COMBINATORICA
Bolyai Society – Springer-Verlag

Combinatorica 1–46

ON THE DEGREE OF UNIVARIATE POLYNOMIALS OVER THE INTEGERS

GIL COHEN, AMIR SHPILKA*, AVISHAY TAL

Received June 9, 2012
Revised September 9, 2014

We study the following problem raised by von zur Gathen and Roche [6]:

What is the minimal degree of a nonconstant polynomial $f\colon \{0,\dots,n\}\to\{0,\dots,m\}$*?*

Clearly, when $m=n$ the function $f(x)=x$ has degree 1. We prove that when $m=n-1$ (i.e. the point $\{n\}$ is not in the range), it must be the case that $\deg(f)=n-o(n)$. This shows an interesting threshold phenomenon. In fact, the same bound on the degree holds even when the image of the polynomial is any (strict) subset of $\{0,\dots,n\}$. Going back to the case $m=n$, as we noted the function $f(x)=x$ is possible, however, we show that if one excludes all degree 1 polynomials then it must be the case that $\deg(f)=n-o(n)$. Moreover, the same conclusion holds even if $m=O(n^{1.475-\epsilon})$. In other words, there are no polynomials of intermediate degrees that map $\{0,\dots,n\}$ to $\{0,\dots,m\}$.

Furthermore, we give a meaningful answer when m is a large polynomial, or even exponential, in n. Roughly, we show that if $m<\binom{n/c}{d}$, for some constant c, and $d\le 2n/15$, then either $\deg(f)\le d-1$ (e.g., $f(x)=\binom{x-n/2}{d-1}$ is possible) or $\deg(f)\ge n/3-O(d\log n)$. So, again, no polynomial of intermediate degree exists for such m. We achieve this result by studying a discrete version of the problem of giving a lower bound on the minimal L^∞ norm that a monic polynomial of degree d obtains on the interval $[-1,1]$.

We complement these results by showing that for every integer $k=O(\sqrt{n})$ there exists a polynomial $f\colon \{0,\dots,n\}\to\{0,\dots,O(2^k)\}$ of degree $n/3-O(k)\le\deg(f)\le n-k$.

Our proofs use a variety of techniques that we believe will find other applications as well. One technique shows how to handle a certain set of diophantine equations by working modulo a well chosen set of primes (i.e., a Boolean cube of primes). Another technique shows how to use lattice theory and Minkowski's theorem to prove the existence of a polynomial with a somewhat not too high and not too low degree, for example of degree $n-\Omega(\log n)$ for $m=n-1$.

Mathematics Subject Classification (2000): .... Fill in, please

* This research was partially supported by the Israel Science Foundation (grant number 339/10).

1. Introduction

In this paper we study the following problem that was raised by von zur Gathen and Roche [6].

What is the minimal degree of a nonconstant polynomial

$$f\colon \{0,\ldots,n\} \to \{0,\ldots,m\}?$$

As f is defined over $n+1$ points, its degree is at most n, so the question basically asks whether the degree can be much smaller than n. The answer must of course depend on the choice of m. For example, when $m=n$ we have the polynomial $f(x)=x$ whereas when $m=1$ the degree of f is at least $n-n^{0.525}$ [6]. Von zur Gathen and Roche observed an obvious lower bound on the degree of nonconstant polynomials $f\colon \{0,\ldots,n\} \to \{0,\ldots,m\}$, that follows from the pigeonhole principle, namely, $\deg(f) \geq (n+1)/(m+1)$. They also noted that their techniques for the case $m=1$ cannot yield bounds better than $n-\Omega(n)$ for larger values of m. Thus, prior to this work no lower bounds of the form $n-o(n)$ were known on the degree of polynomials $f\colon \{0,\ldots,n\} \to \{0,\ldots,m\}$, when $m>1$. We note that von zur Gathen and Roche were mainly interested in the case that m is independent of n, but the problem is also relevant when $m=n-1$ and in fact even for $m \geq n$. In such cases, one should omit other 'trivial' examples besides the constant functions. The reason that a meaningful answer can be obtained is that the requirement that f takes values in the domain $\{0,\ldots,m\}$ restricts the freedom that the coefficients of f a priori had and puts a severe limitation on their structure. In this paper we focus on the case of large m, although our results clearly hold for small values of m as well.

The goal to better understand the degree of polynomials is well motivated by the important role that polynomials (both multivariate and univariate) play in theoretical computer science. For example, polynomials are prominent in areas such as circuit complexity [16,19,2], learning theory [12,15], decision tree complexity and quantum query complexity [3], Fourier analysis of Boolean functions [11,18], explicit constructions (see e.g., [8]) and more. Understanding the complexity of univariate polynomials is one of the most important problems in algebraic complexity as it is closely related to the question of hardness of integer factorization (see e.g., Section B.3 in [7]).

The degree of polynomials is probably the most simple and natural complexity measure that is associated with them. Indeed, a basic question in the study of polynomials that attracted a lot of interest concerns the minimal degree that a polynomial, belonging to some predetermined family of polynomials, can have. This fundamental question was studied before in the context of multivariate real polynomial approximation of Boolean functions (see the survey [3]), in the study of representations of symmetric Boolean functions as univariate polynomials [6] (where the problem that we study here was raised) and in relation to learning symmetric juntas [15,11,18]. In [18] it was showed that in order to better understand the Fourier spectrum of symmetric functions one needs to study polynomials $f\colon \{0,\dots,n\}\to\{0,1,2\}$ and prove lower bounds on their degree, which is exactly the question that we study here for the case $m=2$.

Besides its connection to complexity theory, the question of understanding univariate polynomials is important from an approximation theory point of view. A different angle to look at our problem is asking, for a given degree d how small can the range of a degree d polynomial mapping $\{0,\dots,n\}$ to $\mathbb{N}$ be. This question is a discrete version of a fundamental question in approximation theory concerning the minimal L^∞ norm of monic polynomials[1] over the real interval $[-1,1]$. That is, the question is what is $\min_f \max_{x\in[-1,1]} |f(x)|$, where f ranges over all monic polynomials of degree d. It is well known that Chebyshev polynomials are the only extremal example. The problem that we study in this paper basically asks for the minimum L^∞ norm that a monic polynomial of degree d attains at the points $I_n=\{-1,-1+\frac{2}{n},\dots,1\}$, namely, $\min_f \max_{x\in I_n} |f(x)|$, where f ranges over all monic polynomials of degree d. There is a significant difference from the original question as we allow the polynomial to take arbitrarily high values on other points in the interval. While for $d<\sqrt{n}$ one can get a good estimate using the classical theory of Chebyshev polynomials, this is not the case for larger values of d. We discuss this connection in more detail in Section 5.1.

1.1. Our results

We prove two main results concerning the degree of polynomials mapping integers to integers. Both results present a dichotomy behavior. That is, given a function $f\colon \{0,\dots,n\}\to\{0,\dots,m\}$, either $\deg(f)$ is very small (we consider those cases as ‘trivial’) or $\deg(f)$ is very high. The first result gives a strong lower bound when m is not too large (but still larger than n).

Theorem 1.1. *For every $\epsilon>0$ there exists n_ϵ such that for every $n>n_\epsilon$ and $f\colon \{0,1,\dots,n\}\to\{0,1,\dots,n^{1.475-\epsilon}\}$, either $\deg(f)\le 1$ or $\deg(f)\ge n-4n/\log\log n$.*

As an immediate corollary we get that if a polynomial tries to “compress” the domain even by one value, then it must have a nearly full degree.

[1] A polynomial is monic if its leading coefficient is 1.

Corollary 1.2. *Let $S \subsetneq \{0,\dots,n\}$ and $f\colon \{0,\dots,n\} \to S$ be a nonconstant polynomial. Then,* $\deg(f) \geq n - 4n/\log\log n$.

Note that such a strong result cannot hold for $m \geq n$ as, for example, the function $f(x) = x$ maps $\{0,\dots,n\}$ to itself. Our second main result concerns larger values of m at the price of a slightly weaker dichotomy.

Theorem 1.3. *There exists a constant n_0 such that if d,n are integers satisfying $d \leq \frac{2}{15}n$ and $n > n_0$, then the following holds. If $f\colon \{0,\dots,n\} \to \left\{0,\dots,\left\lfloor \frac{1}{\sqrt{7d}} \cdot \left(\frac{n-d}{2d}\right)^d \right\rfloor\right\}$ is a polynomial, then $\deg(f) \leq d-1$ or $\deg(f) \geq \frac{1}{3}n - 1.2555 \cdot \left(d \ln(\frac{n-d}{2d}) - \frac{1}{2}\ln(\frac{n}{d})\right)$.*

In other words, besides the ("trivial") case where $\deg(f) \leq d-1$, the only other option is that f has a relatively high degree.

The proof of Theorem 1.3 relies on the following theorem that gives a lower bound on the maximum value that *any* monic polynomial must obtain on the points $\{0,\dots,n\}$.

Theorem 1.4. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a degree d* monic *polynomial. Then, $\max_{i=0,1,\dots,n} |f(i)| > \left(\frac{n-d}{2e}\right)^d$. In particular, if $f\colon \mathbb{Z} \to \mathbb{Z}$ is a degree d polynomial (not necessarily monic), then*

$$\max_{i=0,1,\dots,n} |f(i)| > \frac{1}{d!} \cdot \left(\frac{n-d}{2e}\right)^d \geq \frac{1}{\sqrt{7d}} \cdot \left(\frac{n-d}{2d}\right)^d.$$

As mentioned before, this question is a discrete analog of a question from approximation theory asking for the minimal L^∞ norm of a monic polynomial of degree d over the real interval $[-1,1]$.

Our next result gives an *upper bound* on the degree when the range is of size at most $\exp(O(\sqrt{n}))$.

Theorem 1.5. *For every large enough integer $n > 0$ and an integer $k = O(\sqrt{n})$ there exists $f\colon \{0,\dots,n\} \to \{0,\dots,O(2^k)\}$ of degree $2k < \deg(f) \leq n-k$.*

In particular, by Theorem 1.3, it holds that $n/3 - k \leq \deg(f) \leq n-k$. We note that in [6] von zur Gathen and Roche conjectured that any such nonconstant polynomial to $\{0,1\}$ must be of degree $n - O(1)$. While this conjecture is still open, Theorem 1.5 shows that one can get polynomials of lower degree when the range is larger, even after excluding the obvious examples.

Finally, we consider polynomials $f\colon \{0,\dots,n\} \to \{0,1\}$, where $n = p^2 - 1$ and p is a prime number. We are able to show that in this case $\deg(f) \geq p^2 - p > n - \sqrt{n}$. This improves the result of [6] for this special case.

Lower Bounds on Degree			
Ref.	Range of f	"Trivial" case	Excluding "Trivial" case
[6]	$\{0,1\}$	f is constant	$\deg(f)=n$ when $n=p-1$, p is prime
[6]	$\{0,1\}$	f is constant	$\deg(f)\geq n-n^{0.525}$
Thm. 1.6	$\{0,1\}$	f is constant	$\deg(f)\geq n-\sqrt{n}$ when $n=p^2-1$, p is prime
Cor. 1.2	$S\subsetneq\{0,\dots,n\}$	f is constant	$\deg(f)\geq n-4n/\log\log n$
Thm. 1.1	$\left\{0,1,\dots,n^{1.475-\epsilon}\right\}$	$\deg(f)\leq 1$	$\deg(f)\geq n-4n/\log\log n$
Cor. 5.7	$\left\{0,\dots,\left\lfloor\frac{n^2-4\Gamma(n)^2}{8}\right\rfloor\right\}$	$\deg(f)\leq 1$	$\deg(f)\geq n/2-2n/\log\log n$
Thm. 5.6	$\left\{0,1,\dots,n^{2.475-\epsilon}\right\}$	$\deg(f)\leq 2$	$\deg(f)\geq n/2-2n/\log\log n$
Thm. 1.3	$\left\{0,\dots,\left\lfloor\frac{1}{\sqrt{7d}}\cdot\left(\frac{n-d}{2d}\right)^d\right\rfloor\right\}$ $d\leq\frac{2}{15}n$	$\deg(f)\leq d-1$	$\deg(f)\geq\frac{1}{3}n-1.2555\cdot\left[d\ln\left(\frac{n-d}{2d}\right)-\frac{1}{2}\ln\left(\frac{n}{d}\right)\right]$
Upper Bounds on Degree			
Ex. 5.2	$\left\{0,\dots,\binom{\frac{n+d-1}{2}}{d}\approx\left(\frac{e(n+d)}{2d}\right)^d\right\}$	$f=\binom{x-\frac{n-d+1}{2}}{d}$	
Thm. 1.5	$\left\{0,\dots,O\left(2^k\right)\right\}$ $k=O(\sqrt{n})$	$\deg(f)\leq O(\frac{k}{\log n})$	$\deg(f)\leq n-k$ (and $n/3-O(k)\leq\deg(f)$)

Table 1. Summary of Results

Theorem 1.6. *Let p be a prime number, $n=p^2-1$ and $f\colon\{0,\dots,n\}\to\{0,1\}$ be nonconstant. Then* $\deg(f)\geq p^2-p>n-\sqrt{n}$.

We summarize our results in Table 1.

1.2. Related work

The most relevant result is the aforementioned work of von zur Gathen and Roche [6] that raised and studied the question of bounding (from below) the minimal degree that a real polynomial representing a nonconstant symmetric Boolean function can have. As any symmetric function $f\colon\{0,1\}^n\to\{0,1\}$ is actually a function of the number of ones in x, it can be represented by a unique polynomial $f\colon\{0,\dots,n\}\to\{0,1\}$ (we abuse notations here and think of f both as a univariate polynomial and as a symmetric function). Thus, von zur Gathen and Roche basically studied the question of giving a lower bound on the minimal degree of nonconstant polynomials $f\colon\{0,\dots,n\}\to\{0,1\}$. They showed that when $n=p-1$, p prime, it must be the case that $\deg(f)=n$ (when f is not constant). Using the density of prime numbers (see Theorem 2.6) they concluded that $\deg(f) \geq n - o(n)$ for every n (in the notations of Theorem 2.6, $\deg(f) \geq n - \Gamma(n)$). For the case of polynomials taking values in $\{0, \ldots, m\}$, von zur Gathen and Roche observed that $\deg(f) \geq (n+1)/(m+1)$ and mentioned that their techniques cannot give any result of the form $\deg(f) = n - o(n)$. However, they suggested that "...for each m there is a constant C_m such that $\deg(f) \geq n - C_m$ for all n." In particular, when $m = O(1)$, this amounts to having $\deg(f) \geq n - O(1)$. This conjecture is still open, even for the case $m = 1$.

Another line of work concerning symmetric Boolean functions

$$f \colon \{0,1\}^n \to \{0,1\},$$

has focused on bounding from *above* the minimal size of a nonempty set S such that $\hat{f}(S) \neq 0$, where $\hat{f}(S)$ is the Fourier coefficient of f at S. We do not want to delve into the definition of the Fourier transform, so we only mention that when f is *balanced*, i.e. takes the values 0 and 1 equally often, this is the same as bounding from below the degree of $f \oplus$PARITY, see [11] for details. As symmetric Boolean functions can be represented by univariate polynomials from $\{0, \ldots, n\}$ to $\{0,1\}$, this problem is closely related to the questions studied here.

A motivation for studying the case $m > 1$ was given in [18] where it was shown that bounding from below the degree of univariate polynomials to $\{0,1,2\}$, will give an upper bound on the size of such a set S (for which $\hat{f}(S) \neq 0$), even when f is *not balanced*. Thus, an advance in understanding the degree of polynomials mapping integers to integers, that obtain more than two values, may shed new light on a well studied problem concerning the Fourier spectrum of symmetric Boolean functions.

1.3. Techniques

The proofs of Theorems 1.1, 1.4 and 1.5 use a completely different set of techniques. In the proof of Theorem 1.1 we rely on solving systems of diophantine equations by working modulo a well chosen set of primes. The proof of Theorem 1.4 is more elementary and follows from some averaging argument. For the proof of Theorem 1.5 we use lattice theory and Minkowski's theorem to prove the existence of a polynomial with the required properties. We shall now extend more on each of the proofs.

We give a very rough sketch of the idea of the proof of Theorem 1.1. Our goal is to show that every nonlinear polynomial $f \colon \{0, \ldots, n\} \to \{0, \ldots, m\}$, for $m \sim n^{1.475}$, must have high degree. As the coefficients of f are determined by the set of values $\{f(0), f(1), \ldots, f(n)\}$ if $\deg(f) \leq n$, and in fact are linear combinations of them, a natural approach is to look at these dependencies and prove that one of the coefficients of high degree monomials cannot be zero. Specifically, representing f in the basis of the *Newton polynomials* (see Definition 2.2) we get an explicit and nice formula for each coefficient. If f is not of high degree, many of those coefficients vanish and this gives a set of *linear equations* that the values $\{f(0), f(1), \ldots, f(n)\}$ must satisfy. In fact, we manage to get many linear equations from *every* zero coefficient. The idea is that if the degree of f is smaller than a prime number p, then the values $f(r)$ and $f(r+p)$ must be strongly correlated for $r \in \{0, \ldots, n-p\}$. Using such correlations for many different primes, we obtain a set of special linear equations (which we call *linear recurrence relations*) on the values of f. A similar approach was taken in [11] (and arguably also in [6]) where the authors used different primes to obtain information for the case $m=1$.

It is not clear, however, how to exploit the information from the different primes. We manage to do so by considering prime numbers that form a 'nice' and 'rigid' structure that we call a *cube of primes*. An r-dimensional cube of primes is a set $P = P_{p;\delta_1,\ldots,\delta_r} \subseteq \{1, \ldots, n\}$ of the form

$$P = \left\{ p + \sum_{i=1}^{r} a_i \delta_i \;\middle|\; a_1, \ldots, a_r \in \{0,1\} \right\},$$

such that all the elements of P are prime numbers. The idea is that we can partition P, in many different ways, to pairs of primes such that the differences, between the primes in each pair, are the same. This enables us to combine the different linear recurrences obtained from each prime in a way that reveals more information on the values that f takes.

Theorem 1.3 is an immediate corollary of Theorem 1.4 whose proof goes along completely different lines than the proof of Theorem 1.1. The idea is to observe that since f has at most d roots in the interval $\{0, \ldots, n\}$, some point in that interval is relatively far from all roots of f. This immediately implies that f obtains a large value at this point.

To prove Theorem 1.5 we note that polynomials of degree at most $D = n - k$ evaluated on $0, 1, \ldots, n$ form a lattice. Since we are interested in the polynomials that have small coordinates, our problem corresponds to finding a short vector in a lattice with respect to the L^∞ norm. Using Minkowski's theorem, we can prove the existence of a non-trivial polynomial (i.e. of a not too low and not too high degree) with a small L^∞ norm.

1.4. Organization

The paper is organized as follows. In Section 2 we give the basic definitions and discuss mathematical tools that we shall later use. In Section 3 we demonstrate our general technique by considering the case of 2-dimensional cube of primes. In Section 4 we prove Theorem 1.1 and conclude Corollary 1.2. In Section 5 we prove Theorems 1.3 and 1.4 and discuss their tightness. We then present the connection to Chebyshev polynomials in Section 5.1 and conclude Theorem 5.5 that improves Theorem 1.4 for $d \le \sqrt{n}/2$. We prove Theorem 1.5 in Section 6. Finally, in section 7 we consider the case $m=1$ and $n=p^2-1$ for a prime p. We note that the results in Sections 4, 5 and 6 are independent of each other so it is not required to read the paper in a linear order.

2. Preliminaries

For two integers a,b we denote with $[a,b]$ the set of all integers between a and b. Namely, $[a,b] \triangleq \{c \in \mathbb{Z} \mid a \le c \le b\} = \{a, a+1, \ldots, b\}$. We also denote $[m] \triangleq [1,m]$. We sometimes abuse notation and speak of the *real interval* $[a,b]$ (in this case $[a,b] = \{a \le x \le b \mid x \in \mathbb{R}\}$). We will always mention the words 'real interval' whenever we speak of the real interval.

For a prime number p and integers a,b we denote $a \equiv_p b$ when a and b are equal modulo p. For a polynomial $f(x) = \sum_{i=0}^{n} a_i x^i$ we denote with $\mathrm{spar}(f)$ the number of monomials in f, i.e. the number of nonzero a_i's. We denote the family of all polynomials from $[0,n]$ to $[0,m]$ by $\mathcal{F}_m(n)$. Namely,

$$\mathcal{F}_m(n) = \{f \in \mathbb{Q}[x] \mid \deg(f) \le n, f\colon [0,n] \to [0,m]\}.$$

Throughout the paper we avoid the use of floor and ceiling in order not to make the equations even more cumbersome. This does not affect our results and only makes the reading easier.

We denote by $\log(\cdot)$ and $\ln(\cdot)$ the logarithms to the base 2 and to the base e (that is, the natural logarithm) respectively.

In the next subsections we present some well known technical tools that we require for our proofs.

2.1. Stirling's formula

We shall make use of the well known Stirling approximation for the factorial function.

Theorem 2.1 (Stirling's formula). *For every natural number $n \in \mathbb{N}$ it holds that*

$$n! = \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n \cdot e^{\lambda_n}$$

with

$$\frac{1}{12n+1} < \lambda_n < \frac{1}{12n}.$$

A proof of this theorem can be found, e.g., in [17] (see also pages 50-53 of [5]).

2.2. Newton basis

Definition 2.2. *For every $k \in \mathbb{N}$, define the polynomial $\binom{x}{k}$ as follows*

$$\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}.$$

The set of polynomials $\left\{\binom{x}{k} : k \in \mathbb{N}\right\}$ is called the Newton basis.

It is easy to see that $\left\{\binom{x}{k} : k = 0, 1, \ldots, d\right\}$ forms a basis of the vector space of polynomials of degree at most d. An interesting property of the Newton basis is given in the next theorem (see e.g., problem 36 in [10]).

Theorem 2.3. *Let $f \in \mathbb{Q}[x]$ be a polynomial of degree $\leq n$. Then f can be represented as*

$$f(x) = \sum_{d=0}^{n} \gamma_d \cdot \binom{x}{d} \quad \text{where} \quad \gamma_d = \sum_{j=0}^{d} (-1)^{d-j} \cdot \binom{d}{j} \cdot f(j).$$

As noted in [6], Theorem 2.3 implies that a polynomial f is of degree smaller than d iff for all $d \leq s \leq n$ it holds that

$$\sum_{j=0}^{s} (-1)^j \binom{s}{j} f(j) = (-1)^s \gamma_s = 0.$$

As an immediate corollary we get the following useful lemma.

Lemma 2.4. *Let $f\colon [0,n] \to \mathbb{Z}$ be such that $\deg(f) < d$. Then for all $r \in [0, n-d]$ we have that*

$$\sum_{j=0}^{d} \binom{d}{j} \cdot (-1)^j \cdot f(j+r) = 0.$$

Proof. For $r\in[0,n-d]$ set $g_r(x)=f(x+r)$. We think of g_r as a function $g_r\colon[0,n-r]\to\mathbb{Z}$. As $\deg(g_r)=\deg(f)<d$, and $d\le n-r$ Theorem 2.3 implies that

$$\sum_{j=0}^{d}(-1)^j\binom{d}{j}f(j+r)=\sum_{j=0}^{d}(-1)^j\binom{d}{j}g_r(j)=0.$$ ∎

2.3. Lucas' theorem

The following theorem of Lucas [13] allows one to compute a binomial coefficient modulo a prime number.

Theorem 2.5 (Lucas' theorem). *Let $a,b\in\mathbb{N}\setminus\{0\}$ and let p be a prime number. Denote with*

$$a=a_0+a_1p+a_2p^2+\cdots+a_kp^k,$$
$$b=b_0+b_1p+b_2p^2+\cdots+b_kp^k,$$

their base p expansion. Then

$$\binom{a}{b}\equiv_p\prod_{i=0}^{k}\binom{a_i}{b_i},$$

where $\binom{a_i}{b_i}=0$ if $a_i<b_i$.

2.4. The gap between consecutive primes

Denote with p_n the n-th prime number. Understanding the asymptotic behavior of $p_{n+1}-p_n$ is a long standing open question in number theory. Cramér conjectured that $p_{n+1}-p_n=O((\log p_n)^2)$ and, assuming the correctness of Riemann hypothesis, he proved that $p_{n+1}-p_n=O(\sqrt{p_n}\log p_n)$ [4]. The strongest unconditional result is due to Baker et al. [1].[2] Denote with $\pi(n)$ the number of primes numbers less than or equal to n.

Theorem 2.6 ([1]). *For any large enough integer n and any $y\ge n^{0.525}$ we have that*

$$\pi(n)-\pi(n-y)\ge\frac{9}{100}\cdot\frac{y}{\log n}.$$

[2] The main theorem of [1] only claims that there exists a prime number in the interval $[n-n^{0.525},n]$, however they actually prove the stronger claim that is stated here.

For convenience, we denote

$$\Gamma(n) \triangleq n^{0.525}.$$

We will usually apply the theorem above to claim, for some integer n, that there exists a prime number $p \in [n-\Gamma(n), n]$.

2.5. Linear recurrence relations

Definition 2.7. *Let $\Phi(t) = \sum_{i=0}^{s} \alpha_i t^i$ be a polynomial with rational coefficients.*[3] *For $f \in \mathbb{Q}[x]$ we define the action of Φ on f as*

$$(\Phi \circ f)(x) \triangleq \sum_{i=0}^{s} \alpha_i \cdot f(x+i).$$

When we consider Φ as an operator acting on other polynomials, we call Φ a linear recurrence *polynomial.*

From now on we will always denote linear recurrence polynomials with capital Greek letters: Φ, Ψ, Υ. Following is a list of properties of linear recurrence polynomials.

Lemma 2.8. *For polynomials f, g and linear recurrences Φ, Φ' the following claims hold.*

1. $\Phi \circ f \in \mathbb{Q}[x]$.
2. $\deg(\Phi \circ f) \leq \deg(f)$.
3. $(\Phi + \Phi') \circ f = \Phi \circ f + \Phi' \circ f$.
4. $\Phi \circ (f+g) = \Phi \circ f + \Phi \circ g$.
5. $(\Phi \cdot \Phi') \circ f = \Phi \circ (\Phi' \circ f)$.

Proof. Properties 1-4 follow trivially from the definition. Property 5 follows by a simple calculation. Denote, w.l.o.g., $\Phi(t) = \sum_{i=0}^{d} \alpha_i x^i$ and $\Phi'(t) = \sum_{j=0}^{e} \beta_j x^j$. We have that

$$\begin{aligned}\left(\Phi \cdot \Phi'\right) \circ f(x) &= \left(\sum_{i=0}^{d} \sum_{j=0}^{e} \alpha_i \beta_j x^{i+j}\right) \circ f(x) \\ &= \sum_{i=0}^{d} \sum_{j=0}^{e} \alpha_i \beta_j f(x+i+j)\end{aligned}$$

[3] There is nothing special about $\mathbb{Q}$ and the only reason that we use it is that in our proofs we encounter rational coefficients.

$$= \sum_{i=0}^{d} \alpha_i \underbrace{\left(\sum_{j=0}^{e} \beta_j f\left(x+i+j\right) \right)}_{(\Phi' \circ f)(x+i)}$$
$$= \big(\Phi \circ (\Phi' \circ f)\big)\,(x). \qquad \blacksquare$$

While property 2 of Lemma 2.8 states the obvious fact that applying a linear recurrence cannot increase the degree, the following lemma assures that the degree can decrease by (roughly) at most the number of monomials in the linear recurrence polynomial.

Lemma 2.9. *Let $f \in \mathbb{Q}[x]$ be a nonconstant polynomial and let $\Phi(t) = \sum_{i=1}^{s} \alpha_i \cdot t^{d_i}$ be some linear recurrence, $\Phi \neq 0$. Then, for $g = \Phi \circ f$ we have that*

$$\deg(f) \leq \begin{cases} s-2 & g \equiv 0 \\ s + \deg(g) - 1 & \textit{otherwise.} \end{cases}$$

Proof. As $\Phi \neq 0$ we can assume w.l.o.g. that the exponents $d_1, \ldots, d_s$ are distinct (indeed if they are not distinct then we can rewrite Φ as a polynomial with $s' < s$ monomials and obtain stronger results). Similarly, if $\deg(f) \leq s-2$ then we are done. So, we may assume w.l.o.g. that $\deg(f) \geq s-1$. Let $f(x) = \sum_{\ell=0}^{D} b_i x^i$, where $b_D \neq 0$. Let L be a $(D+1)\times(D+1)$ lower triangular matrix whose (i,j) entry (for $i,j = 0, \ldots, D$) is $L_{i,j} \triangleq b_{D+j-i} \cdot \binom{D+j-i}{j}$ (where $b_{D+j-i} = 0$ if $j > i$). This is clearly a lower triangular matrix with a nonzero diagonal. Let V be a $(D+1)\times s$ Vandermonde matrix defined by $V_{i,j} \triangleq (d_j)^i$ for $i = 0, \ldots, D$ and $j = 1, \ldots, s$. It is now easy to verify that the coefficients of the polynomial $g = \Phi \circ f$ are the result of the matrix-vector multiplication $L \cdot V \cdot \vec{\alpha}$ where $\vec{\alpha} = (\alpha_1, \ldots, \alpha_s)$. Namely, if $g(x) = \sum_{i=0}^{D} c_i x^i$, then $(c_D, \ldots, c_0) = L \cdot V \cdot \vec{\alpha}$. Thus $c_{D-r} = (L \cdot V \cdot \vec{\alpha})_r$. Indeed,

$$(\Phi \circ f)(x) = \sum_{i=1}^{s} \alpha_i f(x+d_i) = \sum_{i=1}^{s} \alpha_i \sum_{j=0}^{D} b_j (x+d_i)^j$$
$$= \sum_{i=1}^{s} \alpha_i \sum_{j=0}^{D} b_j \sum_{k=0}^{j} \binom{j}{k} d_i^{j-k} x^k$$
$$= \sum_{k=0}^{D} x^k \sum_{j=k}^{D} b_j \binom{j}{k} \sum_{i=1}^{s} \alpha_i d_i^{j-k}$$
$$= \sum_{k=0}^{D} x^k \sum_{\ell=0}^{D-k} b_{\ell+k} \binom{\ell+k}{k} \sum_{i=1}^{s} \alpha_i d_i^{\ell}.$$

Hence, the coefficient of x^{D-r} is

$$\sum_{\ell=0}^{r} b_{\ell+D-r}\binom{\ell+D-r}{D-r}\sum_{i=1}^{s}\alpha_i d_i^{\ell}=\sum_{\ell=0}^{r} L_{r,\ell}(V\cdot\vec{\alpha})_{\ell}=$$
$$\sum_{\ell=0}^{D} L_{r,\ell}(V\cdot\vec{\alpha})_{\ell}=(L\cdot V\cdot\vec{\alpha})_r.$$

As the first s rows (recall that $D+1=\deg(f)+1\geq s$) of $L{\cdot}V$ form an invertible matrix (as a product of a Vandermonde matrix with a lower triangular matrix that has a nonzero diagonal), we see that the top s coefficients of g are zero iff $\vec{\alpha}=0$ (which is a contradiction to the assumption that $\Phi\neq 0$). Hence, the degree of g is at least $D-s+1=\deg(f)-s+1$. ∎

3. Warm up

In this section we prove some preliminary results that give good intuition to the proofs of Theorem 1.1 (and also to the proof of Theorem 5.6). Similarly to other works that studied the degree of polynomials mapping integers to integers [6,11], we shall consider properties of the polynomial modulo different prime numbers.

As a first step we show that if $f\in\mathcal{F}_{n-1}(n)$ is of low degree then it is actually a constant function. The proof of the lemma already contains some of the ingredients that we will later use in a more sophisticated manner.

Lemma 3.1. *Let $f\in\mathcal{F}_{n-1}(n)$ be such that $\deg(f)<n/6-\Gamma(n)$, then f is a constant.*

Proof. Let $p\in[n/2,n/2+\Gamma(n)]$ be a prime number, guaranteed to exist by Theorem 2.6. Since $\deg(f)<p$, Lemma 2.4 implies that for all $r\in[0,n/2-\Gamma(n)]\subseteq[0,n-p]$ we have that

$$0=\sum_{k=0}^{p}(-1)^k\binom{p}{k}f(k+r)\equiv_p f(r)-f(p+r).$$

In particular, if we define g by $g(r)=\frac{f(r)-f(p+r)}{p}$, then we have that $g\colon[0,n/2-\Gamma(n)]\to[-1,1]$ (indeed, $f(r)-f(p+r)\in[-n+1,n-1]$). Clearly, $g+1\in\mathcal{F}_2(n)$. Note that if g is not constant then its degree must be at least $(n/2-\Gamma(n))/3$ as one of the values in its range is obtained at least that many times. Since in this case $n/6-\Gamma(n)<\deg(g)\leq\deg(f)$ we get a contradiction. Therefore, g must be constant. However, in this case we get by Lemma 2.9 that $\deg(f)\leq\deg(g)+2-1=1$. Indeed, for $\Phi(t)=\frac{1}{p}-\frac{1}{p}t^p$, it holds that $g = \Phi \circ f$. Hence, $\deg(f) \leq 1$. Since the range of f is smaller than its domain (and f takes integer values), f must be constant. ∎

Clearly, for $m \geq n$, we cannot expect such a strong behavior (that is, degree 0 as opposed to degree $\Omega(n)$). However, the following lemma, which relies on Lemma 3.1, shows that a slightly weaker dichotomy behavior exists for m which is roughly quadratic in n. We later strengthen this result (Corollary 5.7).

Lemma 3.2. *Let $m < \frac{n^2 - 4\Gamma(n)^2}{8}$ be an integer and $f \in \mathcal{F}_m(n)$ be such that $\deg(f) < n/12 - \Gamma(n)$, then $\deg(f) \leq 1$.*

Proof. Let $p \in [\frac{n}{2} - \Gamma(n), \frac{n}{2}]$ be a prime number, guaranteed to exist by Theorem 2.6. As before, Lemma 2.4 implies that for all $r \in [0, n-p]$ we have that

$$0 = \sum_{k=0}^{p} (-1)^k \binom{p}{k} f(k+r) \equiv_p f(r) - f(p+r).$$

In particular, if we define g by $g(r) = \frac{f(r) - f(p+r)}{p}$, then we have that $g\colon [0, n-p] \to [-m/p, m/p]$. Clearly, $g + \frac{m}{p} \in \mathcal{F}_{\frac{2m}{p}}(n-p)$, and

$$2\frac{m}{p} < \frac{(\frac{n}{2} - \Gamma(n))(\frac{n}{2} + \Gamma(n))}{p} \leq n - p.$$

Hence, $g + \frac{m}{p}$ is actually in $\mathcal{F}_{n-p-1}(n-p)$, and

$$\deg(g + \frac{m}{p}) \leq \deg(f) \leq \frac{n}{12} - \Gamma(n) \leq \frac{n-p}{6} - \Gamma(n-p).$$

Now we can apply Lemma 3.1 to conclude that $g + \frac{m}{p}$ is constant. From Lemma 2.9 it follows that $\deg(f) \leq 1$ which completes the proof. ∎

We note that the choice $m < \frac{n^2 - 4\Gamma(n)^2}{8}$ is very close to being tight. Indeed, assume that n is odd and consider the function $f\colon [0, n] \to [0, \frac{n^2-1}{8}]$ defined as $f(x) = \binom{x - \frac{n-1}{2}}{2}$.

An important ingredient in the proof of Theorem 1.1 is the use of prime numbers that form a structure analogous to a cube. To illustrate our approach, consider four prime numbers of the form $p < p + \delta_1 < p + \delta_2 < p + \delta_1 + \delta_2$. Using Theorem 2.6 one can show that such primes exist and that we can even choose them so that they all lie in an interval of the form $[n/3 - o(n), n/3]$.

Lemma 3.3. *Let n be a large enough integer. Then, there exist four prime numbers*

$$\frac{n}{3} - \Gamma(n) \leq p < p + \delta_1 < p + \delta_2 < p + \delta_1 + \delta_2 \leq \frac{n}{3}.$$

Proof. The lemma follows from the more general Lemma 4.1 that is proved in Section 4.1, however, for clarity we prove this special case here.

Theorem 2.6 guarantees that for a large enough n there are at least[4] $\Gamma(n)/12\log(n)$ prime numbers in the interval $[n/3-\Gamma(n),n/3]$. Consider all possible differences between two primes in this set. There are at least, say, $\frac{1}{3}(\Gamma(n)/12\log(n))^2$ such differences. As all the differences are smaller than $\Gamma(n)$ it follows that one of the differences is obtained for at least $\frac{\frac{1}{3}(\Gamma(n)/12\log(n))^2}{\Gamma(n)} \geq \frac{\Gamma(n)}{500\log^2(n)}$ many pairs of primes. Denote the i-th pair with $(p_{i,1},p_{i,2})$ where $p_{i,1} < p_{i,2}$. Consider any two distinct pairs in the set, $(p_{1,1},p_{1,2})$ and $(p_{2,1},p_{2,2})$. Denote $\delta_1 = p_{1,2}-p_{1,1} = p_{2,2}-p_{2,1}$ and $\delta_2 = |p_{1,1}-p_{2,1}| > 0$. We have that $0 < \delta_1+\delta_2 < \Gamma(n)$. In particular, $\{p_{1,1},\dots,p_{2,2}\}$ is the required cube.[5] ∎

As a warmup for our main result and to demonstrate our proof technique we shall prove here the following easier theorem.

Theorem 3.4. *If $f\in\mathcal{F}_m(n)$, where $m<n/7$, is nonconstant then* $\deg(f)\geq 2n/3-2\Gamma(n)$.

Although the theorem is much weaker than Theorem 1.1, its proof demonstrates our general technique and, hopefully, will make the proof of Theorem 1.1 easier to follow.

Proof. Let p,δ_1,δ_2 be as guaranteed in Lemma 3.3. Assume for a contradiction that $f\in\mathcal{F}_m(n)$ is such that $\deg(f)<2n/3-2\Gamma(n)\leq 2p$. Consider the identity guaranteed by Lemma 2.4 modulo each of the four primes. For example, taking $d=2p$ (in the notations of Lemma 2.4), we get that for all $r=0,\dots,n-2p$

$$0=\sum_{k=0}^{2p}(-1)^k\binom{2p}{k}f(k+r)\equiv_p f(r)-2f(p+r)+f(2p+r). \tag{1}$$

Since $|f(r)-2f(p+r)+f(2p+r)|<2n/7<p$, Equation (1) is actually satisfied over the integers. Namely, $f(r)-2f(p+r)+f(2p+r)=0$. In the same manner we get, for all $r\in[0,n-2(p+\delta_1+\delta_2)]$

$$\begin{aligned} f_{0,0}(r) &\triangleq f(r)-2f(p+r)+f(2p+r)=0, \qquad (2)\\ f_{1,0}(r) &\triangleq f(r)-2f(p+\delta_1+r)+f(2p+2\delta_1+r)=0,\\ f_{0,1}(r) &\triangleq f(r)-2f(p+\delta_2+r)+f(2p+2\delta_2+r)=0, \end{aligned}$$

[4] There is nothing special about 12, it is just a large enough constant.

[5] We can of course make sure that $p_{2,1}\neq p_{1,2}$, and hence $\delta_1\neq\delta_2$, by 'throwing' away one pair.

$$f_{1,1}(r) \triangleq f(r) - 2f(p + \delta_1 + \delta_2 + r) + f(2p + 2\delta_1 + 2\delta_2 + r) = 0.$$

We now show how to combine these equations in a way that will give information not only for small values of r (i.e. $r \le n - 2(p + \delta_1 + \delta_2)$) but also for larger values of r. By considering the following linear combinations of the equalities $f_{0,0}, \dots, f_{1,1}$ we get that for $r \in [0, n - 2(p + \delta_2 + 2\delta_1)]$ it holds that

$$\begin{aligned} 0 = f_{0,0}(r + 2\delta_1) - f_{1,0}(r) &= f(r + 2\delta_1) - f(r) - 2f(p + r + 2\delta_1) \\ &\quad + 2f(p + r + \delta_1), \\ 0 = f_{0,1}(r + 2\delta_1) - f_{1,1}(r) &= f(r + 2\delta_1) - f(r) - 2f(p + r + 2\delta_1 + \delta_2) \\ &\quad + 2f(p + r + \delta_1 + \delta_2). \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &= (f_{0,0}(r + 2\delta_1 + \delta_2) - f_{1,0}(r + \delta_2)) - (f_{0,1}(r + 2\delta_1) - f_{1,1}(r)) \\ &= f(r + 2\delta_1 + \delta_2) - f(r + \delta_2) - f(r + 2\delta_1) + f(r). \end{aligned}$$

Similarly,

$$\begin{aligned} 0 &= -\frac{1}{2} \cdot ((f_{0,0}(r + 2\delta_1) - f_{1,0}(r)) - (f_{0,1}(r + 2\delta_1) - f_{1,1}(r))) \\ &= f(p + r + 2\delta_1) - f(p + r + \delta_1) - f(p + r + 2\delta_1 + \delta_2) \\ &\quad + f(p + r + \delta_1 + \delta_2) \end{aligned}$$

and

$$\begin{aligned} 0 &= f_{0,0}(r + \delta_1) - f_{1,0}(r) - f_{0,1}(r + \delta_1) + f_{1,1}(r) \\ &= f(2p + r + \delta_1) - f(2p + r + 2\delta_1) - f(2p + r + \delta_1 + 2\delta_2) \\ &\quad + f(2p + r + 2\delta_1 + 2\delta_2). \end{aligned}$$

We thus get the following equations for every $0 \le r \le n - 2(p + \delta_1 + \delta_2)$:

$$0 = f(r + 2\delta_1 + \delta_2) - f(r + \delta_2) - f(r + 2\delta_1) + f(r) \tag{3}$$

$$\begin{aligned} 0 &= f(p + r + 2\delta_1) - f(p + r + \delta_1) - f(p + r + 2\delta_1 + \delta_2) \\ &\quad + f(p + r + \delta_1 + \delta_2) \end{aligned} \tag{4}$$

$$\begin{aligned} 0 &= f(2p + r + \delta_1) - f(2p + r + 2\delta_1) - f(2p + r + \delta_1 + 2\delta_2) \\ &\quad + f(2p + r + 2\delta_1 + 2\delta_2). \end{aligned} \tag{5}$$

These equations give linear recurrence relations on the values of f on the intervals $[0, n - 2p]$, $[p, n - p]$ and $[2p, n]$. Indeed, Equations (4) and (5) are equivalent to

$$0 = f(r + 2\delta_1) - f(r + \delta_1) - f(r + 2\delta_1 + \delta_2) + f(r + \delta_1 + \delta_2) \tag{6}$$

$$0 = f(r+\delta_1) - f(r+2\delta_1) - f(r+\delta_1+2\delta_2) + f(r+2\delta_1+2\delta_2) \quad (7)$$

for $r \in [p, n-p-2(\delta_1+\delta_2)]$ and $r \in [2p, n-2(\delta_1+\delta_2)]$, respectively. Let

$$\begin{aligned}\Phi(t) = {} & (t^{2\delta_1+\delta_2} - t^{\delta_2} - t^{2\delta_1} + 1)\cdot \\ & (t^{2\delta_1} - t^{\delta_1} - t^{2\delta_1+\delta_2} + t^{\delta_1+\delta_2})\cdot \\ & (t^{\delta_1} - t^{2\delta_1} - t^{\delta_1+2\delta_2} + t^{2\delta_1+2\delta_2}). \end{aligned} \quad (8)$$

It follows that $(\Phi \circ f)(r) = 0$ for all

$$r \in [0, n-2p-6(\delta_1+\delta_2)] \cup [p, n-p-6(\delta_1+\delta_2)] \cup [2p, n-6(\delta_1+\delta_2)]$$

(see Property 5 in Lemma 2.8).[6] We have two cases:

- The three ranges are distinct. In this case, $\Phi \circ f$ has at least $3\cdot(n-2p-6(\delta_1+\delta_2)) \geq n-18(\delta_1+\delta_2)$ many roots.
- The three ranges overlap. In this case, $\Phi \circ f$ has at least $n-6(\delta_1+\delta_2)$ many roots.

Either way, $\Phi\circ f$ has at least $n-18(\delta_1+\delta_2)$ many roots. We conclude that either $\Phi\circ f \equiv 0$ or $\deg(\Phi\circ f) \geq n-18(\delta_1+\delta_2)$. As $\deg(\Phi\circ f) \leq \deg(f) < \frac{2}{3}n < n-18(\delta_1+\delta_2)$ it must be the case that $\Phi \circ f \equiv 0$. Hence, by Lemma 2.9 it follows that $\deg(f) = O(1)$. However, at this point we can apply Lemma 3.1 and conclude that f is constant. ∎

In the general case, we will not be able to deduce that in (the analogous equation to) Equation (2) the sum is equal to 0, but rather we will only bound it from above. Furthermore, we will work with $2^{\Omega(\log\log n)}$ many prime numbers that form a structure of an $\Omega(\log\log n)$-dimensional cube (in the sense that $\{p, p+\delta_1, p+\delta_2, p+\delta_1+\delta_2\}$ is a 2-dimensional cube). This will make the construction of the relevant Φ more complicated, but the high level ideas will be similar.

4. Proof of Theorem 1

In this section we prove Theorem 1.1. We begin by giving a proof overview.

Let $f\colon [n] \to [m]$, where $m = n^{1.475-\epsilon}$, such that $\deg(f) \leq n - \frac{4n}{\log\log n}$. We shall find a linear recurrence Υ with the following two properties:[7]

[6] The change in the range of r occurs since we want all the evaluations points of $\Phi \circ f$ to be inside the interval $[0, n]$.

[7] Previous techniques take $\Upsilon(t) := \frac{t^p-1}{p}$ for $p \in [\deg(f), n]$ as the recurrence, which is range reducing, but not of low-degree. We shall combine information from several primes to establish this goal.

1. **Low Degree.** Υ is of degree $\leq n^{\epsilon+o(1)}$ and of sparsity $n^{o(1)}$.
2. **Range Reducing.** The polynomial $g = \Upsilon \circ f$ maps $[n'] \to [-m', m']$ where $n' = n - O(n^{\epsilon+o(1)})$ and $m' \leq \frac{m}{n^{1-\epsilon-o(1)}} \leq \sqrt{n}$.

By applying the linear recurrence on again, this time on g, we get a polynomial $h = \Upsilon \circ g$ that maps $[n''] \to [-m'', m'']$, where $n'' = n - O(n^{\epsilon+o(1)})$ and $m'' = \frac{m'}{n^{1-\epsilon-o(1)}} < 1$, i.e. h has as least n'' roots. By Lemma 2.8, $\deg(h) \leq \deg(g) \leq \deg(f) < n''$, and we get that $h \equiv 0$. Using Lemma 2.9, we get that $\deg(g) \leq \mathrm{spar}(\Upsilon) - 2$ and by applying the lemma again we get that $\deg(f) \leq \mathrm{spar}(\Upsilon) + \deg(g) - 1 \leq 2 \cdot \mathrm{spar}(\Upsilon) - 3 < 2 \cdot \deg(\Upsilon)$ which means that f is of much lower degree than we were promised initially. This allows us to apply Lemma 3.2 and conclude that $\deg(f) \leq 1$.

Proof of Theorem 1.1. For convenience, set $\mu = \log\log(n)/2$ and $m = n^{1.475-\epsilon}$. Let $f \in \mathcal{F}_m(n)$ be a function such that

$$\deg(f) < n \cdot \left(1 - \frac{2}{\mu}\right) = n - \frac{4n}{\log\log n}.$$

As was demonstrated in Section 3, we will consider the behavior of f modulo various prime numbers that form a high dimensional cube of primes. The existence (and properties) of this structure is guaranteed by the next lemma.

Lemma 4.1. *Let $0 < \epsilon < 1/2$, there exists $n_0(\epsilon)$ such that for any $n > n_0(\epsilon)$ and $\mu = \log\log(n)/2$, there exists a set*

$$\begin{aligned} P_{p;\delta_0,\delta_1,\delta_2,\ldots,\delta_\mu} &= \left\{ p + \sum_{i=0}^{\mu} a_i \cdot \delta_i \mid \forall i \ a_i \in \{0,1\} \right\} \\ &\subseteq \left[\frac{n}{\mu+1} - 4\Gamma(n), \frac{n}{\mu+1} - \Gamma(n) \right] \end{aligned}$$

with the following properties:

1. *Every $q \in P_{p;\delta_0,\delta_1,\delta_2,\ldots,\delta_\mu}$ is a prime number.*
2. *$\delta_i > 0$ for all $i = 1, \ldots, \mu$.*
3. *$\Delta \triangleq \sum_{i=1}^{\mu} \delta_i \leq n^{\epsilon}$.*
4. *$\delta_0 \in [\Gamma(n), 3\Gamma(n)]$.*

We defer the proof of the lemma to Section 4.1 and continue with the proof of Theorem 1.1. We shall consider two subcubes of $P_{p;\delta_0,\delta_1,\delta_2,\ldots,\delta_\mu}$. Denote $\mathcal{B} \triangleq P_{p;\delta_1,\delta_2,\ldots,\delta_\mu}$ and $\mathcal{B}_0 \triangleq P_{p+\delta_0;\delta_1,\delta_2,\ldots,\delta_\mu}$. Note that in both $\mathcal{B}, \mathcal{B}_0$ we do not consider shifts by δ_0. Let $q \in P_{p;\delta_0,\delta_1,\delta_2,\ldots,\delta_\mu} = \mathcal{B} \cup \mathcal{B}_0$ be a prime number. From the construction of $P_{p;\delta_0,\delta_1,\delta_2,\ldots,\delta_\mu}$ it follows that (for a large enough n)

$$\deg(f) < n \cdot \left(1 - \frac{2}{\mu}\right) < \frac{n}{\mu+2} \cdot \mu < q\mu. \tag{9}$$

Combining Lemma 2.4 and Lucas' theorem (Theorem 2.5) we get that for every $r \in [0, n - q\mu]$ it holds that

$$0 = \sum_{j=0}^{q\mu} \binom{q\mu}{j} \cdot (-1)^j \cdot f(j+r) \equiv_q \sum_{j=0}^{\mu} \binom{\mu}{j} \cdot (-1)^j \cdot f(qj+r). \tag{10}$$

Notice that this equality is analogous to Equation (1) from the proof of Theorem 3.4. Since $f \in \mathcal{F}_m(n)$ we can rewrite Equation (10) as

$$\sum_{j=0}^{\mu} \binom{\mu}{j} \cdot (-1)^j \cdot f(qj+r) = K_{q,r}(f) \cdot q, \tag{11}$$

where $K_{q,r}(f)$ is an integer satisfying:

$$\begin{aligned} |K_{q,r}(f)| &< \frac{2^\mu \cdot m}{q} < \frac{2^\mu \cdot m}{n/(\mu+2)} = \frac{m}{n} \cdot 2^\mu \cdot (\mu+2) \\ &< \frac{m}{n} \cdot 2^{2\mu} = n^{0.475-\epsilon} \cdot 2^{2\mu}. \end{aligned} \tag{12}$$

Thus, instead of summing to 0 as was the case in Equation (2), we get that the sum equals a relatively small (i.e., at most $\log(n) \cdot n^{0.475-\epsilon}$) multiple of q. In the language of linear recurrence, when applying the linear recurrence

$$\Psi_q(t) = \sum_{j=0}^{\mu} \binom{\mu}{j} \cdot (-1)^j \cdot t^{qj} \tag{13}$$

to f we get

$$(\Psi_q \circ f)(r) = K_{q,r}(f) \cdot q \tag{14}$$

for every $r \in [0, n - q\mu]$. We now combine all the different Ψ_q's to obtain a linear recurrence in an analogous way to the way that we combined the different equalities in (2) to create the linear recurrences given by (3),(4) and (5). Let $\tilde{p}$ be either p or $p + \delta_0$. We will cancel out all the monomials of the linear recurrence except those whose exponents lie in a small range: $[\tilde{p}k, \tilde{p}k + \mu\Delta]$ (recall that $\Delta = \sum_{i=1}^{\mu} \delta_i \le n^\epsilon$). Consider the following linear recurrence for $k \in [0, \mu]$

$$\begin{aligned} \Phi'_{\tilde{p},k}(t) = &\sum_{\vec{a} \in \{0,1\}^\mu} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot \Psi_{(\tilde{p} + \sum_{i=1}^{\mu} a_i \cdot \delta_i)}(t) \\ &\cdot t^{\sum_{i=1}^{k} (1-a_i)\cdot(i-1)\cdot\delta_i + \sum_{i=k+1}^{\mu} (1-a_i)\cdot i \cdot \delta_i}. \end{aligned} \tag{15}$$

The reason for this complicated looking expression will become clear soon when we show that this linear recurrence give information about $f(r)$ for $r \in [\tilde{p}k, \tilde{p}k + n - \mu(\tilde{p} + \Delta)]$. The following claim shows that indeed $\Phi'_{\tilde{p},k}$ has the required property. To simplify the statement of the claim let[8]

$$c_{\vec{a},k,k}(i) \triangleq \begin{cases} k & \text{if } a_i = 1 \\ i-1 & \text{if } a_i = 0 \text{ and } i \le k \\ i & \text{if } a_i = 0 \text{ and } i \ge k+1. \end{cases} \tag{16}$$

Claim 4.2.

$$\Phi'_{\tilde{p},k}(t) = t^{k\tilde{p}} \cdot (-1)^k \cdot \binom{\mu}{k} \cdot \sum_{\vec{a} \in \{0,1\}^\mu} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot t^{\sum_{i=1}^{\mu} c_{\vec{a},k,k}(i) \cdot \delta_i}.$$

To ease the reading we postpone the proof of the claim to Section 4.2 and proceed with the proof of Theorem 1.1. Claim 4.2 has two interesting consequences. The first is that $\tilde{p}$ only appears in the term $t^{k\tilde{p}}$. The second is that $\Phi'_{\tilde{p},k}$ is actually divisible by $t^{k\tilde{p}}$. In particular if we set

$$\Phi_{\tilde{p},k}(t) \triangleq \Phi'_{\tilde{p},k}(t)/t^{k\tilde{p}} \tag{17}$$

then we get that $\Phi_{\tilde{p},k}$ gives a recurrence relation for every $r \in \tilde{p}k + [0, n - \mu(\tilde{p}+\Delta)] = [\tilde{p}k, \tilde{p}k+n-\mu(\tilde{p}+\Delta)]$. This is similar to the way that we obtained Equations (6),(7) from Equations (3),(4) and (5). Furthermore, since we factored out the term $t^{k\tilde{p}}$, it follows that

$$\Phi_{p,k} = \Phi_{p+\delta_0,k}. \tag{18}$$

We now wish to better understand the value of $\Phi_{\tilde{p},k} \circ f$. Equations (14),(15) and (17) imply that one can write $(\Phi_{\tilde{p},k} \circ f)(r)$ as

$$(\Phi_{\tilde{p},k} \circ f)(r) = \sum_{\vec{a} \in \{0,1\}^\mu} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot K_{(\tilde{p} + \sum_{i=1}^{\mu} a_i \cdot \delta_i), r'_{\vec{a}}}(f) \cdot (\tilde{p} + \sum_{i=1}^{\mu} a_i \cdot \delta_i), \tag{19}$$

where

$$r'_{\vec{a}} \triangleq r - k\tilde{p} + \sum_{i=1}^{k} (1-a_i) \cdot (i-1) \cdot \delta_i + \sum_{i=k+1}^{\mu} (1-a_i) \cdot i \cdot \delta_i.^{9}$$

[8] In the proof of Claim 4.2 we use the more general notation $c_{\vec{a},j,k}(i)$.

[9] Notice that $r'_{\vec{a}} \in [0, n - \mu(\tilde{p} + \sum_{i=1}^{\mu} a_i \cdot \delta_i)]$.

Rewriting (19) gives

$$(\Phi_{\tilde{p},k} \circ f)(r) = L_{\tilde{p},r}(f) \cdot \tilde{p} + \sum_{i=1}^{\mu} M_{\tilde{p},i,r}(f) \cdot \delta_i, \tag{20}$$

where

$$L_{\tilde{p},r}(f) \triangleq \sum_{\vec{a} \in \{0,1\}^{\mu}} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot K_{(\tilde{p}+\sum_{i=1}^{\mu} a_i \cdot \delta_i), r'_{\vec{a}}}(f) \tag{21}$$

and

$$M_{\tilde{p},j,r}(f) \triangleq \sum_{\vec{a} \in \{0,1\}^{\mu} : a_j = 1} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot K_{(\tilde{p}+\sum_{i=1}^{\mu} a_i \cdot \delta_i), r'_{\vec{a}}}(f). \tag{22}$$

From the bound in Equation (12) it follows that

$$|L_{\tilde{p},r}(f)| < 2^{3\mu} \cdot n^{0.475-\epsilon} \quad \text{and} \quad |M_{\tilde{p},i,r}(f)| < 2^{3\mu-1} \cdot n^{0.475-\epsilon}. \tag{23}$$

The following claim shows that we actually have $L_{p,r}(f) = L_{p+\delta_0,r}(f) = 0$, so, in fact,

$$(\Phi_{\tilde{p},k} \circ f)(r) = \sum_{i=1}^{\mu} M_{\tilde{p},i,r}(f) \cdot \delta_i. \tag{24}$$

Therefore,

$$|(\Phi_{\tilde{p},k} \circ f)(r)| \leq 2^{3\mu-1} \cdot n^{0.475-\epsilon} \cdot \Delta \leq 2^{3\mu-1} \cdot n^{0.475}. \tag{25}$$

Claim 4.3. $L_{p,r}(f) = L_{p+\delta_0,r}(f) = 0$.

We defer the proof of the claim to Section 4.2 and proceed with the proof of the theorem. The good thing about Equation (25) is that it will allow us to reduce to the case of a polynomial with a bounded range. This somewhat resembles the way that we concluded the proof of Theorem 3.4, although it is done in a slightly more involved manner. Let

$$\Upsilon(t) = \prod_{i=0}^{\mu} \Phi_{p,i}(t) \quad \text{and} \quad \Upsilon_k(t) = \frac{\Upsilon(t)}{\Phi_{p,k}}.$$

We now bound the value of

$$g(r) \triangleq (\Upsilon \circ f)(r)$$

for $r \in [kp, kp + n - \mu(p + \Delta) - \deg(\Upsilon_k)]$. Notice that $g(r) = (\Upsilon_k \circ (\Phi_{p,k} \circ f))(r)$. Furthermore, $\Upsilon_k(t) = \prod_{i \neq k} \Phi_{p,i}(t)$. Claim 4.2 implies that each $\Phi_{p,i}(t)$ contains 2^{μ} monomials [10], and that its coefficients are upper bounded (in

[10] Note that here we allow different monomials with the same exponent.

absolute value) by 2^{μ}. Therefore, since $\Upsilon_k(t)$ is a product of μ such $\Phi_{p,i}$'s, it follows that $\Upsilon_k(t)$ is a sum of 2^{μ^2} monomials with coefficients upper bounded (in absolute value) by 2^{μ^2} . Moreover, as a polynomial, the degree of each $\Phi_{p,i}(t)$ is at most $\mu \cdot \Delta$ (this follows as $c_{\vec{a},k,k} \le \mu$). Hence, the degree of $\Upsilon_k(t)$ is at most $\mu^2 \cdot \Delta$. Thus, we have that $\Upsilon_k(t) = \sum_{i=1}^{2^{\mu^2}} \alpha_i \cdot t^{d_i}$ where $0 \le d_i \le \mu^2 \cdot \Delta$ and $|\alpha_i| \le 2^{\mu^2}$. This implies that for every $k \in [0,\mu]$ and every[11]

$$r \in I_k \triangleq [kp, kp + n - \mu(p + \Delta) - \deg(\Upsilon_k)],$$

we have that

$$\begin{aligned} |g(r)| = |(\Upsilon_k \circ (\Phi_{p,k} \circ f))(r)| &= \left| \sum_{i=1}^{2^{\mu^2}} \alpha_i \cdot (\Phi_{p,k} \circ f)(r + d_i) \right| \\ &\le \sum_{i=1}^{2^{\mu^2}} |\alpha_i| \cdot |(\Phi_{p,k} \circ f)(r + d_i)| \le 2^{\mu^2} \cdot 2^{\mu^2} \cdot 2^{3\mu - 1} \cdot n^{0.475} \le n^{0.475 + o(1)}, \end{aligned} \tag{26}$$

where we also used the bound on $|\Phi_{p,k} \circ f|$ given in (25). Notice that the size of the interval I_k satisfies

$$\begin{aligned} |I_k| &= n - \mu(p + \Delta) - \deg(\Upsilon_k) + 1 \\ &> n - \mu\left(\frac{n}{\mu + 1} - \Gamma(n)\right) - \deg(\Upsilon_k) + 1 > \frac{n}{\mu + 1} > p \end{aligned}$$

and therefore every two consecutive intervals I_k and I_{k+1} have a nonzero intersection. Hence, we conclude that for every $r \in [0, n - \mu\Delta - \deg(\Upsilon_\mu)]$ (note that $n - \mu\Delta - \deg(\Upsilon_\mu)$ is the endpoint of I_μ) it holds, by (26), that $|g(r)| \le n^{0.475 + o(1)} < n^{0.5}$. We thus have that

$$g\colon [0, n - \mu\Delta - \deg(\Upsilon_\mu)] \to [-n^{0.5}, n^{0.5}]. \tag{27}$$

In addition we have (by Lemma 2.8) that

$$\deg(g) \le \deg(f) < \mu p. \tag{28}$$

We now would like to show that $\deg(g)$ is much smaller than μp and then use Lemma 2.9 and Lemma 3.2 to conclude that f is of degree at most 1. Before applying Lemma 2.9, we must ensure that $\Phi_{p,k}(t) \neq 0$.

Claim 4.4. *For every $k \in [0,\mu]$ it holds that $\Phi_{p,k}(t) \neq 0$.*

[11] The drop by $\deg(\Upsilon_k)$ in the range of relevant r's is so that $r + d_i$ will be in the range $[kp, kp + n - \mu(p + \Delta)]$.

We defer the proof of Claim 4.4 and continue with the proof of the Theorem. Assume first that g is not a constant. The point is that now we can repeat the whole proof for g instead of f, with $n'=n-\mu\Delta-\deg(\Upsilon_\mu)$ instead of n. Note that due to the bound on the range of g we get that Equation (12), applied to g instead of f, gives

$$|K_{q,r}(g)| < \frac{2^\mu \cdot n^{0.5}}{q} < \frac{2^\mu \cdot n^{0.5}}{n/(\mu+2)} < 1.$$

Thus $K_{q,r}(g) = 0$. Continuing, we see that $(\Phi_{\tilde{p},k} \circ g)(r) = 0$ for $r \in [\tilde{p}k, \tilde{p}k + n' - \mu(\tilde{p}+\Delta)]$. Therefore, if we define $h = \Upsilon \circ g$ then for every $k \in [0,\mu]$ and $r \in I'_k \triangleq [kp, kp + n' - \mu(p+\Delta) - \deg(\Upsilon_k)]$ we have that $h(r)=0$. As before, we see that any two consecutive intervals I'_k and I'_{k+1} have a nonzero intersection. Indeed

$$\begin{aligned} |I'_k| &= n' - \mu(p+\Delta) - \deg(\Upsilon_k) + 1 \\ &= n - \mu p - 2\mu\Delta - \deg(\Upsilon_k) - \deg(\Upsilon_\mu) + 1 \\ &>^{(*)} n - \mu\left(\frac{n}{\mu+1} - \Gamma(n)\right) - 2(\mu\Delta + \mu^2\Delta) \\ &> \frac{n}{\mu+1} > p, \end{aligned}$$

where inequality $(*)$ follows from the properties of the construction in Lemma 4.1. It therefore follows that $h(r)$ is zero for all $r \in [0, n' - \mu\Delta - \deg(\Upsilon_\mu)]$. Since

$$\deg(h) \le \deg(g) \le \deg(f) < (\mu+1)p < n' - \mu\Delta - \deg(\Upsilon_\mu),$$

we get that $h \equiv 0$. By Lemma 2.9,

$$\deg(g) \le \operatorname{spar}(\Upsilon) - 2.$$

Applying Lemma 2.9 again yields that[12]

$$\deg(f) \le \deg(g) + \operatorname{spar}(\Upsilon) - 1 \le 2 \cdot \operatorname{spar}(\Upsilon) - 3 \le 2^{\mu^2+\mu+1} - 3 = o(n). \quad (29)$$

Lemma 3.2 now implies that f is of degree at most 1. This completes the proof of the theorem (the omitted proofs are given in Sections 4.1 and 4.2). ∎

Corollary 1.2 follows immediately from Theorem 1.1. Indeed, as S is contained in and not equal to the domain $[0,n]$, any function with degree at most 1 is in fact a constant function.

[12] If $g \equiv 0$ then one needs to replace $\deg(g)$ by -1 in (29).

4.1. A cube of primes

We shall now prove Lemma 4.1. As in the proof of Lemma 3.3, the proof of Lemma 4.1 is by the pigeonhole principle and relies on Theorem 2.6.

Proof of Lemma 4.1. The high level idea is the same as in the proof of Lemma 3.3. However, since we are looking for μ-dimensional 'cubes' it will be convenient to first prove the following combinatorial lemma. Note that the lemma does not necessarily concern prime numbers.

Lemma 4.5. *Let $A \subseteq [a_1, a_2]$ and let*

$$\ell = a_2 - a_1, \quad \alpha = |A|/\ell.$$

Then, if $r \leq \log\log(\ell) - \log\log(\frac{4}{\alpha})$, there is an r-dimensional 'cube' which is a subset of A

$$P_{x;\delta_1,\ldots,\delta_r} \triangleq \left\{ x + \sum_{i=1}^{r} a_i \cdot \delta_i \mid \forall i \; a_i \in \{0,1\} \right\} \subseteq A,$$

where $\delta_i > 0$ for $i = 1, 2, \ldots, r$.

Note that we do not require that the δ_i's are distinct.

Proof. We shall prove, by induction on r that for every $r \in [0, \log\log(\ell) - \log\log(\frac{4}{\alpha})]$, there exist $\delta_1, \ldots, \delta_r$ such that there are at least $\frac{\ell \cdot \alpha^{2^r}}{4^{2^r - 1}}$ r-dimensional cubes $P_{x;\delta_1,\ldots,\delta_r}$ (with different x's) inside A.

The case $r = 0$: This case is trivial as there are exactly $\ell \cdot \alpha = |A|$ elements in A, each is a 0-dimensional 'cube'.

The induction step: Assume that we already proved the claim for r and we wish to prove it for $r+1$. Consider the smallest number in each r-dimensional cube that was found in the r-th step. By the induction hypothesis we have $\frac{\ell \cdot \alpha^{2^r}}{4^{2^r - 1}}$ such different numbers, all of which in $A \subseteq [a_1, a_2]$. Looking at all the differences between those numbers, we get that if $\frac{\ell \cdot \alpha^{2^r}}{4^{2^r - 1}} \geq 2$ then there are at least $\binom{\frac{\ell \cdot \alpha^{2^r}}{4^{2^r - 1}}}{2} \geq \frac{1}{4} \left(\frac{\ell \cdot \alpha^{2^r}}{4^{2^r - 1}} \right)^2$ many such differences, all between 1 and ℓ. Using the pigeonhole principle, we conclude that there is a 'popular' difference, δ_{r+1}, with at least $\frac{1}{\ell} \cdot \frac{1}{4} \cdot \left(\frac{\ell \cdot \alpha^{2^r}}{4^{2^r - 1}} \right)^2$ many occurrences. For such a 'popular' difference δ_{r+1} and every pair of cubes at distance δ_{r+1} we have that

$$P_{x;\delta_1,\delta_2,\ldots,\delta_r} \cup P_{x+\delta_{r+1};\delta_1,\delta_2,\ldots,\delta_r} = P_{x;\delta_1,\delta_2,\ldots,\delta_r,\delta_{r+1}}.$$

This gives the required

$$\frac{1}{4\ell}\cdot\left(\frac{\ell\cdot\alpha^{2^r}}{4^{2^r-1}}\right)^2=\frac{\ell\cdot\alpha^{2^{r+1}}}{4^{2^{r+1}-1}}$$

$(r+1)$-dimensional cubes.

To conclude the proof of Lemma 4.5 we need to show that for $r \le \log\log(\ell)-\log\log(\frac{4}{\alpha})$, it holds that $\frac{\ell\cdot\alpha^{2^r}}{4^{2^r-1}}\ge 2$, which is equivalent to showing that $\ell \ge 2\cdot 4^{2^r-1}\cdot(\frac{1}{\alpha})^{2^r}$. It is clearly enough to show that $\ell \ge (\frac{4}{\alpha})^{2^r}$, which follows since $r\le\log\log(\ell)-\log\log(\frac{4}{\alpha})$. This completes the proof of the lemma. ∎

We now proceed with the proof of Lemma 4.1. Recall that we have to find δ_0 that will be much larger than the other δ_i's (in fact, it has to be much larger than their sum, as we consider ϵ which is relatively small). We therefore start by first choosing δ_0 and only then apply Lemma 4.5.

Let p,q be prime numbers such that:

$$q\in I_q \triangleq \left[\frac{n}{\mu+1}-2\Gamma(n),\frac{n}{\mu+1}-\Gamma(n)\right],$$
$$p\in I_p \triangleq \left[\frac{n}{\mu+1}-4\Gamma(n),\frac{n}{\mu+1}-3\Gamma(n)\right].$$

Clearly, $|I_p|=|I_q|=\Gamma(n)$ and $\Gamma(n)\le q-p\le 3\Gamma(n)$ for any such p and q. Theorem 2.6 implies that each of the intervals I_q,I_p contains at least $\frac{9}{100}\cdot\frac{\Gamma(n)}{\log n}$ different prime numbers. By the pigeonhole principle, each of the intervals I_p,I_q has a sub-interval of length n^ϵ that contains at least $\frac{1}{12}\cdot\frac{n^\epsilon}{\log n}$ many prime numbers. Denote these sub-intervals as I'_p,I'_q respectively:

$$I'_p=[r_p,r_p+n^\epsilon] \qquad I'_q=[r_q,r_q+n^\epsilon].$$

Looking at all the differences between pairs of primes in $I'_q\times I'_p$ we get that there are at least $(\frac{n^\epsilon}{12\cdot\log n})^2$ many differences, each of which is between $r_q-r_p-n^\epsilon$ and $r_q-r_p+n^\epsilon$. Hence, one of the differences occurs at least $(\frac{n^\epsilon}{12\cdot\log n})^2/2n^\epsilon=\frac{n^\epsilon}{2(12\cdot\log n)^2}$ many times. Let δ_0 be that popular difference. Clearly, property 4 holds from this choice of δ_0 . Consider the following set

$$A\triangleq\left\{x\in I'_p \mid x+\delta_0\in I'_q,\ x \text{ and } x+\delta_0 \text{ are primes}\right\}.$$

Obviously, $A\subseteq I'_p$, and by the choice of δ_0 we are guaranteed that $|A|\ge\frac{n^\epsilon}{2(12\cdot\log n)^2}$. Let $\alpha=|A|/|I'_p|\ge\frac{1}{2(12\cdot\log n)^2}$. Note that

$$\log\log(n^\epsilon)-\log\log\left(\frac{4}{\alpha}\right)$$

$$\geq \log\log(n) - \log\log\log(n) - \log(1/\epsilon) - O(1) > \frac{\log\log n}{2} = \mu.$$

We now apply Lemma 4.5 with parameters

$$\ell = |I'_p| = n^{\epsilon} \quad \text{and} \quad \alpha = |A|/|I'_p| \geq \frac{1}{2(12 \cdot \log n)^2}$$

and obtain that there exists an μ-dimensional cube $\mathcal{B} = P_{x;\delta_1,\ldots,\delta_\mu} \subseteq A$. By the definition of A it follows that all the elements in $\mathcal{B}+\delta_0 \triangleq \{b+\delta_0 \mid b \in \mathcal{B}\}$ are prime numbers. Our final $(r+1)$-dimensional cube is therefore,

$$P_{x;\delta_0,\delta_1,\ldots,\delta_\mu} = \left\{ x + \sum_{i=0}^{\mu} a_i \cdot \delta_i \mid \forall i \ a_i \in \{0,1\} \right\}.$$

We note that Lemma 4.5 also guarantees that all the δ_i's are positive and that

$$\varDelta \triangleq \sum_{i=1}^{n} \delta_i \leq |I'_p| = n^{\epsilon}.$$

∎

4.2. Omitted proofs

We now give the proofs of Claims 4.2, 4.3 and 4.4.

Proof of Claim 4.2. Recall that

$$\varPhi'_{\tilde{p},k}(t) = \sum_{\vec{a} \in \{0,1\}^\mu} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot \varPsi_{(\tilde{p}+\sum_{i=1}^{\mu} a_i \cdot \delta_i)}(t) \cdot t^{\sum_{i=1}^{k}(1-a_i)\cdot(i-1)\cdot\delta_i + \sum_{i=k+1}^{\mu}(1-a_i)\cdot i \cdot \delta_i}. \tag{30}$$

Denote

$$c_{\vec{a},j,k}(i) \triangleq \begin{cases} j & \text{if } a_i = 1 \\ i-1 & \text{if } a_i = 0 \text{ and } i \leq k \\ i & \text{if } a_i = 0 \text{ and } i \geq k+1 \end{cases}.$$

This is consistent with the previous definition of $c_{\vec{a},k,k}$ (see Equation (16)). By expanding $\varPsi$ (recall Equation (13)) and using the $c_{\vec{a},j,k}$'s we get that

$$\varPhi'_{\tilde{p},k}(t) = \sum_{\vec{a} \in \{0,1\}^\mu} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot \sum_{j=0}^{\mu} (-1)^j \cdot \binom{\mu}{j} \cdot$$

$$\cdot\, t^{j\tilde{p}+\sum_{i=1}^{\mu} c_{\vec{a},j,k}(i)\cdot\delta_i}$$

Considering the coefficients for different j's we have the following cases.

Case 1: $j < k$. For every $\vec{a} = (a_1,\ldots,a_j,0,a_{j+2},\ldots,a_\mu)$, let $\vec{b} = (a_1,\ldots,a_j,1,a_{j+2},\ldots,a_\mu)$. It is easy to verify that $c_{\vec{a},j,k} = c_{\vec{b},j,k}$. As $(-1)^{\sum_{i=1}^{\mu} a_i} = -(-1)^{\sum_{i=1}^{\mu} b_i}$ we get that $\vec{a}$ and $\vec{b}$ cancel each other.

Case 2: $j > k$. Quite similarly, for every $\vec{a} = (a_1,\ldots,a_{j-1},0,a_{j+1},\ldots,a_\mu)$, let $\vec{b} = (a_1,\ldots,a_{j-1},1,a_{j+1},\ldots,a_\mu)$. Again, $\vec{a}$ and $\vec{b}$ cancel each other.

Case 3: $j = k$. This is the only case where coefficients do not get canceled out. We therefore get that

$$\Phi'_{\tilde{p},k} = \sum_{\vec{a}\in\{0,1\}^\mu} (-1)^{\sum_{i=1}^{\mu} a_i} \cdot (-1)^k \cdot \binom{\mu}{k} \cdot t^{k\tilde{p}+\sum_{i=1}^{\mu} c_{\vec{a},k,k}(i)\cdot\delta_i},$$

as claimed. ∎

We now proceed to proving Claim 4.3. The specific properties of the cube (that may have seemed somewhat arbitrary) play a major role in this proof.

Proof of Claim 4.3. Recall that $\Phi_{p,k} = \Phi_{p+\delta_0,k}$ (Equation (18)). Therefore,

$$\begin{aligned} L_{p,r}(f)\cdot p + \sum_{i=1}^{\mu} M_{p,i,r}(f)\cdot\delta_i &= \Phi_{p,k}(r) = \Phi_{p+\delta_0,k}(r) \qquad (31)\\ &= L_{p+\delta_0,r}(f)\cdot(p+\delta_0) + \sum_{i=1}^{\mu} M_{p+\delta_0,i,r}(f)\cdot\delta_i. \end{aligned}$$

Rearranging (31) gives

$$\begin{aligned} &(L_{p,r}(f) - L_{p+\delta_0,r}(f))\cdot p\\ &= L_{p+\delta_0,r}(f)\cdot\delta_0 + \sum_{i=1}^{\mu} (M_{p+\delta_0,i,r}(f) - M_{p,i,r}(f))\cdot\delta_i. \end{aligned}$$

Recall that

$$|L_{p,r}(f)|, |L_{p+\delta_0,r}(f)| < 2^{3\mu}\cdot n^{0.475-\epsilon}$$

and

$$|M_{p,i,r}(f)|, |M_{p+\delta_0,i,r}(f)| < 2^{3\mu-1}\cdot n^{0.475-\epsilon}$$

(Equation (23)). By our choice of parameters we have that

$$\left| L_{p+\delta_0,r}(f)\cdot\delta_0 + \sum_{i=1}^{\mu} (M_{p+\delta_0,i,r}(f) - M_{p,i,r}(f))\cdot\delta_i \right|$$

$$\begin{aligned}&\leq 2^{3\mu} \cdot n^{0.475-\epsilon} \cdot (\delta_0 + \sum_{i=1}^{\mu} \delta_i)\\ &= n^{0.475-\epsilon} \cdot \Gamma(n) \cdot \operatorname{poly}\log(n) = n^{1-\epsilon} \cdot \operatorname{poly}\log(n) < p.\end{aligned}$$

As $(L_{p,r}(f) - L_{p+\delta_0,r}(f)) \cdot p$ is an integer multiple of p, it must be the case that $L_{p,r}(f) - L_{p+\delta_0,r}(f) = 0$. We now show that $L_{p+\delta_0,r}(f) = 0$ which will conclude the proof.

As we just proved that $L_{p,r}(f) - L_{p+\delta_0,r}(f) = 0$ we can rewrite (31) as

$$L_{p+\delta_0,r}(f) \cdot \delta_0 = -\sum_{i=1}^{\mu} (M_{p+\delta_0,i,r}(f) - M_{p,i,r}(f)) \cdot \delta_i.$$

Similarly to the previous argument we note that $L_{p+\delta_0,r}(f) \cdot \delta_0$ is an integer multiple of δ_0 and that, by our choice of parameters (Lemma 4.1)

$$\begin{aligned}&\left|\sum_{i=1}^{\mu} (M_{p+\delta_0,i,r}(f) - M_{p,i,r}(f)) \cdot \delta_i\right|\\ &< 2 \cdot 2^{3\mu-1} \cdot n^{0.475-\epsilon} \cdot \sum_{i=1}^{\mu} \delta_i \leq 2^{3\mu} \cdot n^{0.475} < \Gamma(n) \leq \delta_0.\end{aligned}$$

Hence, $L_{p+\delta_0,r}(f) = 0$. This completes the proof of the claim. ∎

Proof of Claim 4.4. By claim 4.2, $\Phi_{p,k}(t)$ is the sum of 2^{μ} (not necessarily different) monomials. To prove that the different monomials do not cancel each other we will show that there is a unique monomial of maximal degree. Note that for every $\vec{a} \in \{0,1\}^{\mu}$ we have a monomial of degree $\sum_{i=1}^{\mu} c_{\vec{a},k,k}(i) \cdot \delta_i$ in $\Phi_{p,k}(t)$. Let

$$\vec{a} \triangleq (\underbrace{1, 1, \ldots, 1}_{k}, \underbrace{0, 0, \ldots, 0}_{\mu-k}).$$

Then, for every other binary vector $\vec{a} \neq \vec{b} \in \{0,1\}^{\mu}$ we have the following: For $i \leq k$, $c_{\vec{b},k,k}(i) \leq k = c_{\vec{a},k,k}(i)$ and the inequality is strong if $b_i = 0$. For $i \geq k+1$, $c_{\vec{b},k,k}(i) \leq i = c_{\vec{a},k,k}(i)$ and the inequality is strong if $b_i = 1$.

As $\vec{a} \neq \vec{b}$, it follows that $c_{\vec{b},k,k} < c_{\vec{a},k,k}$. Namely,

$$\forall i \in [1,\mu]\colon c_{\vec{b},k,k}(i) \leq c_{\vec{a},k,k}(i) \text{ and } \exists i \in [1,\mu]\colon c_{\vec{b},k,k}(i) < c_{\vec{a},k,k}(i).$$

Since all the δ_i's are positive, we get that $\sum_{i=1}^{\mu} c_{\vec{b},k,k}(i) \cdot \delta_i < \sum_{i=1}^{\mu} c_{\vec{a},k,k}(i) \cdot \delta_i$, and the monomial that corresponds to $\vec{a}$ is the unique monomial of maximal degree. ∎

5. The range of a degree d polynomial

In this section we prove Theorem 1.3. It will be an easy corollary of Theorem 1.4 which we first prove. The proof is quite elementary and basically follows from averaging arguments. At the end of the section we present a possible approach for improving our results using the Chebyshev polynomials, however at this stage we get more general results using our simple argument. To ease the reading we repeat the statement of Theorem 1.4.

Theorem 5.1 (Theorem 1.4). *Let $f:\mathbb{R}\to\mathbb{R}$ be a degree d* monic *polynomial. Then, $\max_{i\in[0,n]}|f(i)|>\left(\frac{n-d}{2e}\right)^d$. In particular, if $f:\mathbb{Z}\to\mathbb{Z}$ is a degree d polynomial (not necessarily monic) then*

$$\max_{i\in[0,n]}|f(i)|>\frac{1}{d!}\cdot\left(\frac{n-d}{2e}\right)^d\geq\frac{1}{\sqrt{7d}}\cdot\left(\frac{n-d}{2d}\right)^d.$$

Proof of Theorem 1.4. For $d=1$ the theorem holds. So we can assume w.l.o.g that $d\geq 2$. Consider the factorization of f over $\mathbb{C}$,

$$f(x)=\prod_{i=1}^{d}(x-\alpha_i). \tag{32}$$

Recall that if $\alpha_i\in\mathbb{C}$ is a root of f then its conjugate $\bar{\alpha}_i$ is also a root of f. As we are interested in bounding the range of f from below, we can assume w.l.o.g. that all the roots of f are real. Indeed, for any complex α and real x it holds that $(x-\alpha)\cdot(x-\bar{\alpha})\geq(x-\mathcal{R}(\alpha))^2$, where $\mathcal{R}(\alpha)$ is the real part of α.

We would like to give a lower bound on the maximum (absolute) value of f by showing that the product $\prod_{i=0}^{n}f(i)$ is large. However, since some of the i's can be roots of f, or very close to roots of f, we need to remove them from the product first.

Call an element $i\in[0,n]$ an *approximate root* of f if there is a root of f, α_j (in the notations of Equation (32)), such that[13] $\mathrm{round}(\alpha_j)=i$. Clearly, there are at most d approximate roots in the set $[0,n]$. Denote with $S\subseteq[0,n]$ the set of all $i\in[0,n]$ such that i is not an approximate root. Clearly $|S|\geq n+1-d$. Note that

$$\max_{i\in[0,n]}|f(i)|\geq\left[\prod_{i\in S}|f(i)|\right]^{\frac{1}{|S|}}. \tag{33}$$

[13] $\mathrm{round}(x)$ is the integer closest to x, if $x=i+1/2$ then $\mathrm{round}(x)=i$. In other words, $\mathrm{round}(x)=\lceil x-1/2\rceil$.

As

$$\prod_{i \in S} |f(i)| = \prod_{j=1}^{d} \prod_{i \in S} |i - \alpha_j|, \tag{34}$$

it will suffice for our needs to bound from below the value of each product $\prod_{i \in S} |i - \alpha_j|$ and then apply it in Equation 33.

Fix some $j \in [d]$. Notice that the closest element to α_j in S has distance at least $1/2$ from it. The next element has distance at least 1 from it. The next has distance at least $3/2$ from it, etc. In other words, if we sort the elements in S according to their distances from α_j, $S = \{i_1, \ldots, i_{|S|}\}$, then the k element, i_k will be at distance at least $k/2$. Hence,

$$\begin{aligned} \prod_{i \in S} |i - \alpha_j| &\geq \prod_{k=1}^{|S|} |i_k - \alpha_j| \geq \prod_{k=1}^{|S|} \frac{k}{2} = \frac{|S|!}{2^{|S|}} \\ &\geq^* \left(\frac{|S|}{2e}\right)^{|S|} \cdot \sqrt{2\pi |S|}, \end{aligned} \tag{35}$$

where inequality $(*)$ follows from Stirling's formula (Theorem 2.1). Plugging Equation (35) back to Equations (34) and (33) we get

$$\begin{aligned} \max_{i \in [0,n]} |f(i)| &\geq \left[\left[\left(\frac{|S|}{2e}\right)^{|S|} \cdot \sqrt{2\pi |S|} \right]^d \right]^{\frac{1}{|S|}} \\ &= (2\pi |S|)^{\frac{d}{2|S|}} \cdot \left(\frac{|S|}{2e}\right)^d > \left(\frac{n-d}{2e}\right)^d. \end{aligned}$$

This proves the first statement of the theorem. For the second statement we note that if f is a polynomial mapping integers to integers then by Theorem 2.3 the coefficient of x^d in f is an integer multiple of $1/d!$. In particular there is an integer $c \neq 0$ such that $(d!/c) \cdot f(x)$ is monic. Therefore,

$$\begin{aligned} \max_{i \in [0,n]} |f(i)| &= \left|\frac{c}{d!}\right| \cdot \max_{i \in [0,n]} \left|\frac{d!}{c} \cdot f(i)\right| > \frac{1}{d!} \cdot \left(\frac{n-d}{2e}\right)^d \\ &\geq \frac{1}{\sqrt{7d}} \cdot \left(\frac{n-d}{2d}\right)^d, \end{aligned}$$

where we used Stirling's formula (and the assumption that $d \geq 2$) in the last inequality. ∎

We believe that Theorem 1.4 can be improved. Nevertheless, the next example shows that the theorem is not far from being tight.

Example 5.2. For an odd integer n and an even integer $d \leq n$, the polynomial $f(x) = \binom{x - \frac{n-d+1}{2}}{d}$ is a degree d polynomial mapping $[0, n]$ to $[0, \frac{n^d}{2^d \cdot d!}]$.

Proof. It is not difficult to see that since d is even, $f(x) = f(n-x)$. In particular, $f(x) \geq 0$ for all $x \in [0, n]$. Furthermore, for all $r \in [0, n]$

$$f(r) \leq f(n) = \binom{\frac{n+d-1}{2}}{d} < \frac{1}{d!} \cdot \left(\frac{n^2-1}{4}\right)^{d/2} < \frac{n^d}{2^d \cdot d!}. \qquad \blacksquare$$

This upper bound is larger by a factor of (roughly) e^d from the lower bound on the range that is stated in Theorem 1.4. It is an interesting question to understand the 'correct' bound.

To derive Theorem 1.3 we will need the following easy property of the function

$$\mathcal{D}_n(x) \triangleq \frac{1}{\sqrt{7x}} \cdot \left(\frac{n-x}{2x}\right)^x.$$

Lemma 5.3. *In the real interval* $[1, n]$ *the function* $\mathcal{D}_n(x)$ *is first strictly increasing and then strictly decreasing. Furthermore, it attains its maximum at some* $0.135 \cdot n < x < 0.136 \cdot n$ *(for* $n \geq 450$*).*

Proof. It is clearly sufficient to prove that the function

$$\begin{aligned}\ln(\mathcal{D}_n(x)) &= \ln\left(\frac{1}{\sqrt{7x}} \cdot \left(\frac{n-x}{2x}\right)^x\right) \\ &= x\ln(n-x) - x\ln x - x\ln 2 - \frac{1}{2}\ln x - \frac{1}{2}\ln 7\end{aligned}$$

has the claimed property. This will follow from the observation that the second derivative of $\ln(\mathcal{D}_n(x))$ is negative. Indeed,

$$(\ln(\mathcal{D}_n(x)))' = \ln(n-x) - \frac{x}{n-x} - \ln(x) - 1 - \ln(2) - \frac{1}{2x}$$

and

$$(\ln(\mathcal{D}_n(x)))'' = -\frac{1}{n-x} - \frac{n}{(n-x)^2} - \frac{1}{x} + \frac{1}{2x^2} < 0$$

where the last inequality holds since $x \geq 1$.

To see the 'furthermore' part we note that $(\ln(\mathcal{D}_n))'(0.135 \cdot n) > 0$ for $n \geq 450$ and that $(\ln(\mathcal{D}_n))'(0.136 \cdot n) < 0$ for every n. Hence, by the intermediate value theorem, $(\ln(\mathcal{D}_n(x)))' = 0$ for some $0.135 \cdot n < x < 0.136 \cdot n$ (when $n \geq 450$). ∎

We denote the unique maximum point of $\mathcal{D}_n$ as $x_{\mathcal{D}_n}$.

We can now derive Theorem 1.3.

Proof of Theorem 1.3. If $\deg(f) \le d-1$ we are done. We may therefore assume that $\deg(f) \ge d$. If $\deg(f) \le x_{\mathcal{D}_n}$ then by Theorem 1.4 and Lemma 5.3, we get that the maximal value that f attains on $[0,n]$ is larger than $\mathcal{D}_n(\deg(f)) \ge \mathcal{D}_n(d) > \frac{1}{\sqrt{7d}} \cdot \left(\frac{n-d}{2d}\right)^d$, in contradiction to the assumption of the theorem. Since $\mathcal{D}_n(x)$ is decreasing for $x > x_{\mathcal{D}_n}$ we observe, by substituting $x = \frac{1}{3}n - 1.2555 \cdot [d\ln(\frac{n-d}{2d}) - \frac{1}{2}\ln(\frac{n}{d})]$ into $\mathcal{D}_n$, that

$$\mathcal{D}_n\left(\frac{1}{3}n - 1.2555 \cdot \left[d\ln\left(\frac{n-d}{2d}\right) - \frac{1}{2}\ln\left(\frac{n}{d}\right)\right]\right) > \frac{1}{\sqrt{7d}} \cdot \left(\frac{n-d}{2d}\right)^d.$$

Indeed, it is not hard to see that for any c such that $c < n/3 - 0.136 \cdot n$ (which in particular means that $x_{\mathcal{D}_n} < n/3 - c$) it holds that

$$\begin{aligned}
\mathcal{D}_n(n/3-c) &= \frac{1}{\sqrt{7(n/3-c)}} \cdot \left(\frac{n-(n/3-c)}{2n/3-2c}\right)^{n/3-c} \\
&= \frac{1}{\sqrt{7(n/3-c)}} \cdot \left(1 + \frac{3c/2}{n/3-c}\right)^{n/3-c} \\
&\ge^{(*)} \frac{1}{\sqrt{7n/3}} \cdot e^{0.531 \cdot 3c/2} \\
&= \sqrt{3} \cdot \frac{1}{\sqrt{7d}} \cdot e^{0.7965 \cdot c - \frac{1}{2}\ln(n/d)},
\end{aligned}$$

where to prove inequality $(*)$ we used the simple fact that $(1+x) \ge e^{0.531 \cdot x}$ for $x \le 2.1765$, together with the bound on c. In our case, since $d \le \frac{2}{15}n$, it is not hard to verify that $c \triangleq 1.2555 \cdot [d\ln(\frac{n-d}{2d}) + \frac{1}{2}\ln(\frac{n}{d})]$ satisfies $c < n/3 - 0.136 \cdot n$ (for n large enough) as required.

We therefore obtain that

$$\begin{aligned}
&\mathcal{D}_n\left(\frac{1}{3}n - 1.2555 \cdot \left[d\ln\left(\frac{n-d}{2d}\right) - \frac{1}{2}\ln\left(\frac{n}{d}\right)\right]\right) \\
&\ge \sqrt{3} \cdot \frac{1}{\sqrt{7d}} \cdot e^{0.7965 \cdot c - \frac{1}{2}\ln(n/d)} \\
&> \frac{1}{\sqrt{7d}} \cdot e^{d\ln\left(\frac{n-d}{2d}\right)} = \frac{1}{\sqrt{7d}} \cdot \left(\frac{n-d}{2d}\right)^d,
\end{aligned}$$

as claimed. By Lemma 5.3, $\deg(f) \ge \frac{1}{3}n - 1.2555 \cdot \left[d\ln\left(\frac{n-d}{2d}\right) - \frac{1}{2}\ln\left(\frac{n}{d}\right)\right]$. ∎

To summarize, Theorem 1.3 uses the fact that $\mathcal{D}_n$ has a unique maximum, $x_{\mathcal{D}_n}$, and aims to find, for a given degree $d < x_{\mathcal{D}_n}$, another degree $d' > x_{\mathcal{D}_n}$ such that $\mathcal{D}_n(d') \ge \mathcal{D}_n(d)$. In the theorem we gave a relatively simple way to derive d' from d. With more work one can push this result for d's closer to $x_{\mathcal{D}_n}$.

We note that Theorem 1.3 implies that when $\Omega(n) \leq \deg(f) < (1-\epsilon)n/3$ then the range of f is exponential in n. As a corollary of Example 5.2 one can show that if we allow the range to be as large as $O\left(\left(\frac{1+\sqrt{5}}{2}\right)^n\right)$ then f can have any degree. Indeed, taking the maximum over $\binom{\frac{n+d-1}{2}}{d}$, when $d+n$ is odd, we get an upper bound on that range that is smaller than the n-th Fibonacci number, FIB_n.

Lemma 5.4. *For integers d, n such that $n+d$ is odd, let $R_{n,d} \triangleq \binom{\frac{n+d-1}{2}}{d}$, and set*

$$R_n \triangleq \max\{R_{n,d} \mid d \in [0,n], d+n \textit{ is odd }\}.$$

Then, $R_n \leq R_{n-1} + R_{n-2}$ for $n > 2$.

Proof. Since $n > 2$, we can assume that the maximum of $R_{n,d}$ is achieved for some $d > 0$. We use the combinatorial identity $\binom{m}{k} = \binom{m-1}{k} + \binom{m-1}{k-1}$ to conclude that:

$$\begin{aligned} R_{n,d} &= \binom{\frac{n+d-1}{2}}{d} = \binom{\frac{n+d-1}{2} - 1}{d} + \binom{\frac{n+d-1}{2} - 1}{d-1} \\ &= \binom{\frac{(n-2)+d-1}{2}}{d} + \binom{\frac{(n-1)+(d-1)-1}{2}}{d-1} \\ &= R_{n-2,d} + R_{n-1,d-1}. \end{aligned}$$

Maximizing over d in both sides we conclude that $R_n \leq R_{n-2} + R_{n-1}$. ∎

As an immediate corollary, using the fact that $R_1 = R_2 = 1$, we deduce that

$$R_n \leq \mathrm{FIB}_n \leq \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n,$$

which completes our argument.

5.1. A possible route for improvements

In this section we present a possible approach towards improving Theorem 1.3, when $d \leq \sqrt{n}/2$, based on Chebyshev polynomials. We will only give a sketch of the approach and we will not cover all necessary background on Chebyshev polynomials. The interested reader is referred to [14].

A natural approach to proving that a polynomial must take large values is by comparing it to the Chebyshev polynomial of the same degree. Roughly, the Chebyshev polynomial of degree d is defined on the real interval $[-1,1]$ in the following way:

$$T_d(x) = \cos(d \arccos(x)).$$

It is not hard to prove that T_d is a degree d polynomial, having exactly d roots in the interval $[-1,1]$, that its leading coefficient is 2^{d-1} and that it has $d+1$ extremal values in the same interval, on which it is equal, in absolute value, to 1. Specifically, its roots lie on the points $\cos(\frac{\pi(2k-1)}{2d})$ and its extremal points are $\cos(\frac{\pi k}{d})$, on which it alternates between 1 and -1. A well known fact of the Chebyshev polynomials is that among the degree d monic polynomials the polynomial $f_d(x) = 2^{1-d}T_d(x)$ whose maximum on the real interval $[-1,1]$ is the smallest and equals 2^{1-d}.

The problem in using this fact is that we are interested in the maximum of a function on a relatively small set of points. Consider a polynomial $f\colon [0,n] \to [0,m]$. Let $g(x) = f(\frac{n}{2}x + \frac{n}{2})$. Thus $g\colon [-1,1] \to [0,m]$, (where $[-1,1]$ is the real interval) and we are interested in the value of g on the points $\{-1, -1+\frac{2}{n}, -1+\frac{4}{n}, \ldots, 1\}$. Denote for simplicity $x_k = 2k/n - 1$, $k = 0, \ldots, n$. We would like to say that as T_d obtains the smallest maximum on $[-1,1]$ then (after we normalize g by its leading coefficient) it must obtain a value larger than 2^{1-d} on one of the x_k's. However, all that we know is that the maximum of g on the whole interval $[-1,1]$ is large and not necessarily on one of the x_k's.

To tackle this problem one has to prove that the values that T_d obtains on the x_k's is relatively large (close to its overall maximum). A possible way for proving this is by observing that we can find a point x_k near any extremal point and then, since we have a reasonable bound on the derivative of T_d, conclude that T_d obtains a relatively large value there as well. This approach in fact works; Since the derivative of T_d is bounded by d^2 it follows that when $d < \sqrt{n}/2$ there are $d+1$ points among the x_k's on which T_d alternates in sign and obtains absolute value larger than, say, $1/2$. Now, let $\tilde{g} = g/g_d$, where g_d is the leading coefficient of g. Assume that $|\tilde{g}(x_k)| < \frac{1}{2} \cdot 2^{1-d}$, for every k. Then the polynomial $2^{1-d}T_d - \tilde{g}$ has degree at most $d-1$ (it is the difference of two degree d monic polynomials) and it changes sign d times (between the x_k's on which T_d obtains large value), which is a contradiction. It therefore follows that $\max_{k \in [0,n]} |g(x_k)| \geq \frac{1}{2}|g_d| \cdot 2^{1-d}$. As g_d equals $f_d \cdot (n/2)^d$, where f_d is the leading coefficient of f, and since $|f_d| \geq \frac{1}{d!}$, we get that $\max_{k \in [0,n]} |f(k)| = \max_{k \in [0,n]} |g(x_k)| \geq 2^{-d} \cdot (n/2)^d / d! = \frac{n^d}{2^{2d} d!}$. We summarize this in the next theorem.

Theorem 5.5. *There exists a constant* n_0 *such that for every two integers* d,n *such that* $n > n_0$ *and* $d \le \sqrt{n}/2$ *it holds that if* $f\colon \mathbb{Z} \to \mathbb{Z}$ *is a degree* d *polynomial (not necessarily monic) then* $\max_{i\in[0,n]} |f(i)| \ge \frac{n^d}{2^{2d} d!}$.

This result is slightly better than the bound $\max_{i\in[0,n]} |f(i)| \ge \frac{1}{d!} \cdot \left(\frac{n-d}{2e}\right)^d$ that was obtained in the proof of Theorem 1.4, but it holds only for $d \le \sqrt{n}/2$. We note, however, that this approach cannot work for $d = \omega(\sqrt{n})$ as for such large d many roots of T_d are very close to each other. Indeed, the distances among the first roots (and among the last roots) are smaller than $1/n$ while the x_k's are separated from one another. For that reason we cannot use Theorem 5.5 instead of Theorem 1.4; In order to show that the degree must be larger than $\Omega(n)$ we must claim something about the range of polynomials of degree, say, $n/\log(n)$ and Theorem 5.5 does not give any information in this case.

5.2. The case of small degrees

In this section we give two small improvements for the case of polynomials of degrees 1 or 2. The first improvement concerns polynomials whose range is (roughly) $[0, n^{2.475}]$.

Theorem 5.6. *For every* $0 < \epsilon$ *there exists* n_0 *such that for every integer* $n_0 < n$ *the following holds: Every*

$$f\colon [0,n] \to \left[0, n^{2.475-\epsilon}\right]$$

must satisfy $\deg(f) \le 2$ *or* $\deg(f) \ge n/2 - 2n/\log\log n$.

Notice that Theorem 1.3 implies that if the range of f is, say, $[0, n^3/1000]$ then either $\deg(f) \le 2$ or $\deg(f) \ge n/3 – O(\log n)$. Thus, the improvement that Theorem 5.6 gives is that if the range is $[0, n^{2.475-\epsilon}]$ then either $\deg(f) \le 2$ (as before) or it is at least $n/2 - 2n/\log\log n$ (compared to roughly $n/3$). The proof is quite similar to the proof of Lemma 3.1.

Proof. We first explain how n_0 is defined. A corollary of Theorem 1.1 is that there exists n_1 such that for every $n > n_1$ and $f\colon [0,n] \to [0, 17n^{1.475-\epsilon}]$, either $\deg(f) \le 1$ or $\deg(f) > n - 4n/\log\log n$. Define n_2 (guaranteed to exist from Theorem 2.6) such that for every $n > n_2$ it holds that there is a prime number in the range $[\frac{n}{2} - \Gamma(n), \frac{n}{2}]$ and such that $\Gamma(n) = n^{0.525} < \frac{n}{2} - \frac{n}{3}$. We set $n_0 = \max(2n_1, n_2)$.

The proof is by a reduction to Theorem 1.1. Let $p \in [\frac{n}{2} - \Gamma(n), \frac{n}{2}]$ be a prime number. If $\deg(f) \geq p$ then we are done, as in this case

$$\deg(f) \geq p \geq \frac{n}{2} - \Gamma(n) \geq \frac{n}{2} - 2n/\log\log n.$$

Therefore, we may assume that $\deg(f) < p$. By Lemma 2.4, working modulo p, we get that $f(r) \equiv_p f(p+r)$ for every $r \in [0, n-p]$. As in the proof of Lemma 3.1, we consider the polynomial $g(r) = \frac{f(r)-f(r+p)}{p}$ which is defined over $r \in [0, n-p]$. It follows that

$$g\colon [0, n-p] \to \left[\frac{-n^{2.475-\epsilon}}{p}, \frac{n^{2.475-\epsilon}}{p}\right] \subseteq \left[-3 \cdot n^{1.475-\epsilon}, 3 \cdot n^{1.475-\epsilon}\right].$$

In particular, $g + 3 \cdot n^{1.475-\epsilon}$ maps $[0, n/2]$ to

$$\left[0, 6 \cdot n^{1.475-\epsilon}\right] \subseteq \left[0, 17(n/2)^{1.475-\epsilon}\right].$$

Since $n > n_0 \geq 2n_1$ Theorem 1.1 implies that either $\deg(g) \leq 1$ or $\deg(g) > n/2 - 2n/\log\log n$. By Lemma 2.9 we get that $\deg(f) \leq \deg(g) + 1$ and so the case $\deg(g) \leq 1$ translates to $\deg(f) \leq 2$. In the second case where $\deg(g) > n/2 - 2n/\log\log n$ we get the same conclusion for f as $\deg(g) \leq \deg(f)$. ∎

As an immediate corollary we get our second improvement that provides a strengthening of Lemma 3.2.

Corollary 5.7. *There exists a constant n_0 such that if $n > n_0$ and $f\colon [0,n] \to \left[0, \left\lfloor \frac{n^2 - 4\Gamma(n)^2}{8} \right\rfloor\right]$ is a polynomial then $\deg(f) \leq 1$ or $\deg(f) \geq n/2 - 2n/\log\log n$.*

Proof. Lemma 3.2 implies that if $\deg(f) > 1$ then it is at least $n/12 - \Gamma(n)$. However, by Theorem 5.6 we get that actually $\deg(f) \geq n/2 - 2n/\log\log n$. ∎

The example given after Lemma 3.2, $f(x) = \binom{x - \frac{n-1}{2}}{2}$, gives a degree 2 polynomial mapping $[0, n]$ to $\left[0, \frac{n^2-1}{8}\right]$. Thus, up to an additive $O(n^{1.05})$ term, the range in Corollary 5.7 is tight.

6. Proof of Theorem 1.5

In this section we prove Theorem 1.5. The proof is based on a reduction to the Shortest Vector Problem (SVP) in Lattice Theory. In section 6.1 we introduce basic definitions and tools from lattice theory. We then turn to prove Theorem 1.5 in section 6.2.

6.1. Basic properties of lattices

Definition 6.1. *Let $b_1, b_2, \ldots, b_n$ be linearly independent vectors in $\mathbb{R}^m$ (obviously $n \le m$). We define the lattice generated by them as*

$$\Lambda(b_1, b_2, \ldots, b_n) = \left\{ \sum_{i=1}^{n} x_i b_i \colon x_i \in \mathbb{Z} \right\}.$$

We refer to $b_1, b_2, \ldots, b_n$ as a *basis* of the lattice. More compactly, if B is the $m \times n$ matrix whose columns are $b_1, b_2, \ldots, b_n$, then we define

$$\Lambda(B) = \Lambda(b_1, b_2, \ldots, b_n) = \{Bx \colon x \in \mathbb{Z}^n\}.$$

We say that the *rank* of the lattice is n and its *dimension* is m. The lattice is called a *full-rank lattice* if $n = m$. The determinant of $\Lambda(B)$ is defined as $\det(\Lambda(B)) = \sqrt{\det(B^T B)}$. Although a basis of a lattice is not unique, e.g., both $\left\{(0,1)^T, (1,0)^T\right\}$ and $\left\{(1,1)^T, (2,1)^T\right\}$ span $\mathbb{Z}^2$, it can be shown that the determinant of a lattice is independent of the choice of basis.

Definition 6.2. *Let K be a bounded and open convex set in $\mathbb{R}^n$, which is symmetric around the origin. Let Λ be a lattice of rank n. For $i \in [n]$, the i-th* successive minimum *with respect to K is defined as*

$$\lambda_i(\Lambda, K) = \inf\{r \colon \dim(\operatorname{span}(\Lambda \cap rK)) > i\}$$

where $rK = \{rx \colon x \in K\}$.

We shall need the following theorem, due to Minkowski. A proof can be found in, e.g., [9].

Theorem 6.3. *For any full-rank lattice Λ of rank n,*

$$\prod_{i=1}^{n} \lambda_i(\Lambda, K) \cdot \operatorname{vol}(K) \le 2^n \det \Lambda.$$

We will take K to be the set $(-1,1)^n$. Thus, K has volume 2^n, and it is clearly a bounded and open convex set, which is symmetric around the origin. For this K, Theorem 6.3 gives an upper bound on the length of shortest vectors in lattices with respect to the L^∞ norm. Note that this is slightly unusual, as in most applications one considers the shortest vectors with respect to the L^2 norm.

6.2. Proof of Theorem 1.5

The idea behind the proof of Theorem 1.5 is roughly as follows. We identify each function $f\colon [0,n]\to\mathbb{Z}$ with its set of values $(f(0),f(1),\ldots,f(n))$. That is, we think of functions as vectors in $\mathbb{Z}^{n+1}$. We shall construct a lattice in $\mathbb{R}^{n+1}$ which is not full-rank, and contains only points representing polynomials of degree $\deg(f)\le n-k$. We then prove that this lattice has many (at least $2k+2$) linearly independent short vectors with L^∞-norm smaller than $O(2^k)$, i.e. many linearly independent polynomials whose image is (somewhat) bounded. One of these polynomials must be of degree at least $2k+1$. For technical reasons we will not work with the lattice described above but rather we shall consider a full rank lattice obtained by adding 'long' orthogonal vectors to the basis of our initial lattice.

Proof of Theorem 1.5. Set $D=n-k$ and let $m=O(2^k)$.[14] We now describe the basis for the lattice. For $i\in[0,D]$ define the vector $b_i\in\mathbb{R}^{n+1}$ as follows: $(b_i)_j=\binom{j}{i}$, for $j=0,\ldots,n$. Notice that b_i corresponds to the polynomial $f_i(x)=\binom{x}{i}$. Let $b_{D+1},\ldots,b_n\in\mathbb{R}^{n+1}$ be arbitrary vectors of length $M\triangleq(m/2+1)\cdot\sqrt{n+1}$, such that for every $i\in[D+1,n]$, b_i is orthogonal to b_k for all $k\ne i$ (we can find such b_i by, say, the Gram-Schmidt procedure). Denote by B the matrix whose columns are $b_0,\ldots,b_n$ and let $\Lambda_{n,D}=\Lambda(B)$.

Lemma 6.4.

$$\det(\Lambda_{n,D})\le 2^{(n+D+1)(n-D)/2}\cdot M^{n-D}.$$

We defer the proof of the lemma and continue with the proof of Theorem 1.5. By a theorem of Minkowski (see Theorem 6.3) and the choice $K=(-1,1)^{n+1}$, we get

$$\prod_{i=1}^{n+1}\lambda_i(\Lambda_{n,D},K)\cdot\mathrm{vol}(K)\le 2^{n+1}\cdot\det\Lambda_{n,D}. \tag{36}$$

Note that for $i\ge D+2$, $\lambda_i(\Lambda_{n,D},K)\ge M/\sqrt{n+1}$. Indeed, if u is a point in $\Lambda_{n,D}$ with a non-zero coefficient for some b_i, $i\ge D+1$, then by orthogonality and the fact that the length of such b_i is M, we have that u has L^2 norm at least M, and hence its L^∞ norm is at least $M/\sqrt{n+1}$. Combining this observation with Equation (36), the fact that $\mathrm{vol}(K)=2^{n+1}$ and Lemma 6.4, we get

$$\prod_{i=1}^{D+1}\lambda_i(\Lambda_{n,D},K)\le 2^{(n+D+1)(n-D)/2}\cdot(\sqrt{n+1})^{n-D}. \tag{37}$$

[14] The exact value of m will be determined later.

Estimating the LHS from below gives

$$\prod_{i=1}^{D+1} \lambda_i(\Lambda_{n,D}, K) \geq \prod_{i=2k+2}^{D+1} \lambda_i(\Lambda_{n,D}, K) \geq \lambda_{2k+2}(\Lambda_{n,D}, K)^{D-2k}. \tag{38}$$

Combining Equations (37) and (38), we get

$$\begin{aligned} \lambda_{2k+2}(\Lambda_{n,D}, K) &\leq 2^{\frac{(n+D+1)(n-D)}{2(D-2k)}} \cdot (\sqrt{n+1})^{\frac{n-D}{D-2k}} = 2^{\frac{(2n-k+1)k}{2(n-3k)}} \cdot (\sqrt{n+1})^{\frac{k}{n-3k}} \\ &= 2^k \cdot 2^{O\left(\frac{k^2+k\log n}{n-3k}\right)} = O(2^k), \end{aligned} \tag{39}$$

where the last step is due to the assumption that $k = O(\sqrt{n})$. In particular, for a large enough n there is some constant β such that $\lambda_{2k+2}(\Lambda_{n,D}, K) \leq \beta 2^k$. Letting $m = 2\beta 2^k$, we get that $\lambda_{2k+2}(\Lambda_{n,D}, K) \leq m/2$. Hence, by definition of λ_{2k+2}, there are $2k+2$ linearly independent vectors, in $\Lambda_{n,D}$ whose L^∞-norm is not greater than $m/2$, i.e. they all lie in $\Lambda_{n,D} \cap [-m/2, m/2]^{n+1}$.

Let v be any such vector. Denote with $v = \sum_{i=0}^{n} \alpha_i b_i$ its representation according to the basis B. Recall that all the coefficients α_i are integers. As $\|v\|_2 \leq \|v\|_\infty \cdot \sqrt{n+1} \leq m/2 \cdot \sqrt{n+1} < M$ and since for every $j > D$, $\|b_j\|_2 = M$, we get, by orthogonality, that $\alpha_{D+1} = \alpha_{D+2} = \cdots = \alpha_n = 0$. Hence, for $\ell \in [0, n]$, the ℓ-th coordinate of v is equal to $v_\ell = \sum_{i=0}^{D} \alpha_i \binom{\ell}{i}$. Therefore, the polynomial $f_v(x) = \sum_{i=0}^{D} \alpha_i \binom{x}{i}$ satisfies $f_v(\ell) = v_\ell$ for every $\ell \in [0, n]$. As $v \in [-m/2, m/2]^{n+1}$ we get that $f_v(x) \colon [0, n] \to [-m/2, m/2]$ is a polynomial of degree at most D.

To complete the proof we need to show that we can pick v such that $\deg(f_v) \geq 2k+1$. Indeed, since there are $2k+2$ linearly independent vectors in $\Lambda_{n,D} \cap [-m/2, m/2]^{n+1}$, we get $2k+2$ linearly independent polynomials f_v. Consequently, there must exist $v \in \Lambda_{n,D} \cap [-m/2, m/2]^{n+1}$ such that $\deg(f_v) \geq 2k+1$. The polynomial we were looking for is therefore, $f(x) = f_v(x) + m/2$.

This completes the proof of Theorem 1.5. ∎

Remark 6.5. Note that when k is a constant integer, we get from (39) that there is a nonconstant polynomial $f \colon [n] \to [2 \cdot 2^k]$ of degree $\deg(f) \leq n-k$, for a large enough n (specifically, $n \geq c \cdot k^2 \cdot 2^k$ for some global constant c is enough). Combining this with Theorem 1.1 we conclude that

$$n - O\left(\frac{n}{\log\log n}\right) \leq \deg(f) \leq n - k.$$

Also note that Theorem 1.5 implies that for $k=\log(n)-O(1)$ there is a nonconstant polynomial $f\colon [n]\to[n-1]$ of degree $2k\le\deg(f)\le n-k$. Again, combining with Theorem 1.1 we conclude that

$$n - O\left(\frac{n}{\log\log n}\right) \le \deg(f) \le n - \log(n) + O(1).$$

Remark 6.6. Even for $k \le n/10$ we would get from (39) that $m = 2^{O(k)}$. Combining this with Example 5.2 for k in the range $[n/10,n]$, it follows that for any integer $1\le k\le n$ there is a nontrivial polynomial of $\deg(f)\le n-k$ and range bounded by $m=2^{O(k)}$.

We now prove Lemma 6.4.

Proof of Lemma 6.4. By the orthogonality of $b_{D+1},\dots,b_n$

$$\begin{aligned}\det \Lambda_{n,D} &= \det(b_0,\dots,b_n)\\ &= \det(b_0,\dots,b_D)\cdot\prod_{i=D+1}^{n}\|b_i\|_2\\ &= \det(b_0,\dots,b_D)\cdot M^{n-D},\end{aligned}$$

and so it is enough to show that $\det(b_0,\dots,b_D) \le 2^{(n+D+1)(n-D)/2}$. Let $B_{n,D}$ be the $(n+1)\times(D+1)$ matrix with columns $b_0,\dots b_D$. By definition, $\det(b_0,\dots,b_D)=\sqrt{\det(B_{n,D}^T B_{n,D})}$. Using basic rows and columns operations on B, one can show that $\det(B_{n,D}^T B_{n,D})=\det(A_{n,D}^T A_{n,D})\cdot\left(\prod_{i=0}^{D} i!\right)^{-2}$, where $A_{n,D}$ is a $(n+1)\times(D+1)$ matrix with entries $(A_{n,D})_{i,j}=i^j$.[15] The matrix $C_{n,D}\triangleq A_{n,D}^T A_{n,D}$ has the form $(C_{n,D})_{i,j}=\sum_{\ell=0}^{n}\ell^{i+j}$ for $0\le i,j\le D$. In [20], the determinant of $C_{n,D}$, which is a *Vandermondian matrix*, was computed.

Theorem 6.7 ([20] subsection 6.10.4.).

$$\Delta_{n,D} \triangleq \det(C_{n,D}) = \sum_{0\le k_0<k_1<\cdots<k_D\le n} \left(V(k_0,k_1,\dots,k_D)\right)^2,$$

where $V(k_0,k_1,\dots,k_D)$ is the determinant of the usual Vandermonde matrix with parameters $k_0,k_1,\dots,k_D$. That is,

$$V(k_0,k_1,\dots,k_D) = \prod_{0\le i<j\le D}(k_j-k_i).$$

[15] It is easy to prove this by, say, induction on j.

To get a more explicit upper bound on the determinant of $C_{n,D}$, $\Delta_{n,D}$, we prove the following lemma.

Lemma 6.8. *For any integer* $\ell > 0$, $\Delta_{D+\ell,D} \leq \Delta_{D+\ell-1,D} \cdot 4^{D+\ell}$.

We postpone the proof of Lemma 6.8 and continue with the proof. We note that

$$\Delta_{D,D} = \left(\prod_{0 \leq i < j \leq D} (j-i) \right)^2 = \left(\prod_{i=1}^{D} i! \right)^2,$$

and so, applying Lemma 6.8 multiple times, we get

$$\begin{aligned} \Delta_{n,D} &\leq \Delta_{n-1,D} \cdot 4^n \leq \Delta_{n-2,D} \cdot 4^{n+(n-1)} \leq \cdots \\ &\cdots \leq \Delta_{D,D} \cdot 4^{n+(n-1)+\cdots+(D+1)} \\ &= \left(\prod_{i=1}^{D} i! \right)^2 \cdot 2^{(D+n+1)(n-D)}. \end{aligned}$$

Therefore,

$$\begin{aligned} (\det(b_0, \ldots, b_D))^2 &= \det(B_{n,D}^T B_{n,D}) \\ &= \det(C_{n,D}) \cdot \left(\prod_{i=1}^{D} i! \right)^{-2} \\ &= \Delta_{n,D} \cdot \left(\prod_{i=1}^{D} i! \right)^{-2} \leq 2^{(D+n+1)(n-D)}. \end{aligned}$$

Taking the square root of both sides we obtain Lemma 6.4. ∎

We now prove Lemma 6.8.

Proof of Lemma 6.8. We shall map each of the sequences $0 \leq k_0 < k_1 < k_2 < \ldots < k_D \leq D+\ell$ to a sequence $0 \leq k'_0 < k'_1 < k'_2 < \ldots < k'_D \leq D+\ell-1$ as follows:

1. If $k_D \leq D+\ell-1$, then $\forall i \in [0,D]$: $k'_i = k_i$.
2. If $1 \leq k_0$, then $\forall i \in [0,D]$: $k'_i = k_i - 1$.
3. Otherwise, let $0 \leq t < D$ be the first index satisfying $k_t < k_{t+1} - 1$. Note that there is such an index since $k_0 = 0$, $k_D = D+\ell$ and $\ell > 0$. We set

$$k'_i := \begin{cases} k_i & \text{if } i \leq t \\ k_i - 1 & \text{otherwise.} \end{cases}$$

Note that $0 \le k'_0 < k'_1 < k'_2 < \ldots < k'_D \le D+\ell-1$, and that at most $D+2$ sequences $0 \le k_0 < k_1 < k_2 < \ldots < k_D \le D+\ell$ were mapped to the same sequence $0 \le k'_0 < k'_1 < k'_2 < \ldots < k'_D \le D+\ell-1$. We now wish to give an upper bound on

$$\frac{V(k_0, k_1, \ldots, k_D)}{V(k'_0, k'_1, \ldots, k'_D)} = \frac{\prod_{i<j} k_j - k_i}{\prod_{i<j} k'_j - k'_i}. \tag{40}$$

In Cases 1,2 Equation (40) equals 1 since the mapping does not affect the differences between the k_i's. In Case 3 we have

$$\begin{aligned}
(40) &= \frac{\prod_{i<j} k_j - k_i}{\prod_{i<j} k'_j - k'_i} \\
&= \prod_{i<j\le t} \frac{k_j - k_i}{k'_j - k'_i} \cdot \prod_{i\le t<j} \frac{k_j - k_i}{k'_j - k'_i} \cdot \prod_{t<i<j} \frac{k_j - k_i}{k'_j - k'_i} \\
&= \prod_{i<j\le t} \frac{k_j - k_i}{k_j - k_i} \cdot \prod_{i\le t<j} \frac{k_j - k_i}{k_j - 1 - k_i} \cdot \prod_{t<i<j} \frac{k_j - k_i}{(k_j-1)-(k_i-1)} \\
&= \prod_{i=0}^{t} \prod_{j=t+1}^{D} \frac{k_j - k_i}{k_j - 1 - k_i} \\
&= \prod_{i=0}^{t} \frac{\prod_{j=t+1}^{D} k_j - k_i}{\prod_{j=t+1}^{D} k_j - 1 - k_i} \\
&= \prod_{i=0}^{t} \frac{k_D - k_i}{k_{t+1} - 1 - k_i} \cdot \frac{\prod_{j=t+1}^{D-1} k_j - k_i}{\prod_{j=t+2}^{D} k_j - 1 - k_i} \\
&\le \prod_{i=0}^{t} \frac{k_D - k_i}{k_{t+1} - 1 - k_i}.
\end{aligned}$$

Note, that by definition of t it must be the case that $k_0 = 0$, $k_1 = 1, \ldots,$ $k_t = t$ and $k_{t+2} \ge t+2$. Therefore,

$$\prod_{i=0}^{t} (k_{t+1} - 1 - k_i) \ge \prod_{i=1}^{t+1} i,$$

and

$$\prod_{i=0}^{t} (k_D - k_i) \le \prod_{i=0}^{t} (D + \ell - i).$$

It follows that

$$(40) \le \prod_{i=0}^{t} \frac{k_D - k_i}{k_{t+1} - 1 - k_i} \le \frac{\prod_{i=0}^{t} D + \ell - i}{\prod_{i=1}^{t+1} i} = \binom{D+\ell}{t+1}$$

$$\leq \binom{D+\ell}{(D+\ell)/2} < \frac{2^{D+\ell}}{\sqrt{1.5\cdot(D+\ell)}},$$

where the last inequality follows from Stirling's approximation for a large enough D. Hence

$$\begin{aligned}\Delta_{D+\ell,D} &= \sum_{0\leq k_0<k_1<\cdots<k_D\leq D+\ell} (V(k_0,k_1,\ldots,k_D))^2 \\ &\leq \sum_{0\leq k_0<\ldots<k_D\leq D+\ell} \left(\frac{2^{D+\ell}}{\sqrt{1.5\cdot(D+\ell)}}\right)^2 \cdot V(k'_0,k'_1,\ldots,k'_D)^2 \\ &= \frac{4^{D+\ell}}{1.5\cdot(D+\ell)} \cdot \sum_{0\leq k_0<\ldots<k_D\leq D+\ell} V(k'_0,\ldots,k'_D)^2 \\ &\leq^{(*)} \frac{4^{D+\ell}\cdot(D+2)}{1.5\cdot(D+\ell)} \cdot \sum_{0\leq k'_0<\ldots<k'_D\leq D+\ell-1} V(k'_0,\ldots,k'_D)^2 \\ &\leq 4^{D+\ell}\cdot \Delta_{D+\ell-1,D},\end{aligned}$$

where inequality $(*)$ holds as at most $D+2$ sequences $0\leq k_0<k_1<k_2<\ldots< k_D\leq D+\ell$ were mapped to the same sequence $0\leq k'_0<k'_1<k'_2<\ldots<k'_D\leq D+\ell-1$, as mentioned above. This completes the proof of the lemma. ∎

7. Back to the Boolean case

In this section we consider the Boolean case. Specifically, let $m=1$ and $n=p^2-1$ for some prime p. We prove that in this case the degree must be at least $n-\sqrt{n}$. For completeness, we also give a proof for the case $n=p-1$, that was previously proved in [6].

Proof of Theorem 1.6. Let f be as in the statement of the theorem and assume that $\deg(f)<p^2-p$. By Lemma 2.4 we get that for all $r\in[0,p-1]$

$$\sum_{k=0}^{p^2-p}(-1)^k\binom{p^2-p}{k}f(k+r)=0. \tag{41}$$

Since $p^2-p=(p-1)\cdot p+0$, it follows, by Lucas' theorem, that if $k=k_1p+k_0$, is the base p representation of k, then $\binom{p^2-p}{k}\equiv_p 0$ when $k_0\neq 0$ and $\binom{p^2-p}{k}\equiv_p(-1)^{k_1}$ when $k_0=0$. Therefore, (41) is equivalent to

$$0=\sum_{k=0}^{p^2-p}(-1)^k\binom{p^2-p}{k}f(k+r)\equiv_p\sum_{k_1=0}^{p-1}f(k_1p+r).$$

Note that the RHS contains exactly p summands. As they are all in $\{0,1\}$ they must all be equal in order for their sum to be 0 modulo p. We thus get that for every $r \in [0,p-1]$, $f(r) = f(p+r) = \ldots = f((p-1)p+r)$. In other words, if we set $g(x) \triangleq f(x+p) - f(x)$ then $g(x)=0$ for $x \in [0,p^2-p-1]$.

If g is identically zero, then Lemma 2.9 implies that $\deg(f)=0$, i.e., that f is constant, as claimed. Otherwise, since g has p^2-p zeroes, it follows that $\deg(g) \geq p^2-p$. This is a contradiction as $\deg(f) \geq \deg(g)$ (in fact, $\deg(f)=\deg(g)+1$). ∎

For completeness we also prove the following result of [6].

Theorem 7.1 ([6]). *Let p be a prime number, $n=p–1$ and $f\colon [0,n] \to \{0,1\}$ be nonconstant. Then $\deg(f)=p-1=n$.*

Proof. Assume that $\deg(f)<n$. As in the proof of Theorem 1.6, we apply Lemma 2.4 and Lucas' theorem to obtain

$$0 = \sum_{k=0}^{p-1} (-1)^k \binom{p-1}{k} f(k+r) \equiv_p \sum_{k=0}^{p-1} f(k).$$

Again, it must be the case that $f(0)=f(1)=\ldots=f(p–1)$, i.e., f is constant. ∎

8. Discussion

We proved that it is 'hard' for polynomials to 'compress' the interval $[0,n]$. Namely, that any such nonconstant polynomial to a strict subset of $[0,n]$ must have degree $n-o(n)$. We also proved that if we allow $m = \frac{1}{d!} \cdot \left(\frac{n-d}{2e}\right)^d$ then f can of course have degree $< d$, but all other polynomials mapping $[0,n]$ to $[0,m]$ must have degree $\geq n/3 - o(n)$. We are not able to prove however that our results are tight. In particular we believe that they can be improved both for the case $m<n$ and for the case of large m. We note that the following question, posed by von zur Gathen and Roche, is still open: "*... for each m there is a constant C_m such that* $\deg(f) \geq n - C_m$". Furthermore, when $m=1$ they raise the possibility that $C_1=3$. As an intermediate goal it will be interesting to manage to break the $n–\Gamma(n)$ upper bound. Specifically, show that when $f \in \mathcal{F}_1(n)$ is nonconstant, $\deg(f) \geq n - \sqrt{n}$. It seems that new techniques are required in order to prove this claim as all current proofs are based on modular calculations and we cannot guarantee the existence of a prime p in the range $[n-\sqrt{n}, n]$. For the special case that $n=p^2-1$ we managed to obtain such a result, and of course when $n=p-1$ a stronger result is known, but the general case is still open.

Another intriguing question is to understand what is the minimal range that a polynomial mapping integers to integers of degree exactly d can have. We note that in Example 5.2 the degree is d and the range is (roughly) of size $\frac{1}{d!} \cdot \left(\frac{n}{2}\right)^d$. Theorem 1.3 asserts that if the degree is d then the range must be larger than (roughly) $\frac{1}{d!} \cdot \left(\frac{n-d}{2e}\right)^d$ (Theorem 5.5 actually improves it to $\frac{1}{d!} \cdot \left(\frac{n}{4}\right)^d$ for $d \leq \sqrt{n}/2$). It is an interesting question to understand the 'correct' bound.

Finally, we think that it will be interesting to find examples that are significantly better than those obtained in Theorem 1.5 and Example 5.2.

Acknowledgements. Gil Cohen would like to thank Orit Ashtamker-Cohen for lots of support. He also thanks Malte Beecken (Bonn), Johannes Mittmann (Bonn) and Pablo Azar (MIT) for helpful discussions on the problem. Avishay Tal would like to thank Benjamin Eliot Klein for lots of support and helpful discussion on the subject - especially in proving Theorem 2.3 and Lemma 6.4. Avishay also thanks Nathan Keller for helpful discussions and for pointing out the possible use of Chebyshev polynomials. The authors wish to thank Joachim von zur Gathen for interesting discussions on the problems studied here.

References

[1] R. C. Baker, G. Harman and J. Pintz: The difference between consecutive primes, II, *Proceedings of the London Mathematical Society* **83** (2001), 532–562.

[2] R. Beigel: The polynomial method in circuit complexity, in: *Structure in Complexity Theory Conference*, 82–95, 1993.

[3] H. Buhrman and R. de Wolf: Complexity measures and decision tree complexity: a survey, *Theor. Comput. Sci.* **288** (2002), 21–43.

[4] H. Cramér: On the order of magnitude of the difference between consecutive prime numbers, *Acta Arithmetica* **2** (1936), 23–46.

[5] W. Feller: *An Introduction to Probability Theory and Its Applications*, volume 1, Wiley, New York, 3rd edition, 1968.

[6] J. von zur Gathen and J. R. Roche: Polynomials with two values, *Combinatorica* **17** (1997), 345–362.

[7] O. Goldreich: *Computational Complexity: A Conceptual Perspective*, Cambridge University Press, 2008.

[8] P. Gopalan: *Computing with Polynomials over Composites*, PhD thesis, Georgia Institute of Technology, August 2006.

[9] P. M. Gruber and C. G. Lekkerkerker: *Geometry of Numbers*, North-Holland, 1987.

[10] D. E. Knuth: *The Art of Computer Programming, Volume III: Sorting and Searching*, Addison-Wesley, 1973.

[11] M. N. Kolountzakis, R. J. Lipton, E. Markakis, A. Mehta and N. K. Vishnoi: On the Fourier spectrum of symmetric Boolean functions, *Combinatorica* **29** (2009), 363–387.
[12] N. Linial, Y. Mansour and N. Nisan: Constant depth circuits, Fourier transform and learnability, *J. ACM* **40** (1993), 607–620.
[13] E. Lucas: Théorie des fonctions numériques simplement périodiques, *American Journal of Mathematics* **1** (1878), 184–196.
[14] J. C. Mason and D. C. Handscomb: *Chebyshev Polynomials*, Chapman & Hall/CRC, Boca Raton, FL, 2003.
[15] E. Mossel, R. O'Donnell and R. A. Servedio: Learning functions of k relevant variables, *J. Comput. Syst. Sci.* **69** (2004), 421–434.
[16] A. A. Razborov: Lower bounds on the size of bounded depth circuits over a complete basis with logical addition, *Math. Notes* **41** (1987), 333–338.
[17] H. Robbins: A Remark of Stirling's Formula, *American Mathematical Monthly* **62** (1955), 26–29.
[18] A. Shpilka and A. Tal: On the minimal Fourier degree of symmetric Boolean functions, *Combinatorica* **34** (2014), 359–377.
[19] R. Smolensky: Algebraic methods in the theory of lower bounds for Boolean circuit complexity, in: *Proceedings of the 19th Annual STOC*, pages 77–82, 1987.
[20] R. Vein and P. Dale: *Determinants and Their Applications in Mathematical Physics*, Springer, 1999.

Gil Cohen, Avishay Tal

Department of Computer Science
and Applied Mathematics
The Weizmann Institute of Science
Rehovot, Israel
{gil.cohen,avishay.tal}@weizmann.ac.il

Amir Shpilka

Faculty of Computer Science
Technion-Israel Institute of Technology
Haifa, Israel
and
Microsoft Research
Cambridge MA, USA
shpilka@cs.technion.ac.il