MATHEMATICS OF COMPUTATION
VOLUME 55, NUMBER 191
JULY 1990, PAGES 345–353

QUADRATIC POLYNOMIALS WHICH HAVE A HIGH DENSITY OF PRIME VALUES

G. W. FUNG AND H. C. WILLIAMS

ABSTRACT. The University of Manitoba Sieve Unit is used to find several values of A (> 0) such that the quadratic polynomial $x^2 + x + A$ will have a large asymptotic density of prime values. The Hardy-Littlewood constants which characterize this density are also evaluated.

1. INTRODUCTION

Let $f_A(x) = x^2 + x + A$ $(A \in \mathbf{Z},\ A > 0)$ and let $P_A(n)$ represent the number of prime values assumed by $f_A(x)$ for $x = 0, 1, \ldots, n$. In 1772 Euler discovered that $P_{41}(39) = 40$. Indeed, the polynomial x^2+x+41 is well known to all students of number theory because of this remarkable property. Consider, however, the much less famous polynomial $x^2+x+27941$ discovered, according to Karst [4], by Beeger in 1938. Here $P_{27941}(39) = 30$; but, $P_{27941}(1000000) = 286128$, whereas $P_{41}(1000000) = 261080$. It appears, then, that while $f_{41}(x)$ starts off very well in the production of primes, the rather more modest $f_{27941}(x)$ begins to better its famous rival as the values of x become large.

This phenomenon could have been predicted from Hardy and Littlewood's [3] Conjecture F. For the case of polynomials of the form $f_A(x)$ this conjecture can be given as

$$P_A(n) \sim C(D)L_A(n), \tag{1.1}$$

where $D = 1 - 4A$, $L_A(n) = 2\int_0^n dx/\log f_A(x)$, and

$$C(D) = \prod_{p \geq 3}(1 - (D/p)/(p-1)). \tag{1.2}$$

The product in (1.2) is taken over all the odd primes p, and by $(\cdot/p)$ we denote the Legendre symbol. Shanks [14] has computed $C(-163) = 3.3197732$ and $C(-111763) = 3.6319998$. Thus, on the basis of Conjecture F one would expect that for sufficiently large values of n, $P_{27941}(n)$ would exceed $P_{41}(n)$,

Received March 10, 1988; revised April 24, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11N32, 11Y60, 11Y40.
The second author's research was supported by NSERC of Canada Grant #A7649.

and this is what we have observed. Indeed, to five significant figures

$$P_{41}(100000)/L_{41}(1000000) = 3.3203,$$
$$P_{27941}(1000000)/L_{27941}(1000000) = 3.6397,$$

both of which are quite close to their respective $C(D)$-values.

The purpose of this note is to find other polynomials $f_A(x)$ which have a high asymptotic density of prime values. We will do this by determining those values of D for which the Hardy-Littlewood constant $C(D)$ should be large, and then evaluating $C(D)$ to eight significant figures. If Conjecture F is true, then the corresponding values of A should provide us with the desired polynomials. We point out here that all previous numerical tests of Conjecture F have tended to confirm its truth (see Shanks [9–11]).

2. Strategy for finding values of D

We first note that since we want $f_A(x)$ to assume prime values, then A must be odd; hence, $-D = 4A - 1 \equiv 3 \pmod 8$. In order to maximize the asymptotic value of $P_A(n)$, we can maximize $C(D)$. According to (1.2) this means that we would want $(D/p) = -1$ for as many of the small primes p as possible. As noted by Lehmer [5], we can also look at this from the point of view of restricting the number of possible small prime divisors of $f_A(x)$. If $(D/p) = -1$, then p cannot divide $f_A(x)$ for any value of x; thus, if $(D/p) = -1$ for many small primes p, then the composite values that $f_A(x)$ can assume are considerably restricted. It follows that $f_A(x)$ should frequently be a prime.

If we let N_r denote the least positive integer such that $N_r \equiv 3 \pmod 8$ and $(N_r/p) = -(-1/p)$ for all odd primes $p \le p_r$, where p_r is the rth prime, then $-N_r$ should be a good candidate for the kind of D-value that we are seeking. This was the strategy used in [5] to find values for D. In Table 2.1 we give all the values of N_r up to $r = 42$.

The values of N_r for $r \le 28$ are given in [5]; the values of N_r for $r \le 38$ are given in Lehmer, Lehmer, and Shanks [6]; and the values of N_r for $r = 39$, 40, 41, 42 were found by D. H. Lehmer but have not been previously published, except for N_{42} which appears in Shanks [15].

To find all these values of N_r, Lehmer made use of mechanized or electronic number sieving devices. Such specialized machines are small, fast, and much less expensive than general purpose computers. In the production of the numbers presented in this paper we made use of the University of Manitoba Sieve Unit (UMSU) [8]. This device solves systems of linear congruences at a trial rate of 133,000,000 numbers per second. Thus, in about three months of continuous use, we were able to examine numbers up to 10^{15}.

TABLE 2.1

r	p_r	N_r	r	p_r	N_r
2	3	19	23	83	114148483
3	5	43	24	89	269497867
4	7	43	25	97	269497867
5	11	67	26	101	269497867
6	13	67	27	103	269497867
7	17	163	28	107	585811843
8	19	163	29	109	52947440683
9	23	163	30	113	52947440683
10	29	163	31	127	71837718283
11	31	163	32	131	229565917267
12	37	163	33	137	229565917267
13	41	77683	34	139	575528148427
14	43	77683	35	149	1432817816347
15	47	1333963	36	151	1432817816347
16	53	2404147	37	157	1432817816347
17	59	2404147	38	163	1432817816347
18	61	20950603	39	167	6778817202523
19	67	36254563	40	173	16501779755323
20	71	51599563	41	179	30059924764123
21	73	96295483	42	181	30059924764123
22	79	96295483			

Note that Beeger's number, 111763, is not in Table 2.1, yet it has a better $C(D)$-value than $C(-77683) = 3.3003388 < C(-163)$.[1] If we put $N_{r,1} = N_r$ above and define $N_{r,i}$ $(i > 1)$ as the least integer greater than $N_{r,i-1}$ such that $N_{r,i} \equiv 3 \pmod 8$ and $(N_{r,i}/p) = -1$ for all odd primes $p \leq p_r$, then $111763 = N_{43,2}$. Thus, instead of attempting simply to tabulate more N_r values than those given in Table 2.1, we used UMSU to compute $N_{r,i}$ for $r \leq 40$ and $i \leq 10$. For $r = 41$ we let UMSU continue to find values of $N_{r,i}$ until these values exceeded 10^{15}. We were thus able to find all $N_{41,i}$ for $i \leq 15$. Having these candidates for D, the next problem is to determine those that yield the largest $C(D)$ values. Unfortunately, the product (1.2) converges very slowly; hence, we must develop an alternative method of computing $C(D)$, especially for large values of D.

3. COMPUTATION OF $C(D)$

Efficient methods for evaluating $C(D)$ have been developed by Shanks [10, 11, 14]. In [10, 11] he discovered a method of finding $C(D)$ which appears to work well when D is fairly small, and in [14] he provided a method of determining $C(D)$ to high accuracy, which will work when D is larger. Indeed, he provides values for $C(-163)$, $C(-77683)$, $C(-111763)$, $C(-289963)$, $C(-991027)$, the latter value being 4.1237067, the largest C-value known

[1] Shanks [14] gets 3.2999354... for $C(-77683)$; however, we evaluated this number in two different ways and still got 3.3003388. Thus we feel that some minor error crept into Shanks' evaluation of $C(-77683)$. Our results agree with all of Shanks' other evaluations.

until now. If we put

$$L(s, \chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s} = \prod_{p} (1 - \chi(p)/p^{s})^{-1},$$

where $\chi(n) = (D/n)$ and $(\cdot/n)$ is the Kronecker symbol, then all of these methods require that $L(s, \chi)$ be computed for various values of s. If h is the class number of $\mathscr{Q}(\sqrt{D})$, the value of $L(s, \chi)$ can be determined fairly readily when the values of h Epstein zeta functions are known. As Shanks [14] can rapidly evaluate these Epstein zeta functions to high accuracy, he can then accurately compute $C(D)$. However, if h is large, this method can be quite slow. It is, however, the best method to use when $C(D)$ is needed to great accuracy. Also, this technique, unlike the one we will discuss below is unconditional.

As it was necessary for us to calculate $C(D)$ for many D-values, some of which were very large, we needed to develop a faster method to compute $C(D)$. We also assumed that evaluating $C(D)$ to eight significant figures would be adequate for the purposes of this note.

We first point out that if we use an idea in [10], it is a simple matter to show that

$$\frac{C(D)L(1, \chi)L(2, \chi)}{\zeta(4)} = \frac{1}{2} \prod_{p|D} \left(1 - \frac{1}{p^4}\right) \prod_{q \geq 3} \left(1 - \frac{2}{q(q-1)^2}\right),$$

where the first product on the right is taken over all the primes p which divide D and the second is taken over all primes q such that $(D/q) = 1$. Since $\zeta(4) = \pi^4/90$ and

$$\pi h/\sqrt{|D|} = L(1, \chi), \tag{3.1}$$

we get

$$C(D) = \frac{\pi^3 \sqrt{|D|}}{180h} \cdot \frac{1}{L(2, \chi)} \cdot \prod_{p|D} \left(1 - \frac{1}{p^4}\right) \prod_{q \geq 3} \left(1 - \frac{2}{q(q-1)^2}\right). \tag{3.2}$$

Put

$$F_1(Q) = \prod_{p \leq Q} p/(p - \chi(p)), \qquad T_1(Q) = \prod_{p > Q} p/(p - \chi(p)),$$

$$F_2(Q) = \prod_{p \leq Q} p^2/(p^2 - \chi(p)), \qquad T_2(Q) = \prod_{p > Q} p^2/(p^2 - \chi(p)),$$

$$F_3(Q) = \prod_{q \leq Q} \left(1 - \frac{2}{q(q-1)^2}\right), \qquad T_3(Q) = \prod_{q > Q} \left(1 - \frac{2}{q(q-1)^2}\right).$$

By examining (3.1) and (3.2), we see that two problems arise in computing $C(D)$: (1) evaluate $L(1, \chi)$ to sufficient accuracy to determine h, (2) find Q

such that

$$C(D) \approx \frac{\pi^3\sqrt{|D|}}{180h}\frac{F_3(Q)}{F_2(Q)}\prod_{p|D}\left(1-\frac{1}{p^4}\right) \tag{3.3}$$

approximates $C(D)$ to eight significant figures.

It is a simple matter to show that

$$|\log T_2(Q)| + |\log T_3(Q)| = \left|\sum_{p>Q}\chi(p)/p^2\right| + \delta_1,$$

where $|\delta_1| < 2/Q^2$ $(Q > 10)$. Also,

$$|\log T_1(Q)| = \left|\sum_{p>Q}\chi(p)/p\right| + \delta_2,$$

where $|\delta_2| < 1/Q$. If we set

$$B(Q) = \log|D|\left(\frac{1}{\pi\log Q} + \frac{5\cdot 3}{(\log Q)^2}\right) + \frac{4}{\log Q} + \frac{1}{\pi},$$

then by using the method of Cornell and Washington [2], we get

$$|\log T_2(Q)| + |\log T_3(Q)| < B(Q)(8 + 13\log Q)/(9Q^{3/2}) + 2/Q^2 \tag{3.4}$$

and

$$|\log T_1(Q)| < B(Q)(4 + 3\log Q)/Q^{1/2} + 1/Q = A(Q). \tag{3.5}$$

It is important to note that the proof of these inequalities requires the truth of the Riemann Hypothesis on $L(s, \chi)$. Thus, the method that we develop here for finding $C(D)$ is contingent on the Extended Riemann Hypothesis.

Now if

$$|\log T_2(Q)| + |\log T_3(Q)| < b,$$

then (3.3) will approximate $C(D)$ to r significant figures if $b < \log((1 + \sqrt{1+4k})/2)$, where $k = 10^{1-r}/2$. Hence, by (3.4), if $D \approx 10^{15}$, then $Q = 10^6$ in (3.3) will yield $C(D)$ to eight figures. Of course, for smaller values of D, smaller values of Q can be used in (3.3). To test this, we evaluated (3.3) for the largest D-values we found with $Q = 10^6$ and $Q = 5\cdot 10^6$. In every case, both computations agreed to eight significant figures.

There remains the problem of determining h. For this problem we used a modification of the idea of Lenstra [7]. If, for a fixed Q, we put

$$\pi B_1 = \sqrt{|D|}F_1(Q)\exp(A(Q)), \qquad \tilde{h} = \mathrm{Ne}(\sqrt{|D|}F_1(Q)/\pi),$$
$$k = \sqrt{|D|}F_1(Q)/\pi - \tilde{h}, \qquad B_2 = |k| + B_1 - \sqrt{|D|}F_1(Q)/\pi,$$

then by (3.1) and (3.5) we have

$$h \le B_1, \qquad |\tilde{h} - h| < B_2.$$

Here, by $\mathrm{Ne}(x)$ we denote the nearest integer to x. Now if we know a divisor h_1 of h such that

$$\tilde{h}/h_1 - B_2/h_1 - [B_2/h_1 + \tilde{h}/h_1] > -1, \tag{3.6}$$

then $h_2 = [B_2/h_1 + \tilde{h}/h_1]$ is the only integer in the interval I given by

$$\tilde{h}/h_1 - B_2/h_1 < x \leq [B_2/h_1 + \tilde{h}/h_1].$$

Since $|h/h_1 - \tilde{h}/h_1| < B_2/h_1$, we see that h/h_1 must be in I. It follows that $h = h_1 h_2$ when (3.6) holds.

If (3.6) does not hold for $h_1 = 1$, we can use the baby step–giant step method of Shanks [12] to find a divisor h_1 of h such that $h_1 > 1$. In fact, since most of the class groups of $Q(\sqrt{D})$ are cyclic or close to it (see Cohen and Lenstra [1]), this technique rapidly provides a value of h_1 which is close to h in value; hence, (3.6) is usually satisfied very quickly.

4. Numerical results

The method described in §3 was programmed in FORTRAN with some assembly language subroutines and run on an Amdahl 5870 computer. For a given D, a value of Q was determined which would guarantee eight figures of accuracy for $C(D)$ by (3.3). The values of $F_1(Q)$, $F_2(Q)$, $F_3(Q)$ were evaluated simultaneously in double precision by the assembly language subroutines.

The $C(D)$-values for the 192 numbers found by UMSU were computed in a total of about three CPU minutes. Denote by $q(D)$ the least prime such that $(D/q(D)) \neq -1$. In Table 4.1 we give all the numbers D found by UMSU with $q(D) \geq 163$. We also provide the corresponding values of h, $C(D)$, and $q(D)$.

Notice that 110587910656507 allows us to extend Table 2.1. In fact, this number is N_{43}, N_{44}, and N_{45}. It is rather unfortunate that $(-N_{43}/199) = 1$ because $(-N_{43}/p) = -1$ for $p = 211$, 223, 227, 229, and 223. Thus, if it were not for the value of the Legendre symbol for 199 we would have $N_{51} = N_{43}$. As it is, the best that we can say here is that $N_{46} > 10^{15}$.

In Table 4.2 we give those values of D from among the 192 such that $C(D) > C(D')$ for all the D', among the 192, which are less than D. We also give the corresponding value of $P_A(1000000)$ and $P_A(1000000)/L_A(1000000)$ (written as $P(1000000)$ and P/L, respectively), where $A = (1 - D)/4$. Also, since the $C(D)$-values are roughly inversely proportional to the respective $L(1, \chi)$-values, we provide these values as $L(D) = L_{-D}(1) = \frac{3}{2}L(1, \chi)$ in order to permit comparison with the results of Shanks [13]. Notice that $P_A(1000000)/L_A(1000000)$ and $C(D)$ are quite close, providing yet further confirmation of Conjecture F.

Let D' denote the last D-value in Table 4.2 ($D' = -531\ldots$). As we would expect, for $D = -N_{43}$ we get a quite large $C(D)$-value. What appears, at first, to be somewhat remarkable is that this $C(D)$-value is less than $C(D')$; however, even though $q(-N_{43}) > q(D')$, beyond $q(D')$ we get a higher density of nonresidues for D' than for $-N_{43}$, hence the larger $C(D)$-value for D'.

TABLE 4.1

D	h	$C(D)$	$q(D)$
−1432817816347	70877	4.4163429	167
−5066580103267	131930	4.4616823	163
−6626709638707	148069	4.5468709	163
−6778817202523	149460	4.5565681	173
−8547099746707	176959	4.3197166	163
−8903633500507	168563	4.6296715	163
−9275311526083	179187	4.4443528	163
−15159061903507	225866	4.5075910	163
−16501779755323	223574	4.7524812	179
−17542900082563	240468	4.5549908	163
−30059924764123	296475	4.8379057	191
−37221595794667	328170	4.8634109	179
−50923056589267	423835	4.4030138	179
−58212094833523	427986	4.6627278	167
−58369246601803	429990	4.6472033	173
−64279195020307	454960	4.6086679	173
−65569185073723	444070	4.7692730	167
−69298004348827	474064	4.5926958	173
−74210430269347	454842	4.9548401	181
−82973459224363	530640	4.4889020	173
−87934318851787	528770	4.6380011	181
−88795060352923	519882	4.7412369	179
−110587910656507	553436	4.9711959	199
−126620398458283	640888	4.5916614	179
−138411891537187	655794	4.6925243	179
−307568240581123	949638	4.8308250	181
−378486993318883	1082408	4.7014317	193
−404210888356867	1123425	4.6808868	191
−414286790833987	1158810	4.5940080	181
−531497118115723	1185668	5.0870883	181
−696687486054883	1410630	4.8947820	191
−772147706149747	1529796	4.7510872	181
−792933985668883	1578348	4.6664594	181
−799705726392763	1554144	4.7590191	181
−850229380873387	1596966	4.7756941	191
−998727466696243	1789388	4.6191771	181

To further illustrate this phenomenon, we give in Table 4.3 some more $C(D)$-values. The D-values in this table are taken from the latter part of Shanks' table of Lochamps (Table 3) in [13]. We have only selected those D-values which are not already in Table 4.2 and are congruent to 5 modulo 8. We have also reproduced the $L(D)$-values given in Shanks' table.

Notice that for $D = -991027$ and $D = -3416131987$ we get larger $C(D)$-values than those given in Table 4.2 for D-values of comparable size. Thus, there may be more numbers $< 10^{15}$ with $C(D)$-values in excess of $C(-N_{43})$. What we can say here is that if Conjecture F holds, then

$$x^2 + x + 132874279528931$$

TABLE 4.2

D	$P(1000000)$	$C(D)$	P/L	$L(D)$
−163	261080	3.3197732	3.3203421	0.3691028
−85507	272102	3.4643422	3.4612190	0.3545382
−111763	286128	3.6319998	3.6396821	0.3383011
−222643	293169	3.7289570	3.7293962	0.3295722
−1333963	300001	3.8123997	3.8169182	0.3223267
−9471067	312436	3.9760501	3.9764927	0.3093093
−10560643	315542	4.0194873	4.0161335	0.3059697
−60408307	318250	4.0501092	4.0531570	0.3037600
−171583003	320126	4.0815068	4.0796515	0.3014727
−269497867	322488	4.1092157	4.1112637	0.2996843
−398158363	325782	4.1579113	4.1548155	0.2961493
−643338763	335224	4.2716019	4.2775772	0.2883454
−1408126003	334712	4.2771747	4.2759778	0.2879549
−1595514187	341572	4.3616794	4.3645752	0.2824327
−4067175907	346057	4.4324788	4.4309683	0.2779060
−71837718283	354875	4.6097143	4.6090901	0.2673146
−85702502803	361841	4.7073044	4.7067227	0.2617208
−16501779755323	326605	4.7524812	4.7559512	0.2593564
−30059924764123	326392	4.8379057	4.8453809	0.2548210
−37221595794667	325086	4.8634109	4.8594354	0.2534793
−74210430269347	323289	4.9548401	4.9413604	0.2488108
−110587910656507	321488	4.9711959	4.9770300	0.2480017
−531497118115723	312975	5.0870883	5.0894316	0.2423560

TABLE 4.3

D	$P(1000000)$	$C(D)$	P/L	$L(D)$
−546067	297046	3.7775732	3.7789730	.32523
−991027	324001	4.1237067	4.1221307	.29822
−1970364883	339556	4.3367305	4.3405407	.28398
−2426489587	343914	4.4024373	4.3981264	.27982
−3416131987	353395	4.5247200	4.5229186	.27227
−8864190043	355373	4.5655590	4.5612380	.26983

is a quadratic polynomial which has a higher asymptotic density of prime values than any other such polynomial known to date.

ACKNOWLEDGMENT

The authors wish to thank D. H. Lehmer for making his unpublished work available to them. They also wish to thank Daniel Shanks for many helpful observations and references.

BIBLIOGRAPHY

1. H. Cohen and H. W. Lenstra, Jr., *Heuristics on class groups of number fields*, Number Theory (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer-Verlag, Berlin and New York, 1984, pp. 33–62.

2. G. Cornell and L. C. Washington, *Class numbers of cyclotomic fields*, J. Number Theory **21** (1985), 260–274.
3. G. H. Hardy and J. E. Littlewood, *Partitio numerorum* III: *On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.
4. E. Karst, *The congruence* $2^{p-1} \equiv 1 \pmod{p^2}$ *and quadratic forms with high density of primes*, Elem. Math. **22** (1967), 85–88.
5. D. H. Lehmer, *On the function of* $x^2 + x + A$, Sphinx **6** (1936), 212–214 and Sphinx **7** (1937), 40.
6. D. H. Lehmer, E. Lehmer, and D. Shanks, *Integer sequences having prescribed quadratic character*, Math. Comp. **24** (1970), 433–451.
7. H. W. Lenstra, Jr., *On the calculation of regulators and class numbers of quadratic fields*, London Math. Soc. Lecture Note Ser., vol. 56, Cambridge Univ. Press, Cambridge and New York, 1982, pp. 123–150.
8. C. D. Patterson and H. C. Williams, *A report on the University of Manitoba Sieve Unit*, Congr. Numer. **37** (1983), 85–98.
9. Daniel Shanks, *A sieve method for factoring numbers of the form* $n^2 + 1$, MTAC **13** (1959), 78–86.
10. ____, *On the conjecture of Hardy and Littlewood concerning the number of primes of the form* $n^2 + a$, Math. Comp. **14** (1960), 320–332.
11. ____, *Supplementary data and remarks concerning a Hardy-Littlewood conjecture*, Math. Comp. **17** (1963), 188–193.
12. ____, *Class number, a theory of factorization and genera*, Proc. Sympos. Pure Math., vol. 20 (1969 Institute on Number Theory), Amer. Math. Soc., Providence, R.I., 1971, pp. 415–440.
13. ____, *Systematic examination of Littlewood's bounds on* $L(1, \chi)$, Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 267–283.
14. ____, *Calculation and applications of Epstein zeta functions*, Math. Comp. **29** (1975), 271–287.
15. ____, *A survey of quadratic, cubic and quartic algebraic number fields*, Congr. Numer. **17** (1976), 15–42.

Department of Computer Science, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2