Number of One Bits or Zero Bits in Their Binary Representation

Samuel S. Wagstaff, Jr.

CONTENTS

This work was supported in part by grants from the National Science Foundation, from the CERIAS Center at Purdue University and from the Lilly Endowment Inc.

Keywords: prime numbers, binary representation

We study the distribution of prime numbers that have a given number of one bits in their binary representation, and of those that have a given number of zero bits. We consider basic questions such as whether there are infinitely many of them, and explain their distribution in residue classes modulo small primes. We prove the unexpected result that, for $m \geq 3$, there is no prime number with precisely 2^m bits, exactly two of which are zero bits.

1. INTRODUCTION

Several authors [Gel'fond 1967/68; Bésineau 1971; Olivier 1971; Fouvry and Mauduit 1996; Dartyge and Mauduit 2000] have studied the binary representation of prime numbers. It is claimed in [Olivier 1971] that the number of primes $\leq x$ having an even number of one bits is asymptotic to $\pi(x)/2$. Although this statement is likely true, Montgomery [1994, Item 67, p. 208] suggests that Olivier's proof is incomplete.

For natural numbers k, let P_k denote the set of all primes with exactly k one bits in their binary representation. Then $P_1 = \{2\}$ and P_2 is the set of all Fermat primes $F_m = 2^{2^m} + 1$. Hardy and Wright [1960] have argued that P_2 is probably finite. It is well known that F_m is prime for $0 \leq m \leq 4$. It is known [Brillhart et al. 1988] that F_m is composite for $5 \leq m < 30$ and for scores of larger m. We will give a heuristic argument that P_k is infinite for each $k \geq 3$. This argument is supported by counts of the primes in P_k for $k = 3$ and 4 given in Tables 1 and 2 (page 269).

One motivation for this work was to discover new prime divisors of F_m. The factor 2424833 of F_9 has only four one bits. (Of course, any factor of F_m must be congruent to 1 modulo 2^{m+2}; see [Brillhart et al.

1058-6458/2001 $0.50 per page

1988]. Thus 10 of the 21 bits of 2424833 must be 0. Many primes of the form

$$t \cdot 2^n + 1$$

with $n > 0$ and small t are known. See [Baillie 1979] and its references. These primes have one more one bit than t does.) This made us wonder whether primes with few one bits are more likely than primes with many one bits to divide Fermat numbers. We tested each prime we found for dividing the possible Fermat numbers. Unfortunately, no new prime divisors of F_m were found.

Additional motivation derives from work of Joel Brenner (personal communication, 1985) and others on the word problem for finite groups. A word $x^a y^b \dots$ from a free group *covers* a given group G if every element of G can be represented by the word when its variables are replaced by suitable elements of G. Suppose G is the symmetric group on s letters and the word is $x^a y^b$. If a and b are both even integers, then the word does not cover G. If a and b are both odd, then the word does cover G. The remaining case with one of a, b even and one odd could be settled if one knew there were infinitely many primes of one of the forms $2^n \pm 1$, $2^n + 2^i \pm 1$. We present some evidence supporting the existence of infinitely many primes $2^n + 2^i + 1$.

Some might think that another use of primes with few one bits would be in the Pohlig–Hellman cryptosystem [1978]. For some reason, a few cryptographers use prime numbers as the secret exponents in this system. The fast exponentiation algorithm used in enciphering and deciphering runs faster when the exponent has few one bits. However, it would be silly to use a prime number with only three or four one bits for this purpose, as it would be easy to guess. Furthermore, there is no obvious advantage to using a prime exponent. The exponent just has to be random.

We also study primes with a fixed number of zero bits. For $k \geq 0$, let Q_k denote the set of all primes having exactly k (non-leading) zero bits in their binary representation. Then Q_0 is the set of all Mersenne primes $2^p - 1$, which is probably an infinite set. Indeed, probably Q_k is infinite for each $k \geq 0$. We support this claim with a heuristic argument and counts of the primes in Q_k for $k = 1$ and 2 given in Tables 3 and 4.

2. SIMPLE HEURISTICS

For natural numbers n and $2 \leq k \leq n+1$ let $A_k(n)$ denote the cardinality of $P_k \cap [2^n, 2^{n+1}]$, that is, the number of primes between 2^n and 2^{n+1} having exactly k one bits.

By the prime number theorem, odd integers near x are prime with probability about $2/(\ln x)$. Assume that odd numbers with exactly k one bits and the same length have this same chance of being prime. Then odd numbers between 2^n and 2^{n+1} with exactly k one bits have probability about $2/(\ln(2^n))$ of being prime. There are $\binom{n-1}{k-2}$ odd numbers between 2^n and 2^{n+1} with exactly k one bits. Thus we have the estimate

$$A_k(n) \approx \binom{n-1}{k-2} \frac{2}{n \ln 2}.$$

For small values of k the estimates are

$$A_2(n) \approx \frac{2}{n \ln 2},$$

$$A_3(n) \approx \frac{(n-1)2}{n \ln 2} \approx \frac{2}{\ln 2} \approx 2.89,$$

$$A_4(n) \approx \frac{(n-1)(n-2)}{n \ln 2} \approx \frac{n}{\ln 2} \approx 1.44n.$$

For $n \geq 1$ and $0 \leq k \leq n$, let $B_k(n)$ denote the cardinality of $Q_k \cap [2^{n-1}, 2^n)$. There are $\binom{n-2}{k}$ odd numbers with n non-leading bits of which exactly k are zero bits. Each has probability about $2/(\ln x)$ of being prime. This leads to the estimate

$$B_k(n) \approx \binom{n-2}{k} \frac{2}{n \ln 2}.$$

For small k the estimates are

$$B_0(n) \approx \frac{2}{n \ln 2},$$

$$B_1(n) \approx \frac{(n-2)2}{n \ln 2} \approx \frac{2}{\ln 2} \approx 2.89,$$

$$B_2(n) \approx \frac{(n-2)(n-3)}{n \ln 2} \approx \frac{n}{\ln 2} \approx 1.44n.$$

Note the similarity of these estimates to those for $A_k(n)$ above. In fact, we expect that

$$B_k(n) \approx A_{k+2}(n)$$

for each $k > 0$ and for large n.

3. EMPIRICAL DATA

As a result of the work factoring Fermat numbers and testing them for primality [Brillhart et al. 1988], it is known that $A_2(n) = 1$ if $n = 2^m$ for $0 \leq m \leq 4$ and that $A_2(n) = 0$ for all other $n < 2^{30}$.

We computed $A_3(n)$ for $2 \leq n \leq 200$ and $A_4(n)$ for $3 \leq n \leq 100$.

The average value of $A_3(n)$ for $2 \leq n \leq 200$ is 2.97, which is higher than the predicted 2.89. The running average $\sum_{n=2}^{x} A_3(n)/(x-1)$ increases steadily over the range of Table 1: It has values 2.63, 2.80, 2.85, 2.97 at $x = 50, 100, 150, 200$. The results suggest that if $A_3(n)$ has an average, it exceeds 3.

n mod 10	0	1	2	3	4	5	6	7	8	9
$\lfloor n/10 \rfloor$			1	2	1	2	3	3	0	4
1	2	3	2	2	2	4	1	3	4	5
2	3	2	1	5	1	0	2	5	2	2
3	8	6	0	5	3	4	2	3	2	2
4	0	3	5	0	1	5	3	7	0	1
5	2	5	1	5	2	6	0	6	0	2
6	3	2	1	2	0	2	3	5	3	6
7	2	2	2	5	2	7	1	3	2	3
8	1	6	2	4	3	3	2	6	1	1
9	5	7	2	4	2	5	0	3	4	3
10	1	2	1	3	0	5	4	6	3	1
11	2	3	0	7	8	1	1	5	2	5
12	0	2	1	2	1	4	4	6	0	4
13	2	4	2	1	0	7	2	7	2	1
14	0	5	1	7	1	0	3	8	2	4
15	5	7	0	10	5	2	1	3	2	6
16	0	6	4	3	2	5	5	3	1	2
17	2	4	3	2	6	10	0	4	1	1
18	4	1	2	7	1	0	8	2	4	3
19	3	4	0	7	3	5	1	3	2	5
20	1									

TABLE 1. $A_3(n)$ for $2 \leq n \leq 200$.

The average value of $A_4(n)/n$ for $3 \leq n \leq 100$ is 1.34, which is lower than the predicted 1.44. The running average of $A_4(n)/n$ increases steadily over the range of Table 2. It has values 1.06, 1.22, 1.28, 1.34 at $x = 25, 50, 75, 100$. It is plausible that it might have a limit near 1.44.

The Mersenne primes have been studied extensively. At present, it is known that $2^p - 1$ is prime for these 38 primes p: 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, 23209, 44497, 86243, 110503, 132049, 216091, 756839, 859433, 1257787, 1398269, 2976221, 3021377 and 6972593. The search for Mersenne primes is being conducted by GIMPS; at http://www.mersenne.org/prime.htm one can see the latest results. At this writing, all exponents p up to about $5 \cdot 10^6$ have been tested.

	0	1	2	3	4	5	6	7	8	9
				0	2	2	5	4	10	6
1	13	11	9	16	16	18	25	15	19	15
2	37	17	37	29	29	32	40	23	49	31
3	51	39	37	30	52	46	40	42	62	43
4	57	42	68	52	78	60	89	54	63	59
5	92	58	79	82	99	73	87	47	99	74
6	72	81	106	56	102	85	117	85	97	64
7	132	93	117	93	117	102	120	101	118	104
8	141	97	157	91	115	113	158	97	152	109
9	187	120	152	83	177	141	118	125	200	127
10	176									

TABLE 2. $A_4(n)$ for $3 \leq n \leq 100$.

Now $B_0(p)$ is 1 if $2^p - 1$ is a Mersenne prime and 0 otherwise. Thus, $B_0(n)$ has been computed for all n up to about $5 \cdot 10^6$, far beyond any calculation in this paper. This has been possible because the Lucas-Lehmer test [Hardy and Wright 1960] for primality of $2^p - 1$ is so swift, and because much more computer time has been devoted to the GIMPS project.

We computed $B_1(n)$ for $1 \leq n \leq 200$ and $B_2(n)$ for $2 \leq n \leq 100$.

The average value of $B_1(n)$ for $1 \leq n \leq 200$ is 3.53, which is higher than the predicted 2.89. The running average

$$\sum_{n=1}^{x} \frac{B_1(n)}{x}$$

increases steadily over the range of Table 3: It has values 3.14, 3.27, 3.55, 3.53 at $x = 50, 100, 150, 200$. The results suggest that if $B_1(n)$ has an average, it exceeds 3.

The average value of $B_2(n)/n$ for $2 \leq n \leq 100$ is 1.08, which is lower than the predicted 1.44. The running average of $B_2(n)/n$ increases steadily over the range of Table 4. It has values 0.78, 0.94, 1.05, 1.08 at $x = 25, 50, 75, 100$. It is plausible that it might have a limit near 1.44.

It was easy to prove the primality of the numbers in these tables which were identified as "probably prime" by a variation of Fermat's theorem. We used

n mod 10	0	1	2	3	4	5	6	7	8	9
$\lfloor n/10 \rfloor$		0	1	1	2	2	3	0	4	4
1	3	1	5	1	4	0	3	2	8	1
2	11	4	5	0	7	1	2	0	1	5
3	0	0	7	5	1	1	9	0	6	0
4	7	1	6	0	4	7	2	1	10	3
5	3	1	2	1	6	0	4	3	0	1
6	8	3	3	0	3	1	8	1	2	2
7	3	0	9	1	5	2	5	8	3	0
8	10	3	0	2	4	4	6	1	4	4
9	9	0	2	3	9	0	6	2	6	5
10	5	2	7	1	7	4	1	2	7	1
11	8	0	8	1	6	3	4	7	4	1
12	6	4	8	0	5	3	5	0	4	3
13	10	2	5	0	4	0	6	4	11	0
14	9	3	3	0	14	6	4	2	0	3
15	7	0	5	7	1	2	3	1	9	0
16	4	3	5	1	4	4	5	0	17	0
17	7	2	0	0	6	0	3	2	2	0
18	9	4	2	0	4	6	4	1	8	5
19	2	0	9	2	7	0	1	4	5	0
20	6									

TABLE 3. $B_1(n)$ for $1 \le n \le 200$.

	0	1	2	3	4	5	6	7	8	9
			0	0	0	1	2	4	0	9
1	5	14	4	16	9	18	0	21	21	21
2	7	41	22	31	5	37	20	33	14	37
3	0	47	0	69	31	36	34	55	34	71
4	10	60	50	69	22	81	52	59	5	97
5	71	79	42	67	86	95	13	103	61	81
6	47	98	50	110	0	108	87	116	36	125
7	98	98	29	126	90	125	46	107	100	125
8	8	158	81	109	65	156	94	131	27	127
9	144	146	38	167	129	137	6	127	112	178
10	76									

TABLE 4. $B_2(n)$ for $2 \le n \le 100$.

the methods of [Brillhart et al. 1988]. The numbers counted in Table 1 were especially easy to prove prime, for if $p = 2^n + 2^i + 1$, then $p - 1 = 2^i(2^{n-i} + 1)$, and the complete factorization of $(2^{n-i} + 1)$ is available in [Brillhart et al. 1988]. Likewise, the numbers counted in Table 3 were easy to prove prime, for if $p = 2^n - 2^i - 1$, then $p + 1 = 2^i(2^{n-i} - 1)$, and $(2^{n-i} - 1)$ is factored in [Brillhart et al. 1988].

4. DEEPER HEURISTICS

We first consider primes having a fixed number of one bits. It is known [Hardy and Wright 1960] that $2^n + 1$ must be composite unless $n = 2^m$. (The proof exhibits a factor of $2^n + 1$ when $n \neq 2^m$.) Thus [Hardy and Wright 1960] the expected number of Fermat primes is

$$\sum_{m=0}^{\infty} A_2(2^m) \approx \sum_{m=0}^{\infty} \frac{2}{2^m \ln 2} = \frac{4}{\ln 2} \approx 5.77 < \infty$$

rather than

$$\sum_{n=1}^{\infty} \frac{2}{\ln(2^n + 1)} \approx \sum_{n=1}^{\infty} \frac{2}{n \ln 2} = \infty.$$

On the other hand, every divisor of F_m is congruent to 1 modulo 2^{m+2}. This restriction on possible divisors might seem to increase F_m's probability of being prime, but in fact the possible divisors divide the F_m's with a higher than expected probability, which just compensates for their reduced number. Do similar considerations influence the primality chances of odd integers with exactly k one bits when $k > 2$? Can we exhibit factors of some of them? Are their possible divisors restricted?

Consider the case $k = 3$. It is easy to see that the prime 3 divides $2^n + 2^i + 1$ if and only if both n and i are even. Thus, 3 divides 1/4 of the odd numbers with exactly 3 one bits. However, for fixed n, either 3 divides $2^n + 2^i + 1$ for half of the i's (when n is even) or 3 divides $2^n + 2^i + 1$ for no i (when n is odd). Thus, $2^n + 2^i + 1$ is more likely to be a multiple of 3, and hence less likely to be prime, when n is even than when n is odd. This difference is easy to observe in Table 1, since

$$\sum_{\substack{n=2 \\ n \text{ even}}}^{200} A_3(n) = 204, \quad \text{while} \quad \sum_{\substack{n=3 \\ n \text{ odd}}}^{199} A_3(n) = 386.$$

Similar analysis of divisibility by 5 and 7 of odd numbers with 3 one bits suggests that $A_3(n)$ should be larger when $n \equiv 2 \pmod 4$ than when $n \equiv 0 \pmod 4$ and that $A_3(n)$ should be larger when 3 divides n than when n lies in one of the other two residues classes modulo 3. Table 1 supports these observations. The sums of $A_3(n)$ with n in a fixed residue class modulo 4 are 61, 171, 143, 215, for classes 0, 1, 2, 3 (mod 4). The sums of $A_3(n)$ with n in a fixed residue class modulo 3 are 254, 165, 171, for classes 0, 1, 2 (mod 3). This analysis (explained below) also concludes that 5 divides 3/16 of the odd numbers with 3 one bits and that 7 divides 2/9 of them.

We now argue heuristically that the fraction of odd numbers with 3 one bits divisible by an odd prime p is on average $1/(p-1)$. If the set of all odd numbers with 3 one bits were sufficiently dense (an arithmetic progression, for example), then we could show by a sieve argument that about $c/(\ln x)$ of these odd numbers $\leq x$ would be prime. This conclusion (with $c = 2$) was our starting point in Section 2.

Let p be an odd prime. Let $l_2(p)$ denote the least $l > 0$ for which $2^l \equiv 1 \pmod p$. Then $l_2(p)$ divides $p-1$ and the function $f(n) = 2^n \bmod p$ is periodic with period $l_2(p)$. The set

$$A = \{2^n \bmod p : 0 \leq n < l_2(p)\}$$

is a subset of $\{1, 2, \ldots, p-1\}$ of size $l_2(p)$. The set

$$B = \{p-1 - (2^i \bmod p) : 0 \leq i < l_2(p)\}$$

is a subset of $\{0, 1, \ldots, p-2\}$ of size $l_2(p)$. There is a one-to-one correspondence between solutions to $2^n + 2^i + 1 \equiv 0 \pmod p$ and elements of $A \cap B$. Each element of A has probability about (size of $B)/(p-1) = l_2(p)/(p-1)$ of also being in B. Assuming independent probabilities, the expected size of $A \cap B$ is $(l_2(p))^2/(p-1)$. Thus, p divides $2^n + 2^i + 1$ with probability

$$\frac{\text{size of } A \cap B}{\text{number of pairs } (n, i)} = \frac{(l_2(p))^2/(p-1)}{(l_2(p))^2} = \frac{1}{p-1}.$$

Note that when 2 is a primitive root for p, $l_2(p) = p-1$, $A = \{1, 2, \ldots, p-1\}$, $B = \{0, 1, \ldots, p-2\}$, and so p divides $2^n + 2^i + 1$ with probability

$$\frac{p-2}{(p-1)^2} \approx \frac{1}{p}.$$

There is a heuristic argument [Lehmer and Lehmer 1962; Hooley 1967] that concludes that 2 is a primitive root for about 37 percent of all primes.

The analysis above consisted of counting the solutions to $2^n + 2^i + 1 \equiv 0 \pmod p$ with $0 \leq n < l_2(p)$ and $0 \leq i < l_2(p)$. For example, for $p = 5$, $l_2(p) = 4$, and the congruence has the three solutions $(n, i) = (0, 3)$, $(1, 1)$ and $(3, 0)$. Thus, 5 divides 3/16 of the odd numbers with three one bits. Similarly, 7 divides 2/9 of the odd numbers with three one bits. There are some primes for which the congruence has no solution, for example, the Mersenne primes larger than 7. In fact, much more is true:

Theorem 1. *No positive integer multiple of* 2^k-1 *has fewer than* k *one bits.*

This theorem is a special case of [Stolarsky 1980, Theorem 2.1].

The theorem restricts the combinations of primes which may divide $2^n + 2^i + 1$: Suppose q is a prime divisor of 2^k-1 for some $k > 3$. Then q may divide $2^n + 2^i + 1$, but not in combination with the factor $(2^k-1)/q$ because 2^k-1 cannot divide $2^n + 2^i + 1$ by the theorem.

Now consider the case $k = 4$. When does 3 divide $2^n + 2^i + 2^j + 1$? The parity of j needed for this to happen depends on those of n and i as follows:

$n \bmod 2$	0	0	1	1
$i \bmod 2$	0	1	0	1
$2^n \bmod 3$	1	1	2	2
$2^i \bmod 3$	1	2	1	2
$(2^n + 2^i + 1) \bmod 3$	0	1	1	2
$2^j \bmod 3$	–	2	2	1
$j \bmod 2$	–	1	1	0

When n is even, both i and j must be odd, and when n is odd, i and j must have different parity, in order for 3 to divide $2^n + 2^i + 2^j + 1$. This means that $2^n + 2^i + 2^j + 1$ is twice as likely to be divisible by 3, and hence less likely to be prime, when n is odd than when n is even. This effect is easy to notice in Table 2, since

$$\sum_{\substack{n=4 \\ n \text{ even}}}^{100} A_4(n) = 4175 \quad \text{while} \quad \sum_{\substack{n=3 \\ n \text{ odd}}}^{99} A_4(n) = 2987.$$

We now consider primes having a fixed number of zero bits, beginning with the case of no zero bits. The simple heuristics predict that the probability that $M_p = 2^p-1$ is prime is about $2/(p \ln 2)$. We explained in [Wagstaff 1983] why the constant $2/\ln 2 \approx 2.89$ should be replaced with $e^\gamma/\ln 2 \approx 2.57$. The same two constants appear in heuristic estimates for the number of Mersenne primes $\leq x$. Simple heuristic arguments suggest that the number $M(x)$ of Mersenne primes $\leq x$ should be about

$$(2/\ln 2) \ln \ln x,$$

while deeper analysis predicts the number should be about

$$(e^\gamma/\ln 2)\ln\ln x.$$

We compared these two predictions in a table [Wagstaff 1983, p. 388] when only 27 Mersenne primes were known. We extend that table here as Table 5. In this table, M_p is the m-th Mersenne prime. The two heuristic analyses predict different limits for the ratio $M(x)/\ln\ln x$. As x increases, this ratio decreases slowly between Mersenne primes and jumps up from $(m-1)/\ln\ln M_p$ to $m/\ln\ln M_p$ at the m-th Mersenne prime M_p. The last two columns give sliding lower and upper bounds for the limit, if any, of the ratio $M(x)/\ln\ln x$.

m	p	$\frac{m-1}{\ln\ln M_p}$	$\frac{m}{\ln\ln M_p}$
27	44497	2.52	2.61
28	86243	2.46	2.55
29	110503	2.49	2.58
30	132049	2.54	2.63
31	216091	2.52	2.60
32	756839	2.35	2.43
33	859433	2.41	2.48
34	1257787	2.41	2.49
35	1398269	2.47	2.54
36	2976221	2.41	2.48
37	3021377	2.47	2.54
38	6972593	2.40	2.47

TABLE 5. Lower and upper estimates for the ratio $M(x)/\ln\ln x$.

If the numbers in the last two columns of Table 5 converge to a limit, that limit is more likely to be near 2.57 than 2.89. The limit 2.57 seemed more plausible with the limited data in [Wagstaff 1983] than it does with the data exhibited here. In fact, the new data suggests that either the limit $e^\gamma/\ln 2 \approx 2.57$ is too large or that one or two Mersenne primes have been missed.

The function $B_1(n)$ counts primes 2^n-2^i-1 with $1 \le i \le n-2$. It is easy to see that 3 divides 2^n-2^i-1 if and only if n is odd and i is even. However, for fixed n, either 3 divides 2^n-2^i-1 for half of the i's (when n is odd) or 3 divides 2^n-2^i-1 for no i (when n is even). Thus, 2^n-2^i-1 is more likely to be a multiple of 3, and hence less likely to be prime, when n is odd than when n is even. This effect is easy to notice in Table 3, since

$$\sum_{\substack{n=2\\ n \text{ even}}}^{200} B_1(n) = 517 \quad \text{while} \quad \sum_{\substack{n=1\\ n \text{ odd}}}^{199} B_1(n) = 184.$$

Similar reasoning explains the variations in the sum of $B_1(n)$ when n ranges over different residue classes modulo small primes.

The analysis for the function $B_2(n)$ is similar to that for $A_4(n)$, with divisibility of $2^n-2^i-2^j-1$ by 3 explaining why $B_2(n)$ tends to be larger when n is odd than when n is even. But there is one surprise: In Table 4, $B_2(n)=0$ whenever n is a power of 2. One can prove that this always happens.

Theorem 2. *For all $m \ge 1$, we have $B_2(2^m)=0$. For all $m \ge 3$, we have $B_1(2^m-1)=0$. This means that, for $m \ge 3$, there is no prime number with precisely 2^m bits, exactly two of which are zero bits. In other words, there is no prime number of the form $2^{2^m}-2^i-2^j-1$, where $1 \le i < j \le 2^m-2$ and $m \ge 1$, and no prime number of the form $2^{2^m-1}-2^i-1$, where $1 \le i \le 2^m-2$ and $m \ge 3$.*

Proof. We see from Table 4 that $B_2(2^1)=B_2(2^2)=0$. Let $m \ge 3$. Let $p = 2^{2^m}-2^i-2^j-1$ with $1 \le i < j \le 2^m-1$. Note that $p = 2^{2^m-1}-2^i-1$ when $j = 2^m-1$. Write $j-i = 2^k e$, where e is odd. We will show that $d = 2^{2^k}+1$ is a proper divisor of p. Note first that $2^k \le j-i \le 2^m-2$, so $k < m$ and

$$1 < d = 2^{2^k}+1 < 2^{2^m-2}-2 < p$$

since $m \ge 3$. Clearly, $2^{2^k} \equiv -1 \pmod d$. Since $k < m$, we have $2^{2^m} \equiv 1 \pmod d$, and so d divides $2^{2^m}-1$. Write $-2^i-2^j = -2^i(2^{j-i}+1) = -2^i(2^{2^k e}+1)$. Since $2^{2^k} \equiv -1 \pmod d$ and e is odd, we have $2^{2^k e} \equiv -1 \pmod d$, and d divides $2^{2^k e}+1$. Therefore, d divides -2^i-2^j and hence also p. It follows that p is composite. $\square$

In the case $j = 2^m-1$ and $i = 2^m-2$, the proof shows that $d = 2^{2^0}+1 = 3$ is a proper divisor of $p = 2^{2^m-2}-1$ when $m \ge 3$. Of course, it is easy to prove this fact directly.

One observes that, in Table 4, $B_2(n)$ is small when n is not a power of 2, but is divisible by a high power of 2. The proof of Theorem 2 explains this phenomenon, too. Suppose $n = 2^m f$, where f is odd and > 1. Then $B_2(n)$ counts primes of the form $p = 2^{2^m f} - 2^i - 2^j - 1$ with $1 \le i < j \le 2^m f - 2$. Write $j - i = 2^k e$, where e is odd. Let $d = 2^{2^k} + 1$. Then the proof above shows that d is a proper divisor of p *provided that* $k < m$. But if $k \ge m$, then $2^{2^m} \not\equiv 1 \pmod{d}$, so d does not divide $2^{2^m} - 1$. However, d still divides $-2^i - 2^j$. (The proof above works.) Therefore, d does not divide p, and p is not prevented from being prime. We see that $B_2(2^m f)$ is small because no number of the form $p = 2^{2^m f} - 2^i - 2^j - 1$ with $1 \le i < j \le 2^m f - 2$ can be prime unless 2^m divides $j - i$. If m is large, this restriction on $j - i$ excludes most candidate p's.

5. OPEN QUESTIONS

Does there exist any k for which we can prove there are infinitely many primes with exactly k one bits? Does there exist any k for which we can prove that there are infinitely many primes with $\le k$ one bits? We conjecture that both questions have the answer, "Yes, any $k \ge 3$ will do."

One may ask the same questions with "one" replaced by "zero." It seems likely that the answers to the "zero" questions are, "Yes, for any $k \ge 0$."

ACKNOWLEDGEMENTS

The author thanks Joel Brenner and the referee for suggesting that the paper consider primes with a given number of zero bits as well as those with a given number of one bits.

REFERENCES

[Baillie 1979] R. Baillie, "New primes of the form $k2^n + 1$", *Math. Comp.* **33** (1979), 1333–1336.

[Bésineau 1971] J. Bésineau, "Sur une problème de Gel'fond relatif à la fonction 'somme des chiffres'", *C. R. Acad. Sci. Paris Sér. A-B* **272** (1971), A453–A456.

[Brillhart et al. 1988] J. Brillhart, D. H. Lehmer, J. L. Selfridge, B. Tuckerman, and S. S. Wagstaff, Jr., *Factorizations of $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ up to high powers*, Second ed., Amer. Math. Soc., Providence, Rhode Island, 1988.

[Dartyge and Mauduit 2000] C. Dartyge and C. Mauduit, "Nombres presque premiers dont l'écriture en base r ne compose pas certain chiffres", *J. Number Theory* **81** (2000), 270–291.

[Fouvry and Mauduit 1996] E. Fouvry and C. Mauduit, "Somme des chiffres et nombres presque premiers", *Math. Annalen* **305** (1996), 571–599.

[Gel'fond 1967/68] A. O. Gel'fond, "Sur les nombres qui ont des propriétés additives et multiplicatives données", *Acta Arith.* **13** (1967/68), 259–265.

[Hardy and Wright 1960] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth ed., Oxford, 1960.

[Hooley 1967] C. Hooley, "On Artin's conjecture", *J. Reine Angew. Math.* **225** (1967), 209–220.

[Lehmer and Lehmer 1962] D. H. Lehmer and E. Lehmer, "Heuristics, anyone?", pp. 202–210 in *Studies in Mathematical Analysis and Related Topics, IV, Essays in Honor of George Pólya*, Stanford Univ. Press, 1962.

[Montgomery 1994] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, Amer. Math. Soc., Providence, Rhode Island, 1994.

[Olivier 1971] M. Olivier, "Sur le développement en base g des nombres premiers", *C. R. Acad. Sci. Paris Sér. A-B* **272** (1971), A937–A939.

[Pohlig and Hellman 1978] S. Pohlig and M. Hellman, "An improved algorithm for computing logarithms over GF(p) and its cryptographic significance", *IEEE Trans. on Info. Theory* **IT-24**:1 (1978), 106–110.

[Stolarsky 1980] K. B. Stolarsky, "Integers whose multiples have anomalous digital frequencies", *Acta Arith.* **38** (1980), 117–128.

[Wagstaff 1983] S. S. Wagstaff, Jr., "Divisors of Mersenne numbers", *Math. Comp.* **40** (1983), 385–397.

Samuel S. Wagstaff, Jr., Center for Education and Research in Information Assurance and Security and Department of Computer Sciences, Purdue University, West Lafayette, IN 47907-1398, United States (ssw@cerias.purdue.edu)

Received July 3, 2000; accepted in revised from November 17, 2000