Canad. Math. Bull. Vol. **56** (4), 2013 pp. 844–849
http://dx.doi.org/10.4153/CMB-2012-021-8

On the Average Number of Square-Free Values of Polynomials

Igor E. Shparlinski

Abstract. We obtain an asymptotic formula for the number of square-free integers in N consecutive values of polynomials on average over integral polynomials of degree at most k and of height at most H, where $H \geq N^{k-1+\varepsilon}$ for some fixed $\varepsilon > 0$. Individual results of this kind for polynomials of degree $k > 3$, due to A. Granville (1998), are only known under the *ABC*-conjecture.

1 Introduction

For a polynomial $f(X) \in \mathbb{Z}[X]$ we denote by $S_f(N)$ the number of positive integers $n \leq N$ such that $f(n)$ is square-free.

It is natural to expect that

$$S_f(N) \sim c_f N \tag{1.1}$$

as $N \to \infty$, where

$$c_f = \prod_{p \text{ prime}} \Big(1 - \frac{\rho_f(p^2)}{p^2}\Big)$$

and $\rho_f(m)$ denotes the number of solutions to the congruence

$$f(n) \equiv 0 \pmod{m}, \qquad 1 \leq n \leq m$$

modulo an integer $m \geq 1$. However for polynomials of degree $k > 0$ only conditional results (under the *ABC*-conjecture) are known; see [2,3,5] and references therein.

A related question of studying the distribution of square-free parts of polynomial values (that is, square-free values of s in the representations $f(n) = r^2 s$ with $r \in \mathbb{Z}$) has also been studied in the literature [1,4].

Here we show that (1.1) holds unconditionally on average over the polynomials of naive height at most H, provided that

$$N^A \geq H \geq N^{k-1+\varepsilon} \tag{1.2}$$

with some fixed A and $\varepsilon > 0$.

For positive integers H and k, let $\mathcal{F}_k(H)$ denote the following family of polynomials over $\mathbb{Z}$ of degree at most k and of height at most H, that is,

$$\mathcal{F}_k(H) = \{a_0 + \cdots + a_k X^k \in \mathbb{Z}[X] \ : \ (a_0, a_1, \ldots, a_k) \in \mathcal{B}_k(H)\}.$$

Received by the editors December 17, 2011; revised February 14, 2012.
Published electronically June 22, 2012.
AMS subject classification: **11N32**.
Keywords: polynomials, square-free numbers.

where

$$\mathcal{B}_k(H) = \{(a_0, \dots, a_k) \in \mathbb{Z}^{k+1} \ : \ \gcd(a_0, a_1, \dots, a_k) = 1, \\ a_i \in \{0, \pm 1, \dots, \pm H\},\ i = 0, 1, \dots, k\}.$$

Theorem 1.1 *For a fixed $k \geq 2$, we have*

$$\frac{1}{\#\mathcal{F}_k(H)} \sum_{f \in \mathcal{F}_k(H)} |S_f(N) - c_f N| \leq N^{1/2} H^{o(1)} + N^{(k+1)/2} H^{-1/2+o(1)},$$

as $H \to \infty$.

We see from Theorem 1.1 that for any constants A and $\varepsilon > 0$ there is $\delta > 0$ such that

$$\frac{1}{\#\mathcal{F}_k(H)} \sum_{f \in \mathcal{F}_k(H)} \left|S_f(N) - c_f N\right| \ll N^{1-\delta}$$

provided H satisfies (1.2).

2 Notation

Throughout the paper, the letter p (possibly subscripted) always denotes a prime; k, m, and n always denote positive integers (as do K, M, and N).

Any implied constants in symbols O, $\ll$, and $\gg$ may occasionally depend, where obvious, on the integer parameter k and are absolute otherwise. We recall that the notations $U = O(V)$, $U \ll V$, and $V \gg U$ are all equivalent to the statement that $|U| \leq cV$ holds with some constant $c > 0$.

Finally, $U = o(V)$ means that $U \leq \psi(V)V$ for some function ψ such that $\psi(V) \to 0$ as $V \to \infty$.

3 Solutions to Polynomial Congruences of Small Height

Let $W_k(m, H, N)$ be the number of solution to the congruence

$$a_0 + a_1 n + \dots + a_k n^k \equiv 0 \pmod{m}, \tag{3.1}$$

in the variables

$$(a_0, \dots, a_k) \in \mathcal{B}_k(H) \qquad \text{and} \qquad n \leq N.$$

Trivially, we have

$$W_k(m, H, N) \ll H^k (H/m + 1) N. \tag{3.2}$$

We now obtain a stronger estimate for large values of m.

Lemma 3.1 *We have*

$$W_k(m, H, N) \leq \left(H^{k+1} N^k / m + H^k\right) (HN)^{o(1)}.$$

https://doi.org/10.4153/CMB-2012-021-8 Published online by Cambridge University Press

Proof We write the congruence (3.1) as the equation

$$a_0 + a_1 n + \cdots + a_k n^k = \lambda m,$$

with some integer λ with $|\lambda| \le (k+1)HN^k/m$.

We fix some a_0 with $|a_0| \le H$ and λ with $|\lambda| \le (k+1)HN^k/m$ and $a_0 - \lambda m \ne 0$. This can be done in $O((HN^k/m + 1)H)$ ways. Since $a_0 - \lambda m \ne 0$, from the bounds on the divisor function, we see that n can take at most $H^{o(1)}$ possible values. After this, $(a_1, \ldots, a_k)$ can take at most $O(H^{k-1})$ values. There are at most $\left(HN^k/m + 1\right) H^k (HN)^{o(1)}$ solutions of this type.

We now fix some a_0 with $|a_0| \le H$ and λ with $|\lambda| \le (k+1)HN^k/m$ with $a_0 - \lambda m = 0$. In this case, a_0 defines λ uniquely and $a_0 \equiv 0 \pmod m$. Hence this can be done in $O\left(H/m + 1\right)$ ways. For $O(H^k)$ possible choices of $(a_1, \ldots, a_k)$ we get $O(1)$ possible values for n.

Thus,

$$\begin{aligned} W_k(m, H, N) &\ll \left(HN^k/m + 1\right) H^k (HN)^{o(1)} + (H/m + 1)H^k \\ &\le \left(HN^k/m + 1\right) H^k (HN)^{o(1)}, \end{aligned}$$

and the result follows. ∎

4 Proof of Theorem 1.1

Clearly we can assume that $H \ge N$, as otherwise the bound is trivial.

Let $T_f(m, N)$ denote the number of solutions to the congruence

$$f(n) \equiv 0 \pmod m, \qquad 1 \le n \le N,$$

modulo an integer $m \ge 1$. Denoting by $\mu(m)$ the Möbius function, by the inclusion exclusion principle, for $f \in \mathcal{F}_k(H)$ we obtain

$$S_f(H) = \sum_{d \le \sqrt{kHN^k}} \mu(d) T_f(d^2, N).$$

We now fix some parameter D and write

$$S_f(H) = M_f(H) + O(R_f(H)),$$

where

$$M_f(H) = \sum_{d \le D} \mu(d) T_f(d^2, N) \quad \text{and} \quad R_f(H) = \sum_{D < d \le \sqrt{kHN^k}} T_f(d^2, N).$$

To estimate $M_f(H)$ we write

$$T_f(d^2, N) = \frac{N}{d^2} \rho_f(d^2) + O(\rho_f(d^2)). \tag{4.1}$$

https://doi.org/10.4153/CMB-2012-021-8 Published online by Cambridge University Press

Clearly, for any p, we have

$$\rho_f(p^2) \le kp \tag{4.2}$$

(as polynomials from $\mathcal{F}_k(H)$ are nontrivial modulo every prime p). Furthermore, using the Hensel lifting, we also see that if $p \nmid \Delta_f$, where Δ_f is the discriminant of f, then

$$\rho_f(p^2) \le k. \tag{4.3}$$

Therefore, combining (4.2) and (4.3), we see that for any square-free integer d, using the multiplicativity of $\rho_f(m)$, we have

$$\rho_f(d^2) \le d\rho_f(d) = d\prod_{p\mid d}\rho_f(p^2) \le k^{\omega(d)}\gcd(d,\Delta_f),$$

where $\omega(d)$ is the number of prime divisors of d. Using the well-known estimate on $\omega(d)$ we derive

$$\rho_f(d^2) \le d^{o(1)}\gcd(d,\Delta_f). \tag{4.4}$$

We now assume that $\Delta_f \ne 0$.

Thus, using (4.1) and (4.4), we obtain

$$M_f(H) = \sum_{d\le D}\mu(d)\Big(N\frac{\rho_f(d^2)}{d^2} + O\big(d^{o(1)}\gcd(d,\Delta_f)\big)\Big).$$

Using the bound (4.4) again and the multiplicativity of $\rho(m)$, we derive

$$\begin{aligned}\sum_{d\le D}\mu(d)\frac{\rho_f(d^2)}{d^2} &= \sum_{d=1}^{\infty}\mu(d)\frac{\rho_f(d^2)}{d^2} + O\Big(\sum_{d>D}\frac{\gcd(d,\Delta_f)}{d^{2+o(1)}}\Big)\\ &= c_f + O\Big(\sum_{d>D}\frac{\gcd(d,\Delta_f)}{d^{2+o(1)}}\Big).\end{aligned}$$

Therefore

$$M_f(H) = c_fN + O\Big(D^{o(1)}\sum_{d\le D}\gcd(d,\Delta_f) + N\sum_{d>D}\frac{\gcd(d,\Delta_f)}{d^{2+o(1)}}\Big). \tag{4.5}$$

We have

$$\sum_{d\le D}\gcd(d,\Delta_f) \le \sum_{e\mid\Delta_f}e\sum_{\substack{d\le D\\ e\mid d}}1 \le D\tau(\Delta_f) = D\Delta_f^{o(1)}, \tag{4.6}$$

https://doi.org/10.4153/CMB-2012-021-8 Published online by Cambridge University Press

where $\tau(m)$ is the number of integer positive divisors of m. Similarly,

$$\sum_{d>D} \frac{\gcd(d,\Delta_f)}{d^{2+o(1)}} \le \sum_{e|\Delta_f} e \sum_{\substack{d>D\\ e|d}} \frac{1}{d^{2+o(1)}} \le \sum_{e|\Delta_f} \frac{1}{e^{1+o(1)}} \sum_{\substack{d>D\\ e|d}} \frac{1}{(d/e)^{2+o(1)}}$$

$$\le \sum_{e|\Delta_f} \frac{1}{e^{1+o(1)}} \min\{(e/D)^{1+o(1)}, 1\}.$$

Considering the cases $e > D$ and $e \le D$, we see that each of $\tau(\Delta_f)$ terms in the last sum is $D^{-1+o(1)}$. Hence

$$\sum_{d>D} \frac{\gcd(d,\Delta_f)}{d^{2+o(1)}} \le D^{-1+o(1)} \Delta_f^{o(1)}. \tag{4.7}$$

Since $\Delta_f = H^{O(1)}$ for $f \in \mathcal{F}_k(H)$, substituting (4.6) and (4.7) in (4.5), we derive

$$M_f(H) = c_f N + O\left(DH^{o(1)} + ND^{-1}H^{o(1)}\right), \tag{4.8}$$

provided that $\Delta_f \ne 0$.

The sum $M_f(H)$ contributes to the main term. We now estimate error term $R_f(H)$ on average on $f \in \mathcal{F}_k(H)$.

Changing the order of summation, we see that

$$\sum_{f\in\mathcal{F}_k(H)} R_f(H) = \sum_{D<d\le\sqrt{kHN^k}} W_k(d^2, H, N).$$

We now choose another parameter $E > D$ and use (3.2) for $d \le E$ and Lemma 3.1 for $d > E$ getting

$$\sum_{f\in\mathcal{F}_k(H)} R_f(H) \ll H^{k+1}N/D + H^k NE + H^{k+1+o(1)}N^k/E + H^{k+o(1)}\sqrt{HN^k}.$$

Taking $E = H^{1/2}N^{(k-1)/2}$, we obtain

$$\sum_{f\in\mathcal{F}_k(H)} R_f(H) \ll H^{k+1}N/D + H^{k+1/2+o(1)}N^{(k+1)/2}. \tag{4.9}$$

Combining (4.8) and (4.9) and estimating the contribution from $O(H^k)$ polynomials $f \in \mathcal{F}_k(H)$ with $\Delta_f = 0$ trivially as $O(N)$, we derive

$$\sum_{f\in\mathcal{F}_k(H)} |S_f(N) - c_f N| \ll DH^{k+1+o(1)} + H^{k+1+o(1)}N/D + H^{k+1/2+o(1)}N^{(k+1)/2}.$$

Taking $D = N^{1/2}$, we conclude the proof.

https://doi.org/10.4153/CMB-2012-021-8 Published online by Cambridge University Press

5 Comments

Since the ultimate goal is obtaining nontrivial estimates on $S_f(N)$, as an intermediate step, it is certainly interesting to reduce the amount of averaging in the sums of Theorem 1.1. For small values of m it is easy to improve the bound (3.2) (for example via bounds of exponential sums). However, the lower limit in condition (1.2) is defined by the bound of Lemma 3.1 that has to be improved for any further progress.

It is also easy to see that for the set of monic polynomials

$$\mathcal{G}_k(H) = \{a_0 + \cdots + a_{k-1}X^{k-1} + X^k \in \mathbb{Z}[X] \ : \ (a_0 \ldots, a_{k-1}) \in \mathcal{B}_{k-1}(H)\}$$

of degree $k \geq 3$, the same method gives

$$\frac{1}{\#\mathcal{F}_k(H)} \sum_{f \in \mathcal{F}_k(H)} \left|S_f(N) - c_f N\right| \ll N^{1/2}H^{o(1)} + N^{(k-1)/2}H^{-1/2+o(1)},$$

which is nontrivial in the range $N^A \geq H \geq N^{k-2+\varepsilon}$.

Finally, the same method also applies to studying square-free values of multivariate polynomials, so one can get unconditional analogues of the result of Poonen [5] (which as in [2] relies on the *ABC*-conjecture) on average over multivariate polynomials of fixed degree and height at most H.

References

[1] P. Cutter, A. Granville, and T. J. Tucker, *The number of fields generated by the square root of values of a given polynomial.* Canad. Math. Bull. **46**(2003), no. 1, 71–79. http://dx.doi.org/10.4153/CMB-2003-007-0

[2] A. Granville, *ABC means we can count squarefrees.* Internat. Math. Res. Notices **19**(1998), no. 19, 991–1009.

[3] C. Hooley, *On the power-free values assumed by polynomial.* Hardy-Ramanujan J. **26**(2003), 30–55.

[4] F. Luca and I. E. Shparlinski, *Quadratic fields generated by polynomials.* Archiv Math. **91**(2008), no. 5, 399–408. http://dx.doi.org/10.1007/s00013-008-2656-2

[5] B. Poonen, *Squarefree values of multivariable polynomials.* Duke Math. J. **118**(2003), no. 2, 353–373. http://dx.doi.org/10.1215/S0012-7094-03-11826-8

Department of Computing, Macquarie University, NSW 2109, Australia
e-mail: igor.shparlinski@mq.edu.au

https://doi.org/10.4153/CMB-2012-021-8 Published online by Cambridge University Press