Product Representations of Polynomials

Jacques Verstraëte *

Abstract

For a fixed polyomial $f \in \mathbb{Z}[X]$, let $\rho_k(N)$ denote the maximum size of a set $A \subset \{1, 2, \ldots, N\}$ such that no product of k distinct elements of A is in the value set of f. In this paper, we determine the asymptotic behaviour of $\rho_k(N)$ for a wide class of polynomials. Our results generalize earlier theorems of Erdős, Sós and Sárközy.

1 Introduction

A polynomial $f \in \mathbb{Z}[X]$ has a *product representation in a set* A if $a_1 a_2 \ldots a_k = f(x)$ for some distinct $a_1, a_2, \ldots, a_k \in A$ and some $x \in \mathbb{Z}$. In other words, the value set $f(\mathbb{Z})$ of f contains some product of distinct elements of A. For a given polynomial $f \in \mathbb{Z}[X]$ and a positive integer k, we are interested in determining the maximum possible size $\rho_k(N; f) := \rho_k(N)$ of a set $A \subset \{1, 2, \ldots, N\}$ such that f has no product representation in A consisting of exactly k integers. This problem is a natural multiplicative analogue of the well-studied problem of representing f as a sum or difference of elements of A, and is motivated by a randomized factoring algorithm known as the *quadratic sieve*, introduced by Lenstra and Pomerance [13]. The study of product representations was initiated by Erdős [4], where the case $f(X) = X^2$ was considered. Erdős, Sós and Sárközy [5] proved that for $f(X) = X^2$ and a positive integer $k \geq 4$,

$$\rho_k(N) \sim \begin{cases} \pi(N) & \text{if } k \equiv 0 \bmod 4 \\ \pi(N) + \pi(N/2) & \text{if } k \equiv 2 \bmod 4 \end{cases} \tag{1}$$

where $\pi(N)$ denotes the number of primes less than or equal to N, and refinements of these results were given by Győri [8] and Sárközy [14]. The tightest results on the asymptotic behaviour of $\rho_k(N)$ may be found in [12].

In this paper, we study the asymptotic behaviour of $\rho_k(N)$ for all polynomials $f \in \mathbb{Z}[X]$. The following definition is required: a polynomial $f \in \mathbb{Z}[X]$ is *k-intersective* if the equation $f(x) = y^k$

*Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1. jverstraete@math.uwaterloo.ca

mod p has a solution for all choices of positive integers y and p. The asymptotic behaviour of $\rho_k(N)$ is divided into two regimes: very roughly speaking, we will show that $\rho_k(N)$ is linear in N or linear in $\pi(N)$, according to whether f is k-intersective. The first result we prove is for polynomials which are not k-intersective:

Theorem 1 *For any polynomial $f \in \mathbb{Z}[X]$ which is not k-intersective, $\rho_k(N)$ is linear in N. If f is irreducible, and of degree at least two, then $\rho_k(N) \sim N$.*

The first statement of the theorem is an immediate consequence of the definition of k-intersective polynomials, whereas the proof of the second requires some number theory. In our next theorem, we deal with polynomials which are k-intersective. The following notation is required: a *d-equipartition* of an integer n is a partition of n into parts of size at least d such that the largest part in the partition is as small as possible. For fixed positive integers d and $n \geq d$, we write $\|n\|$ for the size of the largest part in a d-equipartition of n.

Theorem 2 *Let $f \in \mathbb{Z}[X]$ be a k-intersective polynomial of prime degree d. Then $d \mid k$ and, if $k \geq d^3 - d$ or $d^2 \mid k$, then*

$$\rho_k(N) \;=\; \sum_{j=1}^{\|\frac{k}{d}\|-1} \pi\Big(\frac{N}{j}\Big) \;+\; O(N^{1-\frac{1}{2d}}). \tag{2}$$

Theorem 2 implies the results in (1), by taking $d = 2$. The exponent of the second order term in the theorem is almost best possible: we will prove that the second order term is at least of order $\Omega(N^{1-\frac{1}{d}+\epsilon})$. The exact order of magnitude of this second order term is likely to be difficult to determine. In the proof of Theorem 2, we will determine $\rho_k(N)$ up to a constant factor for all $k \geq d^2$, and it will follow from the Prime Number Theorem that the constant factor is at most about $1 + \frac{1}{\log d}$.

We leave the following as an open problem:

Conjecture 3 *Let $f \in \mathbb{Z}[X]$ and let k be a positive integer. Then, for some constant $\rho = \rho(k, f)$ depending only on k and f, either $\rho_k(N)/N \to \rho$ or $\rho_k(N)/\pi(N) \to \rho$ as $N \to \infty$.*

This paper is organized as follows: in Section 2, we prove Theorem 1. In Section 3, we classify k-intersective polynomials of prime degree, in preparation for the proof of Theorem 2. To obtain an idea of the proof, we give a relatively simple proof of a closely related statement, in Section 4. Additional material from extremal graph theory is required to prove Theorem 2, and we present this in Section 5. Finally, the proof of Theorem 2 is given in Section 6.

2 Proof of Theorem 1

The following elementary proposition shows that if f is not k-intersective, then $\rho_k(N)$ is linear in N, which is the first statement of Theorem 1.

Proposition 4 *Let k be a positive integer, and let $f \in \mathbb{Z}[X]$ be a polynomial which is not k-intersective. Then $\rho_k(N) = \Theta(N)$ as $N \to \infty$.*

Proof. If f is not k-intersective, then $f(x) = y^k \bmod p$ has no solution for some p. Therefore no product of k distinct integers congruent to zero modulo p is in the value set of f, so $\rho_k(N) \geq \lfloor \frac{N-y}{p} \rfloor$. This shows $\rho_k(N)$ is linear in N. ∎

To prove the second statement of Theorem 1, we require a fundamental theorem in class field theory, known as Chebotarev's Density Theorem (see Lenstra [10] for a discussion of this theorem). More precisely, we use the following well-known consequence of Chebotarev's Theorem: if the relative natural density of primes p such that a polynomial $f \in \mathbb{Z}[X]$ has a root modulo p is one, then f is reducible in $\mathbb{Z}[X]$.

Proposition 5 *Let k be a positive integer, and let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree at least two. Then, for all positive integers k, $\rho_k(N) \sim N$ as $N \to \infty$.*

Proof. By Chebotarev's Density Theorem, there exists a set P consisting of a positive relative density of all primes such that for each prime $p \in P$, f has no root mod p. In other words, for all x, no prime $p \in P$ is a factor of $f(x)$, for all x. By inclusion-exclusion, the set A of integers which have a prime factor in P has density

$$d(A) \;=\; 1 - \prod_{p \in P} \Big(1 - \frac{1}{p}\Big)$$

and no product of k distinct elements of A is in the value set of f. It is known (see Tenenbaum [15]) that if P has positive natural density, then $\sum_{p \in P} \frac{1}{p}$ diverges, so $d(A) = 1$. ∎

3 Classification of k-intersective polynomials

The classification of k-intersective polynomials is related to the following problem. Davenport (see Fried [7] page 286) conjectured that if f and g are polynomials with integer coefficients, and the value sets of f and g are equal modulo p, for all primes p, then f and g are linearly related – in other words there are integers a and b such that $f(X) = g(aX + b)$. Polynomials whose value sets are equal are known in the literature (see Fried [7] and also Müller and Völklein [11]) as *Davenport Pairs*. Davenport's conjecture remains open for general polynomials. Fried [6] extended Davenport's conjecture to the case where the value set of f contains the value set of g, and conjectured that such polynomials are linearly related. When $g(X) = X^k$, this is precisely saying that f is k-intersective. In this case, Fried's conjecture is that if $f \in \mathbb{Z}[X]$ is

k-intersective and of degree d, then $f = (X+a)^d$ or $f = (-X+a)^d$ for some integer a. Even in this special case, and even when $k = 2$, both Davenport's and Fried's conjectures are open. Fortunately, via Chebotarev's Theorem and Hilbert's Irreducibility Theorem [9], we can prove that if a k-intersective polynomial $f \in \mathbb{Z}[X]$ has prime degree d, then $d \mid k$ and $f = (X+a)^d$ or $f(X) = (-X+a)^d$ for some integer a. The proposition below was also proved by Fried [6], using a group theoretic approach; we use Hilbert's Irreducibility Theorem.

Proposition 6 *Let $f \in \mathbb{Z}[X]$ be a polynomial of degree d, where d is prime, and suppose that f is k-intersective. Then $d \mid k$ and $f = (X+a)^d$ or $f = (-X+a)^d$ for some integer a.*

Proof. By Chebotarev's Density Theorem, if a polynomial in $\mathbb{Z}[X]$ has a root modulo p for all primes p, then that polynomial is reducible in $\mathbb{Z}[X]$. Applying this to $f(X) - y^k$, we see that $f(X) - y^k$ is reducible in $\mathbb{Z}[X]$ for all integers y. Now Hilbert's Irreducibility Theorem [9] states that if a polynomial $h \in \mathbb{Z}[X,Y]$ is irreducible then there are infinitely many specializations $y \in \mathbb{Z}$ such that $h(X,y)$ is irreducible in $\mathbb{Z}[X]$. We conclude that $f(X) - Y^k$ is reducible in $\mathbb{Z}[X,Y]$. It is not hard to see that this can only happen when $f(X) = g(X)^d$ for some integer $d > 1$ where $d \mid k$ and some $g(X) \in Z[X]$. Note that we have not used the primality of d yet. Now since f has prime degree, the only possibility is that d is prime and g is linear, say $g(X) = (cX + a)$. Now $|c| = 1$, otherwise one of the equations $f(x) = 0 \mod |c|$ and $f(x) = 1 \mod |c|$ has no solution. ∎

The proof of Proposition 6 shows, more generally, that if a polynomial $f \in \mathbb{Z}[X]$ of degree d is k-intersective, then $d \mid k$ and $f = g^d$ for some polynomial $g \in \mathbb{Z}[X]$. This shows that the condition $d \mid k$ in Theorem 2 is necessary. In line with Fried's conjecture, we conjecture that every k-intersective polynomial of degree d is a dth power:

Conjecture 7 *Let $f \in \mathbb{Z}[X]$ be a k-intersective polynomial of degree d. Then $f(X) = (X+a)^d$ or $f(X) = (-X+a)^d$ for some integer a.*

For the remainder of the paper, we wish to estimate $\rho_k(N)$ when f is k-intersective of prime degree $d \mid k$. Since $\rho_k(N)$ is invariant under translation of variables in the polynomial f, we will assume that $f(X) = X^d$ for the remainder of the paper.

4 Sets with no product representations

Let $\rho(N; f) := \rho(N)$ denote the maximum size of a set $A \subset \{1, 2, \ldots, N\}$ with no product representation of a polynomial $f \in \mathbb{Z}[X]$. The difference between this problem and determining $\rho_k(N; f)$ is that the number of factors in the product representation – namely k – is not specified. For example, the reader will observe that the results of Section 2 show that $\rho(N; f) \sim N$ when f is irreducible and of degree at least two. In this section, we determine the asymptotic behaviour

of $\rho(N)$ up to an additive term of order $\pi(N^{\frac{1}{2}})$ for k-intersective polynomials of prime degree, d. From the last section, we may assume $f(X) = X^d$ in this case.

Theorem 8 *Let d be a prime number and $f(X) = X^d$. Then*

$$\rho(N) \;=\; \sum_{j=1}^{d-1} \pi\Big(\frac{N}{j}\Big) + O(\pi(N^{\frac{1}{2}})).$$

The case $d = 2$ of this theorem is very straightforward; there $\rho(N) = \pi(N)$ for all N. For each element $a \in A$ of a set A with no product representations of $f(X) = X^2$ can be assumed squarefree, and then the set of vectors $v(a)$ such that the pth entry of $v(a)$ is the pth valuation of a, for $a \in A$ and p prime, is a linearly independent set of vectors over $\mathbb{F}_2$. Since each vector has $\pi(N)$ entries, it follows that $|A| \leq \pi(N)$, and clearly the primes achieve equality. For the proof of Theorem 8, we require the following special case of Chevalley's Theorem (see Cassels [3]):

Theorem 9 *Let $\mathbb{F}$ be a finite field, and let $f_1, f_2, \ldots, f_n$ be polynomials in a total of m variables over $\mathbb{F}$, such that the zero vector is a common root of $f_1, f_2, \ldots, f_n$, and*

$$\sum_{i=1}^{n} deg(f_i) \;<\; m.$$

Then the polynomials $f_1, f_2, \ldots, f_n$ have a non-zero common root.

Proof of Theorem 8. Let $A \subset \{1, 2, \ldots, N\}$ be a set with no product representation of $f(X) = X^d$. Without loss of generality, N is a dth power. Let $n_i = d^{-i}N$ and $m_i = \min\{d^{i+1}, N^{\frac{1}{2}}\}$, and let $l(a)$ denote the largest prime factor of an integer a. For $0 \leq i \leq \log_d N$, let

$$A_i = \{a \in A : n_{i+1} < l(a) \leq n_i\}.$$

To each prime $p \in \{n_{i+1} + 1, \ldots, n_i\} \cup \{1, 2, \ldots, m_i\}$, we associate a polynomial f_p over the integers modulo d, defined as follows:

$$f_p \;=\; \sum_{j \in A_i} v_p(j) x_j^{d-1},$$

where $v_p(j)$ is the p-adic valuation of j. Note that the total number of variables in the polynomials f_p is exactly $|A_i|$, and the total number of polynomials is $\pi(n_i) - \pi(n_{i+1}) + \pi(m_i)$. If the polynomials f_p have a non-trivial common root, say $f_p(x) = 0$ for all p, then let $B = \{j \in A_i : x_j \neq 0\}$. For every prime $p \in \{n_{i+1} + 1, \ldots, n_i\} \cup \{1, 2, \ldots, m_i\}$,

$$\sum_{j \in B} v_p(j) \;\equiv\; 0 \mod d.$$

For any prime factor $p \neq l(b)$ of $b \in B$, $p \in \{1, 2, \ldots, m_i\}$. We conclude that $\prod_{b \in B} b$ is a dth power. This contradiction shows that the f_p have no non-trivial common root. In order for this to happen, by Theorem 9, the number of variables x_j for $j \in A_i$ satisfies $|A_i| \leq (d-1)\left[\pi(n_i) - \pi(n_{i+1}) + \pi(m_i)\right]$. Therefore

$$\sum_{i=1}^{\log_d N} |A_i| \;\leq\; (d-1)\pi\left(\frac{N}{d}\right) + O(\pi(N^{\frac{1}{2}})). \tag{3}$$

Now observe that all elements of A_0 are a product of a prime greater than N/d and an integer less than d. Let

$$A_{0j} = \{a \in A_0 : N/(j+1) < l(a) \leq N/j\},$$

for $j = 1, 2, \ldots, d-1$. Then $|A_{0j}| \leq j[\pi(N/j) - \pi(N/(j+1))]$, and therefore

$$|A_0| \;=\; \sum_{j=1}^{d-1} |A_{0j}| \;\leq\; \sum_{j=1}^{d-1} \pi\Big(\frac{N}{j}\Big) - (d-1)\pi\Big(\frac{N}{d}\Big). \tag{4}$$

Putting together the bounds (3) and (4), we obtain, as required,

$$\rho(N) \;\leq\; \sum_{j=1}^{d-1} \pi\Big(\frac{N}{j}\Big) + O(\pi(N^{\frac{1}{2}})).$$

The lower bound on $\rho(N)$ is proved via a construction. Let A^* consist of all integers less than or equal to N of the form pj, where $j \in \{1, 2, \ldots, d-1\}$ and $p \in \{d, d+1, \ldots, N\}$ is a prime. To see that no product of distinct integers in A^* is a dth power, observe that each prime $p \geq d$ dividing some a_i satisfies $p^d \mid a_1 a_2 \ldots a_k$, so p must divide d of the a_is. But then two of those a_is are identical, by definition of A^*, which is a contradiction. Therefore no product of distinct elements of A^* is a dth power. Finally,

$$\rho(N) \;\geq\; |A^*| \;=\; \sum_{j=1}^{d-1} \pi\Big(\frac{N}{j}\Big) - (d-1)\pi(d-1)$$

and this completes the proof of Theorem 8. ∎

Remarks. It would be interesting to determine the order of magnitude of

$$\rho(N) - \sum_{j=1}^{d-1} \pi(\tfrac{N}{j})$$

in Theorem 8, and perhaps it is always at most a constant. We also remark that if every integer in the set A in the proof of Theorem 8 is n-smooth [13] – in other words, the prime factors of every integer in A are less than n – then we obtain the bound $|A| \leq (d-1)\pi(n)$, which is stronger than Theorem 8 when n is much less than N. It would be interesting to see if this fact is of any use in factoring algorithms when $d > 2$; the quadratic sieve uses the case $d = 2$.

5 Nonzero k-Sums

Let $\mathbb{F}$ be a finite field. The *weight* of a vector $v \in \mathbb{F}^n$, denoted $\omega(v)$, is the number of non-zero co-ordinates of v. Let $\mathbb{B}_r$ denote the set of vectors with at most r non-zero co-ordinates in $\mathbb{F}^n$. In this section, we are concerned with the problem of finding the maximum possible size of a set of vectors $\mathcal{E} \subset \mathbb{B}_r$ such that the sum of any k distinct vectors in $\mathcal{E}$ is non-zero – we say that $\mathcal{E}$ has *non-zero k-sums*. The case $\mathbb{F} = \mathbb{F}_2$ was studied in [12], and we extend the analysis to all finite fields as follows:

Theorem 10 *Let $\mathbb{F}$ be a finite field of characteristic q, and let $k \geq q^2$ be a positive integer with $q \mid k$. Let $\mathcal{E} \subset \mathbb{F}^n$ be a set of vectors of weight at most r with non-zero k-sums. Then*

$$|\mathcal{E}| \;\leq\; \tfrac{2k}{q}(MN^{1-\frac{1}{q}} + N) \quad \textit{where } M = |\mathbb{B}_{\lfloor \frac{r}{2} \rfloor}| \textit{ and } N = |\mathbb{B}_{\lceil \frac{r}{2} \rceil}|. \tag{5}$$

Furthermore, for any $x \neq 0$, there exists a set $\mathcal{E}^* \subset \{0, x\}^n \subset \mathbb{B}_r$, with non-zero l-sums for all $l \leq k$, such that $|\mathcal{E}^*| = \Omega(MN)^{1-1/q+(q-1)/q(k-1)}$.

We will give a reduction of Theorem 10 to extremal graph theory, by applying the lemma below, which can be deduced from Theorem 2.2 in Alon, Krivelevich and Sudakov [1]:

Lemma 11 *Let $G = (X, Y; E)$ be a bipartite graph which does not contain a bipartite q-regular subgraph with s vertices in each part. Then $|E| \leq s|X||Y|^{1-\frac{1}{q}} + (s-1)|Y|$.*

Proof of Theorem 10. We start with the upper bound, namely (5). For each vector $v \in \mathcal{E}$, consider a partition of v into two vectors v_1 and v_2 where $\omega(v_1) \leq \omega(v_2) \leq \omega(v_1) + 1$. We may consider these partitions as edges in an auxilliary graph whose vertex set is $\mathbb{B}_{r/2}$ when r is even, and an M by N bipartite graph with parts $\mathbb{B}_{\lfloor r/2 \rfloor}$ and $\mathbb{B}_{\lceil r/2 \rceil}$ when r is odd. In these graphs, two vectors v_1, v_2 are joined by an edge if their concatenation is a vector $v \in \mathcal{E}$ and they form the chosen partition (v_1, v_2) of v. These two graphs have exactly $|\mathcal{E}|$ edges, and do not contain a copy of any bipartite graph with k edges and every vertex of degree zero modulo q, since $\mathbb{F}$ has characteristic q. If r is odd, then by Lemma 11 with $s = k/q$, we obtain

$$|\mathcal{E}| \leq sMN^{1-1/q} + (s-1)N,$$

as required. If r is even, then the auxilliary graph has a bipartite subgraph containing at least half its edges, to which Lemma 11 may be applied. This proves (5).

The construction of $\mathcal{E}^*$ is via the first moment method. Since this is now a fairly standard approach, we do not include all the calculations. Consider a random collection $\mathcal{E} \subset \{0, x\}^n \subset \mathbb{B}_r \backslash \mathbb{B}_{r-1}$, where each vector is chosen independently with probability

$$p \;=\; n^{-\frac{r}{q}+\frac{r(q-1)}{q(k-1)}}(rk)^{-q}(4r)^{-\frac{r}{k}}.$$

Let Y and Z be the number of sets of at most k vectors in $\mathcal{E}$ adding up to zero, and the number of vectors in $\mathcal{E}$, respectively. Then $\mathbb{E}[Z] = p\binom{n}{r}$, and a short calculation gives $16\mathbb{E}[Y] < \mathbb{E}[Z]$. Using Markov's inequality and concentration of Z (which has a binomial distribution), we deduce that

$$\mathbb{P}[Z > 2Y \quad \wedge \quad 2Z > \mathbb{E}[Z]] > 0.$$

So we can find $\mathcal{E}$ such that $Z > 2Y$ and $2Z > \mathbb{E}[Z]$. Now we delete all vectors in $\mathcal{E}$ which appear in at least one subset of at most k vectors adding up to zero mod q, to obtain $\mathcal{E}^* \subset \mathcal{E}$. ∎

Remarks. The proof of Theorem 10 is a reduction of the non-zero k-sum problem to extremal graph theory. If $r = 2$ and $k \geq q! + q$, then the existence of norm-graphs (see Alon, Rónyai and Szábo [2]) shows that the upper bound in Lemma 11 and Theorem 10 is tight – consider the incidence vectors of the edges of a norm-graph. However the problem of determining the maximum number of edges in a graph not containing $K_{q,q}$ is a notoriously difficult problem, known as Zarankiewicz's Problem [16], and it is likely that determining the maximum size of a set $\mathcal{E} \subset \mathbb{B}_r$ with non-zero k-sums is even more difficult.

6 Proof of Theorem 2

By the results of Section 3, if f is k-intersective, then $d \mid k$ and we may take $f(X) = X^d$. The proof of Theorem 2 is split into two parts. We begin by showing that (2) is a lower bound for $\rho_k(N)$ when $k \geq d^3 - d$ or $d^2 \mid k$. Thereafter, we show that (2) is an upper bound for $\rho_k(N)$ for all values of $k \geq d^2$. This gives the desired equality in Theorem 2 when $k \geq d^3 - d$ or $d^2 \mid k$.

6.1 Proof of Theorem 2 : A Lower Bound on $\rho_k(N)$

For $k \geq d^3 - d$ or $d^2 \mid k$, we observe that $J = \|\frac{k}{d}\| \in \{d, d+1\}$. To prove (2), we construct a set in $\{1, 2, \ldots, N\}$ without product representations of X^d with k factors. We start with the set B of integers less than or equal N of the form pj, where $p > N^{1/3} > J$ is prime and $j \in \{1, 2, \ldots, J-1\}$. Then

$$|B| \quad \geq \quad \sum_{j=1}^{J-1} \pi\Big(\frac{N}{j}\Big) - (J-1) \cdot \pi(N^{\frac{1}{3}}). \tag{6}$$

Claim 1. No product of k distinct integers in B is a dth power.

Proof. First suppose $J = d$. If $b_1 b_2 \ldots b_k = x^d$ for some x and $b_i \in B$, then $p^d \mid b_1 b_2 \ldots b_k$ for every prime $p \mid b_1 b_2 \ldots b_k$. This means that $b_i = b_j$ for some $i \neq j$, as required. Now suppose $J = d + 1$ – then $d^2 \nmid k$. If $b_1 b_2 \ldots b_k = x^d$ where $b_i \in B$, then for each prime $p \geq N^{1/3}$ dividing x, we have $p^d \mid b_1 b_2 \ldots b_k$. This means that there are exactly d values of i such that $p \mid b_i$. In particular, jp is one of the integers in the product, for all $j \in \{1, 2, \ldots, d\}$, and $b_1 b_2 \ldots b_k$ has

exactly $\frac{k}{d}$ distinct prime factors which are at least $N^{1/3}$. Divide $b_1 b_2 \dots b_k = x^d$ by the product of these prime factors. Then we are left with the equation

$$d!^{\frac{k}{d}} \;=\; y^d$$

for some integer y. Now there is a prime p such that $d/2 < p \le d$, by Bertrand's Postulate. But

$$v_p(d!^{\frac{k}{d}}) = \frac{k}{d} \quad \text{and} \quad v_p(y^d) \equiv 0 \mod d,$$

and since $d^2 \nmid k$, this is a contradiction. This proves Claim 1. $\square$

Let $\mathbb{F} = \mathbb{Z}/d\mathbb{Z}$, and $n = \pi(N^{\frac{1}{3}}) - \pi(J)$. Let

$$\mathcal{E}^* \subset \{0,1\}^n \subset \mathbb{F}^n$$

be a set of vectors of weight three such that no sum of at most k vectors in $\mathcal{E}^*$ is zero, and suppose $\mathcal{E}^*$ has maximum possible size. We index the co-ordinates of vectors in $\mathcal{E}$ by the prime numbers in $\{J, J+1, \dots, \lfloor N^{\frac{1}{3}} \rfloor\}$. According to Theorem 10 with $d = q$ and $r = 3$, we can choose $\mathcal{E}^*$ so that

$$|\mathcal{E}^*| \;=\; \Omega(n^{3-\frac{3}{d}+\frac{3(d-1)}{d(k-1)}}) \;=\; \tilde{\Omega}(N^{1-\frac{1}{d}+\frac{d-1}{d(k-1)}}). \tag{7}$$

To each $v \in \mathcal{E}^*$ we associate the integer $c(v) = pqr$, where $v_p = v_q = v_r = 1$ and p, q, r are distinct primes. Let $C \subset \{1, 2, \dots, N\}$ be the set of integers $c(v)$ for $v \in \mathcal{E}^*$ – then $|\mathcal{E}|^* = |C|$.

Claim 2. No product of k distinct integers in $B \cup C$ is a dth power.

Proof. Consider the equation $a_1 a_2 \dots a_k = x^d$, where x is an integer and $a_1, a_2, \dots, a_k \in B \cup C$ are distinct. By Claim 1, we have $a_i \in C$ for some i, and so $a_i = c(v) = pqr$ for some $v \in \mathcal{E}^*$ with $v_p = v_q = v_r = 1$. The integers p, q and r are primes in $\{1, 2, \dots, \lfloor N^{\frac{1}{3}} \rfloor\}$. Since $p \mid x^d$ we have $p^d \mid x^d$, and therefore $v_p = 1$ for zero mod d vectors $v \in \mathcal{E}^*$, since each a_j has three distinct prime factors. This is valid for all primes $p \in \{J, J+1, \dots, \lfloor N^{\frac{1}{3}} \rfloor\}$ which divide an a_j in the product $a_1 a_2 \dots a_k$. This contradicts the fact that no sum of at most k vectors in $\mathcal{E}^*$ is zero modulo d, and proves Claim 2. $\square$

Finally, we combine (6) and (7) to obtain

$$|B \cup C| \;=\; \sum_{j=1}^{J-1} \pi\left(\frac{N}{j}\right) \;+\; \tilde{\Omega}\Big(N^{1-\frac{1}{d}+\frac{d-1}{d(k-1)}}\Big).$$

This gives the lower bound on $\rho_k(N)$ required for Theorem 2. $\blacksquare$

6.2 Proof of Theorem 2 : An Upper Bound on $\rho_k(N)$

Let $A \subset \{1, 2, \ldots, N\}$ be a set such that no product of k distinct elements of A is a dth power. We will prove (2) by showing, more generally, that for all $k \geq d^2$ such that $d \mid k$,

$$|A| \leq \sum_{j=1}^{\|\frac{k}{d}\|-1} \pi\Big(\frac{N}{j}\Big) + O(N^{1-\frac{1}{2d}}). \tag{8}$$

It is convenient to put $J = \|\frac{k}{d}\|$ and $n = \lfloor \frac{1}{2}\log_2 N - \log_2 J \rfloor$. For $0 \leq i \leq n$, let

$$\begin{aligned} X_i &= \{1, 2, \ldots, J2^{i+1}\} \\ Y_i &= \{p \text{ prime } : \tfrac{N}{J2^{i+1}} < p \leq \tfrac{N}{J2^i}\}. \end{aligned}$$

Form a bipartite graph $G_i = (X_i, Y_i; E_i)$ where

$$E_i = \{xy : x \in X_i, y \in Y_i, xy \in A\}.$$

Then G_i does not contain subgraph with k edges such that every vertex has degree congruent to zero mod d, otherwise the product of the integers in A corresponding to edges in the subgraph is a dth power. Let $(k_1, k_2, \ldots, k_t)$ be a d-equipartition of $\frac{k}{d}$. Then $J = \max\{k_j : 1 \leq j \leq t\} = J$ by definition of $J = \|\frac{k}{d}\|$. Let $\mathcal{H}$ denote the family of bipartite graphs which comprise an edge-disjoint union of t bipartite d-regular graphs, such that the ith graph in the union has parts of size k_i, and $k_1 + k_2 + \cdots + k_t = \frac{k}{d}$. Since $k \geq d^2$, $\mathcal{H}$ is non-empty. Then, for all $H \in \mathcal{H}$, $H \not\subset G_i$.

Claim 1. For all $i \leq n$,

$$|E_i| \leq J \cdot |X_i||Y_i|^{1-\frac{1}{d}} + (J-1)|Y_i| + k. \tag{9}$$

Proof. Suppose that the claim is false. By Lemma 11, with $s = J$, G_i contains a d-regular bipartite subgraph H_1 with k_1 vertices in each part. Remove the edges of H_1 from E_i, to get a new graph G_i'. Applying Lemma 11, again G_i' has enough edges to guarantee a subgraph H_2 with k_2 vertices in each part. We continue this procedure t times to obtain graphs $H_1, H_2, \ldots, H_t$ (this is possible since after each stage, we have not deleted more than k edges, and (9) exceeds the bound in Lemma 11 by k). We have produced a graph $H \in \mathcal{H}$, namely $H_1 \cup H_2 \cup \cdots \cup H_t$, which is contained in G_i. This contradiction proves (9). □

For the next claim, let $l(a)$ be the largest prime factor of $a \in A$.

Claim 2. Let $A_0 = \{a \in A : l(a) \geq N^{\frac{1}{2}}\}$. Then

$$|A_0| \leq \sum_{j=1}^{J-1} \pi\left(\tfrac{N}{j}\right) + O(N^{1-\frac{1}{2d}}). \tag{10}$$

Proof. Let $A_{0i} = \{a \in A_0 : l(a) \in Y_i\}$ for $0 \le i \le n$. By Claim 1,

$$\begin{aligned}\sum_{i=0}^{n} |A_{0i}| &\le \sum_{i=0}^{n} |E_i| \le \sum_{i=0}^{n} \Big\{ J|X_i||Y_i|^{1-\frac{1}{d}} + (J-1)|Y_i| + k \Big\} \\ &= \sum_{i=0}^{n} O(2^{\frac{i}{d}} \cdot N^{1-\frac{1}{d}}) + (J-1)\cdot \pi(\tfrac{N}{J}) + k(n+1) \\ &= O(N^{1-\frac{1}{2d}}) + (J-1)\cdot \pi(\tfrac{N}{J}).\end{aligned}$$

For each $j \le J-1$, define $B_j = \{a \in A_0 : N/(j+1) < l(a) \le N/j\}$. Then

$$|B_j| \le j[\pi(\tfrac{N}{j}) - \pi(\tfrac{N}{j+1})].$$

Together with the bounds for A_{0i} above, we obtain

$$\begin{aligned}|A_0| &= \sum_{j=1}^{J-1} |B_j| + \sum_{i=0}^{n} |A_{0i}| \\ &\le \sum_{j=1}^{J-1} j\left[\pi\left(\tfrac{N}{j}\right) - \pi\left(\tfrac{N}{j+1}\right)\right] + O(N^{1-\frac{1}{2d}}) + (J-1)\cdot\pi\left(\tfrac{N}{J}\right) \\ &= \sum_{j=1}^{J-1} \pi\left(\tfrac{N}{j}\right) + O(N^{1-\frac{1}{2d}}).\end{aligned}$$

This proves Claim 2. □

The prime factors of each element of $A\backslash A_0$ are less than $N^{\frac{1}{2}}$, by definition of A_0. It was observed in [5] that each such integer admits a factorization into two positive integers, each at most $N^{2/3}$. For each $a \in A\backslash A_0$, we choose exactly one such factorization, say $x_a y_a$, where $1 \le y_a \le x_a$. Let

$$A_1 = \{a \in A\backslash A_0 : y_a \le x_a \le N^{\frac{1}{2}}\}.$$

Then the bipartite graph $G = (X, Y; E)$ where $X = \{x_a : a \in A_1\}$ and $Y = \{y_a : a \in A_1\}$ does not contain any graph in $\mathcal{H}$. By Lemma 11, applied in the same way as in Claim 1,

$$|A_1| \le |E| \le J|X||Y|^{1-\frac{1}{d}} + (J-1)|Y| = O(N^{1-\frac{1}{2d}}). \tag{11}$$

Let $A_2 = A\backslash(A_1 \cup A_0)$, and let

$$\begin{aligned}X_i &= \{1, 2, \ldots, 2^i N^{1/3}\} \\ Y_i &= \{1, 2, \ldots, 2^{-i} N^{2/3}\}\end{aligned}$$

where $0 \le i \le m$ and $m = \lceil \frac{1}{6}\log_2 N \rceil$. Form a bipartite graph $F_i = (X_i, Y_i; E_i)$ where E_i is the set of pairs $\{x_a, y_a\}$ where $x_a \in X_i$ and $y_a \in Y_i$, and $a \in A_2$. Then F_i does not contain any

subgraph in $\mathcal{H}$. Therefore the appropriate analog of (9) holds, and

$$\begin{aligned} |A_2| &\leq \sum_{i=0}^{m} \Big\{ J|X_i||Y_i|^{1-\frac{1}{d}} + (J-1)|Y_i| + k \Big\} \\ &= \sum_{i=0}^{m} O(2^i N^{\frac{1}{3}} (2^{-i} N^{\frac{2}{3}})^{1-\frac{1}{d}}) + O((J-1)N^{\frac{2}{3}}) + k(m+1) \\ &= \sum_{i=0}^{m} O(2^{\frac{i}{d}} N^{1-\frac{2}{3d}}) \\ &= O(N^{1-\frac{1}{2d}}). \end{aligned} \tag{12}$$

Since $A = A_0 \cup A_1 \cup A_2$, we may add (10), (11) and (12) to obtain (8). ∎

Remarks. The proof of Theorem 2 given above shows that for all $k \geq d^2$,

$$\sum_{j=1}^{K_1} \pi(\tfrac{N}{j}) \lesssim \rho_k(N) \lesssim \sum_{j=1}^{K_2} \pi(\tfrac{N}{j})$$

where $K_1 = d-1$ if $d^2 \mid k$ and $K_1 = d$ if $d^2 \nmid k$, and $K_2 = \|k/d\| - 1$. Since $\|k/d\| \leq 2d-1$ for all k such that $d \mid k$, the above inequalities determine $\rho_k(N)$ for all $k \geq d^2$ up to a factor of about $1 + \frac{1}{\log d}$, by the Prime Number Theorem.

7 Concluding Remarks

• In line with Fried's conjecture (see Section 3), we conjecture that if $f \in \mathbb{Z}[X]$ is any k-intersective polynomial of degree d, then $f(X) = (X+a)^d$ or $f(X) = (-X+a)^d$ for some integer a. This is open even in the case $d = 2$.

• Theorem 2 gives the asymptotic behaviour of $\rho_k(N)$ for any k-intersective polynomial of prime degree, provided k is relatively large. We conjecture that for any polynomial $f \in \mathbb{Z}[X]$, $\rho_k(N) \sim \rho N$ or $\rho_k(N) \sim \rho\pi(N)$, where $\rho > 0$ depends only on k and f. It would be interesting to determine the value of ρ.

• We defined (see Section 4) $\rho(N)$ to be the maximum size of a set $A \subset \{1, 2, \ldots, N\}$ with no product representation of a polynomial f. In the case $f(X) = X^2$, $\rho(N) = \pi(N)$. It would be interesting to determine $\rho(N)$ precisely for $f(X) = X^d$ and $d > 2$. Perhaps, in this case,

$$\rho(N) = \sum_{j=1}^{d-1} \pi(\tfrac{N}{j}) + O(1).$$

• In addition, we showed (Section 2) that $\rho(N) \sim N$ when f is irreducible and of degree at least two. The asymptotic behaviour of $\rho(N)$ when f is reducible and $f(X) \neq X^d$ is left as an open question.

Acknowledgments. I would like to give special thanks to Josh Greene, László Lovász, and Assaf Naor for helpful comments, as well as an anonymous referee for pointing out corrections to an earlier draft.

References

[1] Alon, N., Krivelevich, M., Sudakov, B. Turán numbers of bipartite graphs and related Ramsey-type questions. Combinatorics, Probability and Computing 12 (2003), 477-494.

[2] Alon, N,. Rónyai, L., Szabó, T. Norm-graphs: variations and applications. J. Combin. Theory Ser. B 76 (1999), no. 2, 280–290.

[3] Cassels, J. W. S. and Scott, J. W. Local Fields. Cambridge University Press, 1986.

[4] Erdős, P. An application of graph theory to combinatorial number theory. (1969)

[5] Erdős, P., Sárközy, A., Sós, V. T. On product representations of powers. I. *European J. Combin.*, 16 (6) 567–588, 1995.

[6] Fried, M. D. Arithmetic properties of value sets of polynomials, Acta Arithmetica XV (1969) 91–115.

[7] Fried, M. D., Jarden, M. Field Arithmetic. Springer Verlag, New York, 1986.

[8] Györi, Ervin, Bipartite graphs and product representation of squares, Graphs and combinatorics (Marseille, 1995), Discrete Math., 165/166 (1997) 371–375.

[9] Hilbert, David, Über die Irreduzibilität ganzer rationaler Functionen, mit gannzzahliger Koeffizienten, Journal f. Reine und Angew. Math. 110 (1892) 104–129.

[10] Lenstra, H. The Chebotarev Density Theorem. `http://math.berkeley.edu/~jvoight/notes/oberwolfach/Lenstra-Chebotarev.pdf`.

[11] Müller, P., Völklein, H. On a question of Davenport. J. Number Theory 58 (1996) 1 46–54.

[12] Naor, A., Verstraëte, J. Parity check matrices and product representations of squares. Submitted to Combinatorica (2005).

[13] Pomerance, C. A tale of two sieves. Notices Amer. Math. Soc. 43 (12) 1473–1485 (1996).

[14] Sárközy, G. N. Cycles in bipartite graphs and an application in number theory, J. Graph Theory, 19 (1995) 3 323–331.

[15] Tenenbaum, G. Introduction to analytic and probabilistic number theory. Cambridge studies in advanced mathematics, 46. Cambridge University Press, 1995.

[16] Zarankiewicz, K. The solution of a certain problem on graphs of P. Turan. Bull. Acad. Polon. Sci. Cl. III. 1, (1953). 167–168.