Publ. Math. Debrecen
62/3-4 (2003), 589–600

Remarks on prime values of polynomials at prime arguments

By IMRE Z. RUZSA (Budapest) and SÁNDOR TURJÁNYI (Debrecen)

Dedicated with gratitude and respect to Professor Lajos Tamássy on the occasion of his 80th birthday

Abstract. We show that for a large number N and for any functions taking integer values bounded by polynomials there exist integers $t_1, t_2, \dots, t_s$ and primes $2 \le p_1 < p_2 < \cdots < p_M \le N$ such that the numbers $f_1(p_i) + t_1, f_2(p_i) + t_2, \dots, f_s(p_i) + t_s$ are primes for $i = 1, 2, \dots, M$ and $M \ge c\frac{N}{\ln N^{s+1}}$ (where c is a positive constant independent of N), as expected. In Theorem 3.1 we show that when $s = 1$ and f_1 is a polynomial taking integer values the order of the magnitude of M can be improved by a factor $\log\log N$. Theorem 4.1 illustrates that there exists a function f_1 which is "not far" from being a polynomial but the value of M never exceeds the expected order of magnitude, so in the improvement it is essential that the function is actually a polynomial.

1. Introduction

It is a generally accepted conjecture that an irrecucible polynomial in one variable with integer coefficents without a constant divisor assumes infinitely many prime values (positive or negative) at integers. This is known

Research of the first author supported by Hungarian National Foundation for Scientific Research Grants No. T 38396 and T 25617. Research of the second author supported by Hungarian National Foundation for Scientific Research Grant No. T 29330.

to hold for linear polynomials (Dirichlet's theorem), even in a quantitative form (the prime number theorem for arithmetic progressions), and it seems to lie beyond the reach of present methods for any polynomial of higher degree. The inaccessibility of this conjecture led several authors to establish at least the existence of polynomials with many prime values.

Sierpiński [19] showed that there exists an integer c_M for any large positive integer M such that the polynomial

$$f(x) = x^2 + c_M$$

takes prime values for M natural numbers at least. His result has been extended in various ways in [1], [4], [7], [8] and [20].

A stronger conjecture is expressed by Schinzel's Hypothesis H, which asserts that for any collection $f_1, \ldots, f_j$ of irreducible polynomials with integral coefficients and positive leading coefficient such that $\prod f_i$ has no constant divisors there are infinitely many positive integers n such that each $f_i(n)$ is a prime. The simplest case of this hypothesis is the twin prime conjecture, also generally considered hopeless. Ribenboim [17] gives a comprehensive account of the connection between primes and polynomials and among other things he treats the hypothesis of Schinzel in detail.

Let $f_1, \ldots, f_s$ be arbitrary functions taking integer values. Denote by $Q(f_1, f_2, \ldots, f_s; N)$ the number of integers n between 0 and N for which each number $f_1(n), f_2(n), \ldots, f_s(n)$ is a prime, and by $P(f_1, f_2, \ldots, f_s; N)$ the number of primes p between 1 and N such that each number $f_1(p)$,$f_2(p) \ldots, f_s(p)$ is a prime.

From the notation it is clear that

$$P(f_1, f_2, \ldots, f_s; N) = Q(f_0, f_1, f_2, \ldots, f_s; N),$$

if $f_0(x) = x$.

With this notation Schinzel Hypothesis H can be reformulated as

$$\lim_{N\to\infty} Q(f_1, f_2, \ldots, f_s; N) = \infty.$$

Bateman and Horn [3] formulated a quantitative version of Hypothesis H in the following form. If the polynomials $f_1, f_2, \ldots, f_s$ satisfy the conditions of Hypothesis H and their degrees are $h_1, h_2, \ldots, h_s$ then

$$Q(f_1, f_2, \ldots, f_s; N) \sim h_1^{-1} h_2^{-1} \ldots h_s^{-1} C(f_1, f_2, \ldots, f_s) \int_2^N (\log u)^{-s} du$$

where

$$C(f_1, f_2, \dots, f_s) = \prod_p \left((1-1/p)^{-s}(1-\omega(p)/p)\right).$$

Here p runs over all the prime numbers and $\omega(p)$ denotes the number of solutions of the congruence

$$f_1(x)f_2(x)\dots f_s(x) \equiv 0 \pmod{p}.$$

As a contrast we mention some results that show the limitations of such hypotheses by exhibiting irreducible polynomials which assume composite values for a long run.

Adleman and Odlyzko pose the following problem in [2]: how can we decide the irreducibility of a polynomial $f(x)$ with integer coefficients by testing the prime property of $f(v)$ for small v.

In [13]–[15] McCurley studied for which M and a can an irreducible polynomial $f(x) = x^d + a$ takes composite values at v for $0 \le v \le M$. He obtained the strongest result in [15] which says that

$$M \ge C(d)\frac{\log a}{\log_3 a}\left(\frac{(\log_2 a)(\log_4 a)}{\log_3 a}\right)^{\tau(d)}$$

for infinitely many values of a, where $c(d)$ is a positive constant depending only on d, $\log_k a$ denotes the logarithm iterated k times, $\tau(d)$ denotes the number of the divisors of d.

This is (apart from the value of the constant) a generalization of a result of Rankin [16] which asserts that there exist infinitely many consecutive primes p_{n+1} and p_n such that

$$p_{n+1} - p_n > (e^\gamma - \varepsilon)\log p_n \frac{\log_2 p_n \log_4 p_n}{(\log_3 p_n)^2}$$

(where γ is Euler's constant and $\varepsilon > 0$).

It is possible to extend McCurley's result to general polynomials, that is, to find a translate which assumes composite values in a long run. We remark that while it is possible to tell explicitly whether a binomial $x^d + a$ is reducible, this is not the case for the translates of a general polynomial. However, one can get around this by using the fact that irreducible polynomials are ubiquitous; for instance, a result of Győry [9] states that an

arbitrary polynomial $f(x)$ has a translate $f + c$ which is irreducible and the upper bound of the absolute values of c depends only on the number of the divisors of the leading coefficients of the polynomial f and its degree. We plan to return to this problem in a future paper.

In Section 2 we show by a simple averaging argument that every function of polinomial order of magnitude has a translation that assumes $\gg N/(\log N)^2$ prime values at prime arguments $\le N$. In Section 3 we show that for an actual polynomial this can be improved by a factor $\log\log N$. In Section 4 we exhibit an example of an almost polynomial function for which such an improvement is impossible.

2. Prime values of translations of general functions

We consider integer-valued functions defined on the set of positive integers. We say that such a function f is polynomially bounded if there is a polynomial g such that $|f(x)| \le g(x)$ for sufficiently large x, or equivalently, if $|f(x)| \le Ax^d$ for large x and fixed A, d.

Theorem 2.1. *If $f_1, f_2, \dots, f_s$ are arbitrary polynomially bounded functions then for sufficiently large N there exist integers $t_1, t_2, \dots t_s$ (depending on N) such that*

$$P(f_1 + t_1, f_2 + t_2, \dots, f_s + t_s, N) \ge c\frac{N}{(\log N)^{s+1}}$$

where the constant c does not depend on N.

Proof. Take d, A such that $|f_j(x)| \le Ax^d$ for large x. Now consider $(s+1)$-tuples $(p, p_1, \dots, p_s)$ of primes such that

$$1 \le p \le N, \quad 1 \le p_i \le AN^d.$$

The number of such tuples is clearly $\pi(N)\pi(AN^d)^s$. For each such tuple there are unique integers $t_1, \dots, t_s$ such that

$$f(p) + t_i = p_i, \imath = 1, \dots, s.$$

These numbers satisfy

$$|t_i| = |p_i - f(p)| \le 2AN^d.$$

Hence the number of s-tuples $(t_1, \dots, t_s)$ is $\leq (4AN^d + 1)^s$. Consequently there is such a tuple which occurs at least

$$\frac{\pi(N)\pi(AN^d)^s}{(4AN^d+1)^s} \geq c\frac{N}{(\log N)^{s+1}}$$

times; the last inequality holds for large N with any $c < (4d)^{-s}$ by the prime number theorem. □

3. Prime values of polynomials at prime arguments

Let f be a polynomial with integral coefficients. Define

$$P(f, N) = \{p \leq N : p \text{ and } f(p) \text{ are primes}\}.$$

It is not known whether $P(f, N) \to \infty$ as $N \to \infty$, except the trivial case $f(x) = \pm x$. The simplest case of this question is the twin prime conjecture, which arises for $f(x) = x + 2$.

On the other hand, if we consider a polynomial together with its translates, it becomes possible to find lower estimates. The averaging argument of the previous section gives the existence of $t \in N$ such that

$$P(f + t, N) \gg \frac{N}{(\log N)^2}$$

(case $s = 1$ of Theorem 2.1).

Our aim is to improve this estimate if f is a polynomial.

Theorem 3.1. *Let f be a polynomial of degree d with integral coefficients. We have*

$$\max_{1 \leq t \leq N^{d+1}} P(f + t, N) \geq c_d \frac{N \log\log N}{(\log N)^2}$$

with a constant $c_d > 0$ depending on the degree only, and all sufficiently large N.

The case $f(x) = x$ is essentially a result of Erdős and Straus [6], and the proof is similar to theirs. The idea is to average $P(f+t, N)$ for the integers belonging to a suitable arithmetic progression $t \equiv a \pmod m$. We want to find such numbers for which the integers $f(p) + t$ have a greater than average chance of being prime. Coprimality to m increases this chance, and first we show how this can be achieved.

Lemma 3.2. *Let f be a polynomial with integral coefficients, and let m be a positive integer. We can find an integer a such that $(f(p)+a, m) = 1$ for every prime p such that $p \nmid m$.*

Proof. Let $p_1, p_2, \dots, p_k$ be the prime divisors of m. For an i, $1 \le i \le k$, consider the integers $f(1), f(2), \dots, f(p_i - 1)$. These occupy at most $p_i - 1$ residue class modulo p_i, hence there is a residue class, say $b_i \pmod{p_i}$, which does not contain any of them.

Any prime $p \nmid m$ satisfies $p \equiv j \pmod{p_i}$ with $1 \le j \le p_i - 1$, hence $f(p) \equiv f(j) \pmod{p_i}$, consequently $f(p) \not\equiv b_i \pmod{p_i}$. If we now select a so that $a \equiv -b_i \pmod{p_i}$ for each $i = 1, 2, \dots, k$, then the above congruences becomes $f(p) + a \not\equiv 0 \pmod{p_i}$. As this holds for every prime divisor of m, we conclude that $(f(p) + a, m) = 1$.

For the estimation we will need the prime number theorem for arithmetic progressions which we quote now, see e.g. Karatsuba [12]. We use $\pi(x, a, m)$ to denote the number of primes $p \le x$ satisfying $p \equiv a \pmod m$. □

Lemma 3.3. *There are positive constants c_1, c_2 such that uniformly for $m \le \exp(c_1\sqrt{\log x}\,)$, $(a, m) = 1$ we have*

$$\pi(x, a, m) = \frac{\operatorname{li}(x)}{\varphi(m)} + O\big(x \exp(-c_2\sqrt{\log x}\,)\big) \tag{3.1}$$

except possibly for those integers m that are multiples of a certain integer m_0 (which may depend on x).

This integer m_0 is characterized by the existence of a Siegel root of an L-function belonging to the modulus m_0.

Proof of Theorem 3.1. Consider the sum

$$S = \sum_{1 \le t \le N^{d+1},\ t \equiv a \pmod m} P(f + t, N);$$

we shall specify a and m later. This counts the number of pairs (p,t) such that $1 \le p \le N$, $1 \le t \le N^{d+1}$, $t \equiv a \pmod m$ and both p and $f(p)+t$ are primes. We rearrange this sum according to the value of p. For a given p, the integers $f(p)+t$, $1 \le t \le N^{d+1}$, run over the interval $[f(p)+1, f(p)+N^{d+1}]$ and the condition $t \equiv a \pmod m$ is equivalent to $f(p)+t \equiv f(p)+a \pmod m$. So the number of prime values is

$$\pi(f(p)+N^{d+1}, f(p)+a, m) - \pi(f(p), f(p)+a, m).$$

Summing this we obtain

$$S = \sum_{p \le N} \Big(\pi(f(p)+N^{d+1}, f(p)+a, m) - \pi(f(p), f(p)+a, m) \Big).$$

To estimate this first observe that $|f(p)| \le AN^d$ with some constant A, hence

$$\pi(f(p)+N^{d+1}, f(p)+a, m) = \pi(N^{d+1}, f(p)+a, m) + O(N^d)$$

and $\pi(f(p), f(p)+a, m) = O(N^d)$, so we can simplify S to

$$S = \sum_{p \le N} \pi(N^{d+1}, f(p)+a, m) + O(N^{d+1}).$$

We will select m so that Lemma 3.3 could be applied, that is, $m_0 \nmid m$ with the m_0 belonging to $x = N^{d+1}$. We take an a such that $(f(p)+a, m)=1$ for every prime $p \nmid m$, which can be done by Lemma 3.2. Then applying Lemma 3.3 to each summand and denoting by $\omega(m)$ the number of prime factors of m we get

$$S \ge (\pi(N) - \omega(m)) \frac{\operatorname{li}(N^{d+1})}{\varphi(m)} + O\big(N^{d+2} e^{-c_2\sqrt{\log N}}\big)$$

as $\omega(m) < \log N$, $\pi(N) \sim \frac{N}{\log N}$ and $\operatorname{li}(N^{d+1}) \sim \frac{N^{d+1}}{(d+1)\log N}$ and this implies

$$S \ge \frac{N^{d+2}}{(d+1)(\log N)^2 \varphi(m)} (1+o(1)).$$

As the number of summands in S is $\frac{N^{d+1}}{m} + O(1)$ we conclude that

$$\max_{1 \le t \le N^{d+1}} P(f+t, N) \ge (1+o(1)) \frac{N}{(\log N)^2} \, \frac{m}{\varphi(m)} \, \frac{1}{d+1}. \tag{3.2}$$

Now we specify m. Let q_0 be the greatest prime factor of m_0. We will put

$$m = \prod_{p \le K,\ p \ne q_0} p,$$

this guarantees $m_0 \nmid m$ where K tending to infinity with N will be chosen later. With this choice

$$\frac{m}{\varphi(m)} = \prod_{p \le K,\ p \ne q_0} \left(1 - \frac{1}{p}\right)^{-1}.$$

As

$$\prod_{p \le K} \left(1 - \frac{1}{p}\right)^{-1} \sim e^{\gamma} \log K$$

according to Mertens' theorem, we have

$$\frac{m}{\varphi(m)} \sim \left(1 - \frac{1}{q_0}\right) e^{\gamma} \log K \ge \left(\frac{e^{\gamma}}{2} + o(1)\right) \log K.$$

In order to achieve $m < \exp(c_1 \sqrt{\log N}\,)$ we put $K = c_3 \sqrt{\log N}$ with $c_3 < c_1$, then $\log K \sim \frac{1}{2} \log\log N$ and

$$\frac{m}{\varphi(m)} \ge \left(\frac{e^{\gamma}}{2} + O(1)\right) \log\log N.$$

Substituting this into (3.2) we obtain the claim of the theorem for any $c_d < \frac{e^{\gamma}}{4(d+1)}$. With a slightly more careful calculation, the value of c_d can be improved to $\frac{e^{\gamma}}{2d}$. Indeed, the range of averaging can be lowered to $N^{d+o(1)}$, and the estimate $q_0 \ge 2$ can be improved. Namely, we know that $m_0 \to \infty$ (in fact $m_0 \gg (\log N)^B$ for any fixed B), hence if m_0 is squarefree, then $q_0 \gg \log m_0 \gg \log\log N$. If m_0 is not squarefree, then

$$m_0 \nmid \prod_{p \le K} p,$$

so there is no need to exclude a prime from the product at all.

4. Prime values of almost polynomials at prime arguments

In this section we exhibit an example of a function f which is very near to a polynomial and for which the analog of Theorem 3.1 fails.

Theorem 4.1. *Let f be a function of the form $f(n) = [g(n)]$, where g is a polynomial of degree $d \geq 1$ such that all coefficients of g are integers, except that of degree 1 which is a quadratic irrational. We have*

$$P(f + t, N) \leq c\frac{N}{(\log N)^2}$$

for all $N \geq N_0(f)$ with an absolute constant c and $1 \leq t \leq N^{d+1}$.

Proof. We can write f as $f(n) = h(n) + [\alpha n]$, where α is a quadratic irrational and h is a polynomial with integer coefficients. □

Lemma 4.2. *The number of positive integers $n \leq N$ satisfying*

$$n \equiv u \pmod{m}, \quad [\alpha n] \equiv v \pmod{m}$$

is $N/m^2 + O(\log N)$.

Proof. Write $n = u + km$. The condition $1 \leq n \leq N$ is essentially equivalent to $1 \leq k \leq N/m$, the difference being at most two numbers. The condition $[\alpha n] \equiv v \pmod{m}$ is equivalent to

$$\left\{\frac{\alpha n - v}{m}\right\} < 1/m,$$

and after substituting $n = u + km$ this becomes

$$\left\{\alpha k + \frac{\alpha u - v}{m}\right\} < 1/m.$$

As the discrepancy of the sequence $\{\alpha k\}$ is known to be of order $O(\log N)$ for quadratic irrationals (or for any real number with bounded partial quotients), this number is $N/m^2 + O(\log n)$. □

Lemma 4.3. *Let m be a squarefree integer, t arbitrary. The number of pairs (u, v) of residue classes modulo m satisfying*

$$u(h(u) + t + v) \equiv 0 \pmod{m} \tag{4.1}$$

is

$$\prod_{p|m}(2p-1).$$

PROOF. First consider the case when m is a prime. The solutions are clearly $u \equiv 0$ and arbitrary v, which yields p solutions, or $u \not\equiv 0$ and $v \equiv -h(u) - t$, which yields further $p-1$. For a composite m the number of solutions is clearly the product of the number of solutions for the prime divisors of m. □

Lemma 4.4. *Let m be a squarefree integer, t arbitrary. The number of integers $n \le N$ such that*

$$n(f(n)+t) \equiv 0 \pmod{m} \tag{4.2}$$

is $\gamma_m N + O(m^2 \log N)$, where γ_m is defined by

$$\gamma_m = \prod_{p|m} \frac{2p-1}{p^2}$$

and the implied constant is independent of t.

PROOF. We split the integers $n \le N$ into m^2 classes according to the residue of n and $[\alpha n]$ modulo m. From Lemma 4.2 we know that the cardinality of each class is $N/m^2 + O(\log N)$. If $n \equiv u$ and $[\alpha n] \equiv v$, then $f(n) \equiv h(u) + v \pmod{m}$, so congruence (4.2) is equivalent to (4.1). Lemma 4.3 tells us that the number of pairs u, v corresponding to a solution is $m^2\gamma_m$ and the claim of the lemma follows. □

Lemma 4.5. *Let A be a set of integers, let*

$$N_d = \big|\{a \in A : d \mid a\}\big|,$$

and let $S(A,y)$ denote the number of those elements of A that are free of prime divisors $\le y$. Assume that there is a nonnegative multiplicative function w such that $w(p)$ is bounded for all primes p and we have

$$N_d = Nw(d)/d + R_d \tag{4.3}$$

for every squarefree d whose all prime divisors are $\le y$. Then we have for arbitrary $u \ge 1$

$$S(A,y) = N\prod_{p\le y}\left(1-\frac{w(p)}{p}\right)\Big(1+O(u^{-u/2})\Big) + O\Big(\sum_{d\le y^u}|R_d|\Big). \tag{4.4}$$

This is a form of the combinatorial sieve, see [10]; we quoted it in a less general form than given there, where w is assumed to be bounded only in a certain average and sieving is done with a general subset of primes rather than all primes $\le y$.

PROOF of Theorem 4.1. We apply Lemma 4.5 to the set of integers $n(f(n)+t)$. If both n and $f(n)+t$ are primes, then this number is free of prime divisors $\le \sqrt{N}$, except when either $n \le \sqrt{N}$ or $|f(n)+t| \le \sqrt{N}$, which gives $O(\sqrt{N}\,)$ possibilites at most. We set $y = N^{1/4}$ and $u = 1$. The assumptions of Lemma 4.5 hold with $w(d) = d\gamma_d$. This function satisfies $w(p) = 2 - 1/p$, thus it is bounded, and a routine calculation reveals that

$$\prod_{p\le y}\left(1-\frac{w(p)}{p}\right) = (c+o(1))(\log y)^{-2}$$

with a positive constant c. Hence the principal term in (4.4) is $O(N(\log N)^{-2})$ as wanted. Since by Lemma 4.4 we have

$$|R_d| = O(d^2 \log N),$$

the remainder term is $O(y^3 \log N) = O(N^{3/4+\varepsilon})$.

The second author wishes to thank B. BRINDZA, K. GYŐRY and the referee of the paper for their valuable advice and help.

References

[1] V. ABEL and H. SIEBERT, Sequence with large numbers of prime values, *Amer. Math. Monthly* **100** (1993), 167–169.

[2] L. M. ADLEMAN and A. M. ODLYZKO, Irreducibility testing and factorization of polynomials, *Math. Comp.* **41** (1983), 699–709.

[3] PAUL T. BATEMAN and ROGER A. HORN, A heuristic asymptotic formula concerning the distribution of prime numbers, *Math. Comp.* **16** (1962), 363–367.

[4] T. BROWN, P. JAU-SHYONG SHIU and X. Y. YU, Sequencces with translates containing many primes, *Canad. Math. Bull.* **41**(1) (1998), 15–19.

[5] YONG-GAO CHEN and I. Z. RUZSA, Prime values of reducible polynomials, I, *Acta Arith.* **95** (2000), 185–193.

[6] P. ERDŐS and E. G. STRAUS, Remarks on the difference between consecutive primes, *Elem. D. Math.* **35** (1980), 115–118.

[7] R. Forman, Sequences with many primes, *Amer. Math. Monthly* **99** (1992), 548–557.

[8] B. Garisson, Polynomials with large numbers of prime values, *Amer. Math. Monthly* **97** (1990), 316–317.

[9] K. Győry, On the irreducibility of neighbouring polynomials, *Acta Arith.* **67** (1994), 283–294.

[10] H. Halberstam and H.-E. Richert, Sieve methods, *Academic Press*, *London*, 1974.

[11] H. Heilbronn, Über die Verteilung der Primzahlen in Polynomen, *Math. Ann.* **104** (1931), 794–799.

[12] A. A. Karatsuba, Basic Analytic Number Theory, *Springer*, *New York*, 1993.

[13] Kevin S. McCurley, Prime values of polynomials and irreducibility testing, *Bull. Amer. Math. Soc. (N. S.)* **11** (1984), 155–158.

[14] Kevin S. McCurley, Polynomials with no small prime values, *Proc. Amer. Math. Soc.* **97** (1986), 393–395.

[15] Kevin S. McCurley, The smallest prime value of $x^n + a$, *Canad. J. Math.* **38** (1986), 925–936.

[16] R. A. Rankin, The difference between consecutive prime numbers, *V. Proc. Edinburgh Math. Soc.* (2) **13** (1962/1963), 331–332.

[17] P. Ribenboim, The Book of Prime Records, *Springer-Verlag*, *New York*, 1988.

[18] A. Schinzel and W. Sierpiński, Sur certaines hypothèses cocernant les nombres premiers, *Acta. Arithm.* **4** (1958), 185–208; erratum ibid. **5** (1959) 259.

[19] W. Sierpiński, Les binômes $x^2 + n$ et les nombres premiers, *Bull. Soc. Roy. Sci. Liège* **33** (1964), 259–260.

[20] S. Turjányi, Polynomials with many prime values, *Acta Acad. Paedagogicae Nyiregyhaziensis* **13**(d) (1992), 11–12.

IMRE Z. RUZSA
MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES
H-1364 BUDAPEST, P.O. BOX 127
HUNGARY

E-mail: ruzsa@renyi.hu

SÁNDOR TURJÁNYI
INSTITUTE OF MATHEMATICS AND INFORMATICS
UNIVERSITY OF DEBRECEN
H-4010 DEBRECEN, P.O. BOX 12
HUNGARY

E-mail: turjanyi@math.klte.hu

(Received November 15, 2002; revised February 12, 2003)