INTERSECTIVE POLYNOMIALS AND THE PRIMES

THÁI HOÀNG LÊ

Abstract. Intersective polynomials are polynomials in $\mathbf{Z}[x]$ having roots every modulus. For example, $P_1(n) = n^2$ and $P_2(n) = n^2 - 1$ are intersective polynomials, but $P_3(n) = n^2 + 1$ is not. The purpose of this note is to deduce, using results of Green-Tao [8] and Lucier [16], that for any intersective polynomial h, inside any subset of positive relative density of the primes, we can find distinct primes p_1, p_2 such that $p_1 - p_2 = h(n)$ for some integer n. Such a conclusion also holds in the Chen primes (where by a Chen prime we mean a prime number p such that $p + 2$ is the product of at most 2 primes).

1. Introduction

In the late 1970s, Sárközy and Furstenberg independently proved the following, which had previously been conjectured by Lovasz:

Theorem 1 (Sárközy [18], Furstenberg [9], [10]). *If A is a subset of positive upper density of $\mathbf{Z}$, then there are two distinct elements of A whose difference is a perfect square.*

While Furstenberg used ergodic theory, Sárközy actually proved the following finitary, quantitative form:

Theorem 2 (Sárközy). *Let $\delta > 0$. Then provided N is sufficiently large depending on δ, $N > N_0(\delta)$, any subset A of $\{1, \dots, N\}$ of size δN contains two distinct elements $a, a' \in A$ such that $a - a'$ is a perfect square.*

We have the same conclusion if the set of the squares is replaced by $\{p + 1 : p \text{ prime}\}$ or $\{p - 1 : p \text{ prime}\}$. More generally, we say that a set $H \subset \mathbf{Z}^+$ is intersective if $H \cap (A - A) \neq \emptyset$ for any set A of positive upper density. We say that a polynomial $h \in \mathbf{Z}[x]$ is intersective if the set $\{h(n) : n \in \mathbf{Z}\} \cap (0, \infty)$ is intersective. Thus Sárközy's theorem says that the polynomial $h(n) = n^2$ is intersective.

Kamae and Mendès France [11] proved a criterion about intersective sets. This gives a necessary and sufficient condition for a polynomial to be intersective:

Theorem 3 (Kamae-Mendès France). *A polynomial $h \in \mathbf{Z}[x]$ is intersective if and only if for every $d > 0$, there exists n such that $P(n) \equiv 0 \pmod{d}$.*

For example, the polynomials x^2 and $x^2 - 1$ are intersective, while $x^2 + 1$ is not (think of obstruction modulo 3). A polynomial having an integer root is certainly intersective, but there are intersective polynomials which do not have an integer root, e.g. the polynomials $(x^3 - 19)(x^2 + x + 1)$, or $(x^2 - 2)(x^2 - 3)(x^2 - 6)$. Berend and Bilu gave in [1] a procedure to determine whether or not a given polynomial is intersective.

If h is an intersective polynomial, denote by $D(h,N)$ the maximal size of a subset A of $\{1,\dots,N\}$ such that we cannot find distinct elements $a,a' \in A$ such that $a-a'=h(n)$ for some integer n. Thus necessarily $D(h,N)=o(N)$. It should be mentioned that like Furstenberg's method, Kamae and Mendès France's is qualitative, i.e., does not give any bound on $D(h,N)$. In the case where $h(n)=n^2$, and more generally $h(n)=n^k$, the best bound is due to Pintz-Steiger-Szemerédi [17] and Balog-Pelikan-Pintz-Szemerédi [3]. They proved that

$$D(n^k,N) \ll_k N(\log N)^{-(1/4)\log\log\log\log N}$$

for N sufficiently large depending on k. Note that this density already includes the primes. For general intersective polynomials, such a quantitative bound was obtained recently by Lucier [16]. He proved that, for any intersective polynomial h of degree k,

$$D(h,N) \ll_h N\frac{(\log\log N)^{\mu/(k-1)}}{(\log N)^{1/(k-1)}}$$

for N sufficiently large depending on h, where $\mu = \begin{cases} 3, & \text{if } k=2; \\ 2, & \text{if } k\geq 3. \end{cases}$

This density is weaker and does not include the primes. It may well be the case that the correct density includes the primes. However, we don't seek to improve upon Lucier's result, but rather use it, coupled with a "transference principle" to deduce a corresponding result for the primes.

Let $\mathcal{P}$ be a subset of N. For any subset $\mathcal{A}\subset\mathcal{P}$, define the upper relative density of $\mathcal{A}$ with respect to $\mathcal{P}$ by $\overline{d}_{\mathcal{P}}(\mathcal{A}) = \lim_{N\to\infty}\frac{\sharp\{n\leq N: n\in\mathcal{A}\}}{\sharp\{n\leq N: n\in\mathcal{P}\}}$. We will obtain the following:

Theorem 4. *For any intersective polynomial h, for any subset $\mathcal{A}$ of positive upper relative density of the primes, there exist distinct elements p_1,p_2 of $\mathcal{A}$ such that $p_1-p_2=h(n)$.*

Remarks 1.1. If $h(0)=0$, then this is a very special case of the result of Tao-Ziegler [20], which says that configurations $a+P_1(d),\dots,a+P_k(d), d\neq 0$ exist in dense subsets of the primes, where $P_i\in\mathbf{Z}[x], P_i(0)=0$. Their starting point is a uniform version of the Bergelson-Leibman theorem, which says that such configurations exist in dense subsets of the integers. Tao-Ziegler's proof of the uniform version uses a lifting to a multidimensional version of the Bergelson-Leibman theorem and relies on the very fact that each $P_i(0)=0$. Therefore, it is not applicable to general intersective polynomials.

Following Green and Tao, let us call a prime p a Chen prime if $p+2$ is either a prime or a product p_1p_2 of primes with $p_1,p_2>p^{3/11}$. The following result is due to Chen [4]:

Theorem 5 (Chen). *Let N be a large integer. The the number of Chen primes in the interval $[1,N]$ is at least $c_1N/\log^2 N$ for some absolute constant $c_1>0$.*

For a proof of Chen's theorem, see [13]. Using this result as a "black box" we can show that the same conclusion holds for dense subsets of the Chen primes:

Theorem 6. *For any intersective polynomial h, for any subset $\mathcal{A}$ of positive upper relative density of the Chen primes, there exist distinct elements p_1,p_2 of $\mathcal{A}$ such that $p_1-p_2=h(n)$.*

The idea of transferring results on dense subsets of the integers to the primes originates with Green [6], in which he proved an analog of Roth's theorem for the primes. Later on, other transference principles have been devised by Green and Tao in [8] in which they proved the analog of Roth's theorem in the Chen primes, and in [7] in which they proved that the primes contains arbitrarily long arithmetic progressions. These machineries have been used in a number of settings, such as random sets ([19], [12]) or the ring of polynomials over a finite field ([14]). We opt for the transference principle in [8] since it is relatively simpler and more general than that in [6]. In a similar spirit, Li and Pan [15] proved that if Q is a polynomial in $\mathbf{Z}[x]$ such that $Q(1) = 0$, then inside any dense subset of the primes, we can find two distinct elements whose difference is of the form $Q(p)$ where p is a prime number. It would be interesting to determine the class of all the polynomials Q such that the same conclusion holds (other than those vanishing at 1).

Acknowlegdments. I would like to thank my advisor Terence Tao for helpful discussions during the preparation of this paper. I would like also to thank Craig Spencer for pointing me to Lucier's paper [16] and for helpful comments on this paper. Part of this research was done when I was visiting the Mathematical Sciences Research Institute, Berkeley, and I am grateful to their hospitality.

2. Notation and Preliminaries

For two quantities A, B, we write $A = O(B)$, or $A \ll B$, or $B \gg A$ if there is an absolute positive constant C such that $|A| \leq CB$. If A and B are functions of the same variable x, we write $A = o_{x\to\infty}(B)$ if A/B tends to 0 as x tends to infinity. If the constant C, (respectively, the rate of convergence of A/B) depends on a parameter, e.g. m, then we write $A = O_m(B)$ (respectively, $A = o_m(B)$). Quantities denoted by the C, c will stand for constants, which may change from line to line. We denote by $\mathbf{Z}_N$ the cyclic group on N elements. This is not to be confused with the ring of p-adic integers, which we also denote by $\mathbf{Z}_p$, since we will make use of the latter very briefly (in the introduction of auxiliary polynomials).

2.1. Fourier analysis on $\mathbf{Z}_N$.

We will work primarily in a group $\mathbf{Z}_N$ where N is a large number. For a function $f : \mathbf{Z}_N \to \mathbf{C}$ let us define its Fourier transform by $\widehat{f}(\xi) = \mathbf{E}_{x\in\mathbf{Z}_N} f(x)e_N(x\xi)$, where $e_N(t) = e^{\frac{2\pi it}{N}}$, and $\mathbf{E}$ is the expectation. If $f, g : \mathbf{Z}_N \to \mathbf{C}$ are two functions, then $f * g$, the convolution of f and g, is defined by $f * g(x) = \mathbf{E}_{y\in\mathbf{Z}_N} f(y)g(x - y)$. We also define the l^p-norm of f to be $\|f\|_p = \left(\sum_{\xi\in\mathbf{Z}_N} |f(\xi)|^p\right)^{1/p}$. We will often use a subset of $\mathbf{Z}_N$ to denote its characteristic function.

We recall the basic properties of the Fourier transform:

- (Fourier inversion formula) $f(x) = \sum_{\xi\in\mathbf{Z}_N} \widehat{f}(x)e_N(-x\xi)$
- (Plancherel) $\sum_{\xi\in\mathbf{Z}_N} \widehat{f}(\xi)\overline{\widehat{g}(\xi)} = \mathbf{E}_{x\in\mathbf{Z}_N} f(x)\overline{g(x)}$
- (Parseval) $\|\widehat{f}\|_2^2 = \sum_{\xi\in\mathbf{Z}_N} |\widehat{f}(\xi)|^2 = \mathbf{E}_{x\in\mathbf{Z}_N} |f(x)|^2$
- (Fourier transform of a convolution) $\widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$ for every $\xi \in \mathbf{Z}_N$

2.2. Intersective polynomials.

Let $h(x) = a_k x^k + \cdots + a_0$ be a fixed intersective polynomial of degree $k \geq 2$ throughout the paper. By a change of variables if need be, we may assume that h and h' are positive and increasing for $x \geq 0$.

If $f(x) = b_k x^k + \cdots + b_0$, let us denote by $b(f) = b_k$ and $B(f) = \frac{2}{|b_k|}(|b_{k-1}| + \cdots + |b_0|)$. Then if $b(f) > 0$, we have $B(f') \leq B(f)$ and

$$\frac{1}{2}b(f)x^k \leq f(x) \leq \frac{3}{2}b(f)x^k \tag{1}$$

for $x \geq B(f)$ ([16, Lemma 3]).

If f has integer coefficients, let us denote by $c(f) = \gcd(b_k, \ldots, b_1)$, the content of f.

Suppose $f = a(x-\eta_1)^{e_1} \cdots (x-\eta_r)^{e_r}$ in some splitting field. Let us denote by $\Delta(f) = a^{2k-2}\prod_{i\neq j}(\eta_i - \eta_j)^{e_i e_j}$, the semidiscriminant of f. The semidiscriminant was first introduced by Chudnovsky [5]. When f is separable then the semidiscriminant is simly the discriminant. It can be shown that $\Delta(f)$ is always a non-zero integer when $f \in \mathbf{Z}[x]$.

In order for the transference principle to work, we need not only one solution to $a - a' = h(n)$, but "many" (i.e., of the "right" order) of them. This is already established by Lucier. Another issue is that we will not be working directly with the primes, but rather affine images of primes (in congruences classes modulo W, where W is a product of small primes meant to absorb obstruction at these primes). This technique is called the "W-trick" and is quite common in situations in arithmetic combinatorics where we want to transfer results on dense subsets of the integers to the primes [6], [8], [7], [20].

Thus instead of a single polynomial h, we will work with a family of polynomials h_W parametrized by W, which are compositions of h with affine maps. Our bounds need to be independent of W. As mentioned earlier, Tao-Ziegler's proof of the uniform version of the Bergelson-Leibman theorem does not apply to general intersective polynomials. Fortunately, the auxiliary polynomials introduced by Lucier serve well our purposes.

Note that the condition that h has roots every modulo is equivalent to saying that h has a root in $\mathbf{Z}_p$ for every prime p, where $\mathbf{Z}_p$ is the ring of p-adic integers. For each p let us fix a root $z_p \in \mathbf{Z}_p$ of h. If m is the multiplicity of z_p as a root of h then we define $\lambda(p) = p^m$. We can then extend λ to a completely multiplicative arithmetic function on $\mathbf{N}$. It is easy to see that for every d, $d|\lambda(d)|d^k$.

Suppose $d = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$ is the prime factorization of d. By the Chinese remainder theorem, let r_d be the unique integer satisfying $-d < r_d \leq 0$ and $r_d \equiv z_p \pmod{p_i^{\alpha_i}\mathbf{Z}_{p_i}}$ for every $i = 1, \ldots, s$.

For any positive integer d we define the polynomial $h_d(x) = \frac{h(r_d + dx)}{\lambda(d)}$. The properties of h_d, proved in [16], are summarized in the following lemma:

Lemma 1. (1) *For every d, h_d is a polynomial with integer coefficients and degree k. Furthermore, h_d is also intersective.*
(2) *The polynomials $h(d), h'(d), h''(d)$ are positive and increasing for $x \geq 1$.*
(3) *For every $d, q > 0$ then $(h_d)_q = h_{dq}$.*
(4) *$b(h_d) \leq b(h_d) \leq d^{k-1}b(h)$.*
(5) *$B(h_d) \leq 2^{k-1}k(B(h) + 2)$.*
(6) *$c(h_d) \leq |\Delta(h)|^{\frac{k-1}{2}}c(h)$, where $\Delta(h)$ is the semidiscriminant of h.*

Remark 2.3. The last property is by far the most important, since our bounds on exponential sums involving h_d will depend on $c(h_d)$. The last two properties ensure that $B(h_d)$ and $c(h_d)$

can be bounded uniformly, no matter what d is. The only quantity that can grow is $b(h_d)$. We will see that this quantity is also within control if we keep d smaller than a small power of N.

3. A uniform version of Lucier's theorem

Let us first recall Lucier's main result in [16]. Let $\delta > 0$ and A be a subset of $\{1, \dots, N\}$ such that $|A| = \delta N$. For every n let $r(h, n, A)$ be the number of couples (a, a') of elements in A such that $a - a' = h(n)$. Let $R(A, h) = \sum_{n \geq 0} h'(n) r(h, n, A)$.

Theorem 7 (Theorem 5, [16])**.** *There is a constant $C(h, \delta)$ depending on h and δ alone such that whenever N is sufficiently large in terms of h and δ, the following estimate holds:*

$$R(A, h) \geq c(h, \delta)|A|^2$$

Actually Lucier obtained the following estimate for $c(h, \delta)$:

$$c(h, \delta) = \exp\left(-c_1 \delta^{-(k-1)} \log^{\mu}\left(\frac{2}{\delta}\right)\right)$$

which is valid for $\delta \geq c_2 \frac{(\log\log N)^{\mu/(k-1)}}{\log N^{1/(k-1)}}$ where c_1, c_2 are constants depending on h alone, and $\mu = \begin{cases} 3, & \text{if } k = 2; \\ 2, & \text{if } k \geq 3. \end{cases}$ As mentioned earlier, we need to work with the family (h_W) rather than with h alone. The following gives a uniform version of Theorem 7:

Theorem 8. *There is a constant $\kappa_1 = \kappa_1(k)$ depending on k alone, and a constant $C(h, \delta)$ depending on h and δ alone such that whenever N is sufficiently large in terms of h, the following estimate holds:*

$$R(A, h_W) \geq C(h, \delta)|A|^2$$

for every $W < N^{\kappa_1}$, where the constant $C(h, \delta)$ is the same as in Theorem 7 (but the range of validity of N may be slightly different).

Proof. Only a minor modification of Lucier's proof is needed. Lucier used a density increment argument based on the following:

Lemma 2 (Lemma 31, [16])**.** *Let $\varrho = \varrho(k)$ be defined by*

$$\varrho = \begin{cases} 1/4, & \text{if } k = 2; \\ 1/(8k^2(\log k + 1.5\log\log k + 4.2)), & \text{if } k \geq 3. \end{cases}$$

Define the function

$$\theta(x) = \begin{cases} \frac{x}{2\log(2x^{-1})}, & \text{if } k = 2; \\ x^{k-1}, & \text{if } k \geq 3. \end{cases}$$

Let N be large in terms of h, and assume that

$$d \leq N^{\rho/4k^2}$$

Let A be a subset of $\{1, \dots, N\}$ with size δN such that

$$\delta \geq N^{-\varrho/2k}$$

If $R(h_d, A) \leq \frac{1}{64}|A|^2$, then there exist positive integers d' and N', and a set $A' \subset \{1, \dots, N'\}$ such that the following holds:

- $W(h_{d'}, A') \leq W(h_d, A)$,
- $\delta' \geq \delta(1 + C_1\theta(\delta))$,
- $C_2\delta^{2k^2}N \leq N' \leq N$,
- $d \leq d' \leq C_3\delta^{-k}d$.

where C_1, C_2, C_3 are positive constants that depend only on h.

Following Lucier, suppose that

$$\delta \geq C\frac{(\log\log N)^{\mu/(k-1)}}{(\log N)^{1/(k-1)}}$$

for C a constant chosen later, that depends on h alone. Let

$$Z = [8C_1^{-1}\delta^{-(k-1)}(\log 2\delta^{-1})^{\mu-1}]$$

Suppose, for a contradiction, that that $R(h, A) \leq \frac{1}{64}(C_2^2\delta^{4k^2})^Z|A|^2$. Lucier constructed a sequence of quadruples $\{(N_i, A_i, \delta_i, d_i)\}_{i=0}^{Z}$, where N_i, d_i are positive integers, $A_i \subset \{1, \ldots, N\}$, $\delta_i = |A_i|/N_i$, satisfying the properties:

- $(N_0, A_0, \delta_0, d_0) = (N, A, \delta, 1)$
- $R(h_{d_i}, A_i) \leq \frac{1}{64}(C_2^2\delta^{4k^2})^{Z-i}|A_i|^2$
- $\delta_i \geq \delta_{i-1}(1 + C_1\theta(\delta_{i-1}))$
- $C_2\delta_{i-1}^{2k^2}N_{i-1} \leq N_i \leq N_{i-1}$
- $d_{i-1} \leq d_i \leq C_3\delta_i^{-k}d_{i-1}$

where C_1, C_2, C_3 are constants as in Lemma 2 above. We can perform the iteration at step l as long as the conditions of Lemma 2 is valid:

(1) N_l is large in terms of h. Indeed, if we choose C large enough we can ensure that $N_l \geq N^{1/2}$ for all $0 \leq l \leq Z$.
(2) $d_l \leq N_l^{\rho/4k^2}$. Indeed, we have the inequality $\log d_l \ll_h C^{-1}\log N + \log d_0$, so if C is large enough in terms of h this is satisfied.
(3) $\delta_l \geq N_l^{-\varrho/2k}$. This too is ensured if C is large enough.

A calculation shows that we will end up with $\delta_Z > 1$, a contradiction.

Now if the initial values are $(N_0, A_0, \delta_0, d_0) = (N, A, \delta, W)$ (instead of $(N, A, \delta, 1)$) then the same iteration goes through. The only thing that needs to be checked is the condition (2) above. But we can ensure this by choosing C sufficiently large depending on h alone, as long as we keep W smaller than N^{κ_1}, for $\kappa_1 = \frac{\varrho}{16k^2}$, say. □

4. A transference principle for intersective polynomials

4.1. An exponential sum estimate.

Lemma 3. *Let $f \in \mathbf{Z}[x]$ be a polynomial of degree k, and assume that f is positive and increasing for $x \leq 0$. Then there is an integer $s_0(k)$ depending on k alone, such that whenever*

$s \geq s_0$, we have

$$\int_{\mathbf{T}} \left| \sum_{n=1}^{N} f'(n) e(\alpha f(n)) \right|^{2s} \ll_s c(f) f(N)^{2s-1}$$

for $N \geq B(f)$.

Remarks 4.2. This is essentially [15, Lemma 2.6], where Li and Pan showed that we can take $s_0 = k2^{k+1}$. This result is standard in the context of Waring's problem, so we will skip the proof. It may be possible to improve upon the value of s_0 using Vinogradov's method, but this is not important since all we need is the existence of such a number s_0. The condition $N \geq B(f)$ is needed in order to guarantee that $\sum_{n=1}^{N} f'(n) \ll f(N)$.

Let us denote $S_f(x) = S_{N,f}(x) = \begin{cases} f'(n), & \text{if } 0 < x < N/2 \text{ and } x = f(n) \text{ for some } n \in \mathbf{Z}; \\ 0, & \text{otherwise.} \end{cases}$ and consider S_f as a function on $\mathbf{Z}_N$.

Corollary 1. *For $s \geq s_0(k)$, and for $N \gg b(f)B(f)^k$, we have*

$$\left\| \widehat{S_f} \right\|_{2s} \ll_s (c(f))^{1/2s}$$

Proof. Let M be the largest integer such that $f(M) < \frac{N}{2}$. In view of (1), if $b(f)B(f)^k \ll N$ then $M \geq B(f)$. We can therefore apply Lemma 3 and have:

$$\begin{aligned} \left\| \widehat{S_{N,f}} \right\|_{2s}^{2s} &= \frac{1}{N^{2s}} \sum_{\xi \in \mathbf{Z}_N} \left| \sum_{x \in \mathbf{Z}_N} S_f(x) e_N(\xi x) \right|^{2s} \\ &= \frac{1}{N^{2s-1}} \sum_{\substack{n_1,\dots,n_s,m_1,\dots,m_s \in \{1,\dots,M\} \\ f(n_1)+\cdots+f(n_s)=f(n_1)+\cdots+f(n_s)}} f'(n_1)\cdots f'(n_s) f'(m_1)\cdots f'(m_s) \\ &= \frac{1}{N^{2s-1}} \int_{\mathbf{T}} \left| \sum_{n=1}^{M} f'(n) e(\alpha f(n)) \right|^{2s} \\ &\ll_s \frac{1}{N^{2s-1}} c(f) f(M)^{2s-1} \\ &\ll_s c(f) \end{aligned}$$

□

From this it immediately follows that

Corollary 2. *There is a constant $\kappa_2 = \kappa_2(k)$ such that for $s \geq s_0$, and for N sufficiently large depending on h, we have*

$$\left\| \widehat{S_{N,h_W}} \right\|_{2s} \ll_{s,h} 1$$

for every $W < N^{\kappa_2}$.

Proof. Lemma 1 tells us that $c(h_W)$ is uniformly bounded in terms of h. Thus we need $b(h_W)B(h_W)^k \ll N$ for all $W \leq N^{\kappa_2}$. But this also follows from Lemma 1. Actually we may take $\kappa_2(k) = 1/k$. □

4.3. **A transference principle.** Let us reformulate Theorem 7 under the following form:

Proposition 1. *There is a constant a constant $c(h,\delta)$ such that the following holds. If $f : \mathbf{Z}_N \to [0,\infty)$ is a function such that $\mathbf{E}_{\mathbf{Z}_N} f \geq \delta$, then*

$$\sum_{a\in\mathbf{Z}_N}\sum_{d\in\mathbf{Z}_N} f(a)f(a+n)S_h(d) \geq c(h,\delta)N^2$$

for N sufficiently large depending on h and δ.

We are now in a position to state the following transference principle for intersective polynomials:

Proposition 2. *Let η,δ,M,q be positive parameters such that $2 < q < \frac{4s_0}{2s_0-1}$, where $s_0 = s_0(k)$ as in Lemma 3. Suppose f,ν are function $\mathbf{Z}_N \to \mathbf{R}$ satisfying the following conditions:*

1. $0 \leq f \leq \nu$
2. $\mathbf{E}_{n\in\mathbf{Z}_N} f(n) \geq \delta$
3. *ν satisfies the pseudorandom condition $|\widehat{\nu}(\xi) - 1_{\xi=0}| \leq \eta$ for all $\xi \in \mathbf{Z}_N$.*
4. $\|\widehat{f}\|_q \leq M$.

Then for N large enough depending on h and δ, we have

$$\sum_{a\in\mathbf{Z}_N}\sum_{d\in\mathbf{Z}_N} f(a)f(a+d)S_h(d) \geq \left(\frac{1}{2}c(h,\delta) - O_{M,q,\delta}(\eta)\right)N^2$$

We proceed as in [8, Proposition 5.1]. Let us recall in the form of a lemma the following decomposition result contained in the proof of [8, Proposition 5.1]:

Lemma 4. *Suppose $0 < \epsilon < 1$. Let*

$$\Omega = \{a \in \mathbf{Z}_N : |\widehat{f}(a)| \geq \epsilon\}$$

and

$$B = B(\Omega,\epsilon) = \{m \in \mathbf{Z}_N : |1 - e_N(am)| \geq \epsilon \text{ for all } a \in \Omega\}$$

Let

$$f_1(n) = \mathbf{E}_{m_1,m_2\in B} f(n + m_1 - m_2)$$

and $f_2 = f - f_1$ is the uniform part. Then f_1 and f_2 satisfy the following properties:

1. $0 \leq f_1 \leq 1 + (N/|B|)\eta$,
2. $\mathbf{E}_{\mathbf{Z}_N}(f_1) = \mathbf{E}_{\mathbf{Z}_N}(f)$
3. $\|\widehat{f_2}(\xi)\|_\infty \leq 3(1+\eta)\epsilon$,
4. *For every $\xi \in \mathbf{Z}_N$, we have $|\widehat{f_1}(\xi)|, |\widehat{f_2}(\xi)| \leq |\widehat{f}(\xi)|$.*

Proof of Proposition 2. We write

$$\begin{aligned}\sum_{a,d\in\mathbf{Z}_N} f(a)f(a+d)S_h(d) &= \sum_{a,d\in\mathbf{Z}_N} f_1(a)f_1(a+d)S_h(d) + \sum_{a,d\in\mathbf{Z}_N} f_1(a)f_2(a+d)S_h(d) \\ &+ \sum_{a,d\in\mathbf{Z}_N} f_2(a)f_1(a+d)S_h(d) + \sum_{a,d\in\mathbf{Z}_N} f_2(a)f_2(a+d)S_h(d)\end{aligned}$$

Note that since $\|\widehat{f}(\xi)\|_q \leq M$, we have $|\Omega| \leq (M/\epsilon)^q$. Also, $|B| \geq (\epsilon/C)^{|\Omega|}$ for some absolute constant C. Thus we have $0 \leq f_1 \leq 1 + (C/\epsilon)^{(M/\epsilon)^q}\eta = 1 + O_{M,\epsilon,q}(\eta)$. Applying Proposition 1 to the function f_1 (possibly modified by $O_{M,q,\epsilon}(\eta)$), we have

$$\sum_{a\in\mathbf{Z}_N}\sum_{d\in\mathbf{Z}_N} f_1(a)f_1(a+d)S_h(d) \geq (c(h,\delta) - O_{M,q,\epsilon}(\eta))N^2$$

Our goal is to show that the three last terms are small in absolute value. We consider the second term; the other two terms are treated similarly. We have

$$\begin{aligned}\left|\sum_{a,d\in\mathbf{Z}_N} f_1(a)f_2(a+d)S_h(d)\right| &= N\left|\sum_{a\in\mathbf{Z}_N} f_1(a)f_2 * S_h(a)\right| \\ &= N^2\left|\sum_{\xi\in\mathbf{Z}_N} \overline{\widehat{f_1}(\xi)}\widehat{f_2}(\xi)\widehat{S_h}(\xi)\right| \\ &\leq N^2\sum_{\xi\in\mathbf{Z}_N} |\widehat{f_1}(\xi)||\widehat{f_2}(\xi)||\widehat{S_h}(\xi)|\end{aligned}$$

By Hölder's inequality,

$$\sum_{\xi\in\mathbf{Z}_N} |\widehat{f_1}(\xi)||\widehat{f_2}(\xi)||\widehat{S_h}(\xi)| \leq \|\widehat{f_2}\|_\infty^t \|\widehat{f_1}\|_q \|\widehat{f_2}\|_q^{1-t} \|\widehat{S_h}\|_{2s_0}$$

where $t > 0$ is such that $\frac{2-t}{q} + \frac{1}{2s_0} = 1$. By Corollary 2 we know that $\|\widehat{S_h}\|_q \ll_q 1$. Thus $\sum_{a,d\in\mathbf{Z}_N} f(a)f(a+d)S_h(d) \ll_q (1+\eta)^t\epsilon^t M^{2-t}$. We have similar estimates for the other two terms. Thus by choosing ϵ sufficiently small depending on M, q, δ, the contribution of the three last terms is less than $\frac{1}{2}c(h,\delta)$.

Therefore,

$$\sum_{a\in\mathbf{Z}_N}\sum_{d\in\mathbf{Z}_N} f(a)f(a+d)S(d) \geq \left(\frac{1}{2}c(P,\delta) - O_{M,q}(\eta)\right)N^2$$

as required. □

From Proposition 2 we immediately have the following:

Corollary 3. *Let $\kappa = \min(\kappa_1, \kappa_2)$ where κ_1 is the constant in Theorem 8 and κ_2 is the constant in Corollary 2. Then under the same hypothesis as in Proposition 2, we have*

$$\sum_{a\in\mathbf{Z}_N}\sum_{d\in\mathbf{Z}_N} f(a)f(a+d)S_{h_W}(d) \geq \left(\frac{1}{2}c(h,\delta) - O_{M,q,\delta}(\eta)\right)N^2$$

for all N large enough depending on h and δ, and $W < N^\kappa$.

5. Construction of a pseudorandom measure that majorizes the primes

In this section we will find functions f, ν satisfying the conditions of Proposition 1 such that f is supported on the Chen primes. This is done exactly the same way as in the proof of [8, Theorem 1.2], the main tool being the Hardy-Littlewood majorant property for objects called "enveloping sieves".

Let us recall the settings from [8]. Consider $F = \prod_{j=1}^{k}(a_j n + b_j)$, a product of k linear factors with integer coefficients, no two linear factors are rational multiples of each other.

Let $X = X(F) = \{n \in \mathbf{Z}^+ : F(n) \text{ is the product of } k \text{ primes}\}$. For any $q \geq 1$, let $X_q = \{n \in \mathbf{Z}_q : (F(n), q) = 1\}$. Thus $X_{R!} = \{n \in \mathbf{Z} : (d, F(n)) = 1 \text{ for all } 1 \leq d \leq R\}$. Let $\gamma(q) = \frac{|X_q|}{q}$. We assume that $\gamma(q) > 0$ for all $q \geq 1$. Let $\mathfrak{S}_F$ be the singular series $\mathfrak{S}_F = \prod_{p \text{ prime}} \frac{\gamma(p)}{\left(1-\frac{1}{p}\right)^k}$.

Proposition 3 (Proposition 3.1, [8])**.** *Let F be as above, with coefficients a_i, b_i satisfying $|a_i|, |b_i| \leq N$. Let $R \leq N$ be a large integer. Then there is a non-negative function $\beta := \beta_R : \mathbb{Z} \to \mathbf{R}^+$, called the envelopping sieve associated to F and R, with the following properties:*

(i) (Majorant property) *We have*
$$\beta(n) \gg_k \mathfrak{S}_F^{-1} \log^k R \mathbf{1}_{X_{R!}}(n) \tag{2}$$
for all integers n. In particular, $\beta(n)$ is non-negative.

(ii) (Crude upper bound) *We have*
$$\beta(n) \ll_{k,\epsilon} N^{\epsilon} \tag{3}$$
for all $0 < n \leq N$ and $\epsilon > 0$.

(iii) (Fourier expansion) *We have*
$$\beta(n) = \sum_{q \leq R^2} \sum_{a \in \mathbf{Z}_q^*} w(a/q) e_q(-an), \tag{4}$$
where $w(a/q) = w_R(a/q)$ obeys the bound
$$|w(a/q)| \ll_{k,\epsilon} q^{\epsilon - 1} \tag{5}$$
for all $q \leq R^2$ and $a \in \mathbf{Z}_q^*$. Also we have $w(0) = w(1) = 1$.

(iv) (Fourier vanishing properties) *Let $q \leq R^2$ and $a \in \mathbf{Z}_q^*$. If q is not square-free, then $w(a/q) = 0$. Similarly, if $\gamma(q) = 1$ and $q > 1$, then $w(a/q) = 0$.*

It should be mentioned that all the implied constants depend on k, but not on F. Moreover, β_R enjoys the following properties:

Proposition 4 (Discrete majorant property, Proposition 4.2, [8])**.** *For every $q > 2$, we have*
$$\left(\sum_{b \in \mathbf{Z}_N} |\mathbf{E}_{1 \leq n \leq N} a_n \beta_R(n) e_N(-bn)|^q \right)^{1/q} \ll_{q,k} \left(\mathbf{E}_{1 \leq n \leq N} |a_n|^2 \beta_R(n) \right)^{1/2}$$

Proposition 5 (Lemma 4.1, [8])**.** *Suppose $R \leq \sqrt{N}$. Then $\mathbf{E}_{1 \leq n \leq N} \beta_R(n) \ll 1$.*

Suppose $\mathcal{A}$ is a subset of positive relative density of the primes. Let t be a large number (independent of N), and $W = W_t = \prod_{p \leq t} p$. We will assume at all times that $W < N^{\kappa}$, where κ is the constant as in Corollary 3. By the pigeonhole principle we can choose $b \in X_W$ such that the set $X = \{0 \leq n \leq N/2 : \lambda(W)n + b \in \mathcal{A}\}$ satisfies

$$\begin{aligned} |X| &\gg \frac{1}{\phi(\lambda(W))}\frac{N\lambda(W)}{\log(N\lambda(W))} \\ &\gg \frac{\lambda(W)}{\phi(\lambda(W))}\frac{N}{\log N} \\ &\gg \prod_{p\leq t}(1-1/p)^{-1}\frac{N}{\log N} \\ &\gg \log t\frac{N}{\log N} \end{aligned} \tag{6}$$

for infinitely many N. We may assume henceforth that N satisfies the inequality (6). Let us now consider the polynomial $F(n) = \lambda(W)n + b$. Then it is easy to see that $\mathfrak{S} = \prod_{p\leq t}(1 - 1/p)^{-1} \ll \log t$.

Now let $R = [N^{1/20}]$ and let $\beta_R : \mathbf{Z} \to R^+$ be the enveloping sieve associated to F and R. Let ν be the restriction of β on $\{1, \ldots, N\}$ which may be regarded as a function on $\mathbf{Z}_N$. Then we have $\nu(n) \gg \mathfrak{S}^{-1}\log N 1_X(n) \gg \frac{1}{\log t}\log N 1_X(n)$.

Lemma 5 (Lemma 6.1,[8])**.** $\widehat{\nu}(a) = \delta_{a,0} + O(t^{-1/2})$.

Proof of Theorem 4. Let us now define the function $f : \mathbf{Z}_N \to \mathbf{R}^+$ by

$$f(n) = c\frac{\log N}{\log t}1_X(n)$$

Let us verify the conditions of Proposition 2. Clearly $0 \leq f \leq \nu$ for c appropriately chosen, and $\mathbf{E}_{\mathbf{Z}_N} f \geq \delta > 0$, where δ depends only on the upper relative density of $\mathcal{A}$ in the primes.

Fix any $2 < q < 4s_0/(2s_0 - 1)$. By Propositions 4 and 5 (for the sequence $a_n = \frac{f(n)}{\nu(n)}$, with the convention that $a_n = 0$ if $f(n) = \nu(n) = 0$), we have

$$\|\widehat{f}\|_q = \left(\sum_{b\in\mathbf{Z}_N} |\mathbf{E}_{1\leq n\leq N} f(n)e_N(-bn)|^q\right)^{1/q} \ll \left(\mathbf{E}_{1\leq n\leq N}\frac{f(n)^2}{\nu(n)}\right)^{1/2} \ll (\mathbf{E}_{1\leq n\leq N}\nu(n))^{1/2} \ll 1$$

Thus the condition (4) of Proposition 2 is satisfied. Finally, the condition (3) of Proposition 2 follows from Lemma 5 with $\eta = O(t^{-1/2})$.

Proposition 2 now tells us that

$$\sum_{a,d\in\mathbf{Z}_N} f(a)f(a+d)S_{h_W}(d) \geq c(h,\delta) - O(t^{-1/2}) \tag{7}$$

for some constant c depending on h and δ, for N sufficiently large depending on h, and for every $W \leq N^\kappa$. Thus for t sufficiently large depending on h and δ, for N sufficiently large depending on t, we have $\sum_{a,d\in\mathbf{Z}_N} f(a)f(a+d)S_{h_W}(d) > 0$, which implies the existence of a couple $a, a' \in X$ and d such that

$$a - a' = h_W(d) = \frac{h(Wd + r_W)}{\lambda(W)} \neq 0$$

A priori, this is an equality in $\mathbf{Z}_N$, but since $a, a', h_W(d) < \frac{N}{2}$, this is an equality in $\mathbf{Z}$. Therefore, $h(Wd + r_W) = (\lambda(W)a + b) - (\lambda(W)a' + b)$ is the difference of two elements of $\mathcal{A}$, as desired. □

Proof of Theorem 6. The proof goes along the lines that of Theorem 4. Suppose $\mathcal{A}$ is a subset of positive relative density of the Chen primes. This time, we consider $X = \{0 \le n \le N/2 : \lambda(W)n + b \in \mathcal{A}\}$ for some appropriately chosen b, $F = (\lambda(W)n + b)(\lambda(W)n + b + 2))$, and $f = c\frac{\log N}{\log^2 t} 1_X(n)$. □

Remarks 5.1. What we have proved so far is that not only is there a couple p_1, p_2 such that $p_1 - p_2 = h(n)$ for some n, but the number of such couples is of the correct magnitude. More precisely, if $\mathcal{A}$ is a subset of positive upper relative density of the primes, then we have

$$\sharp\{(p_1, p_2) : p_1, p_2 \in \mathcal{A}, p_1, p_2 \le N, p_1 - p_2 = h(n) \text{ for some } n\} \gg \frac{N^{1+1/k}}{\log^2 N}$$

where the implied constant depends only on h and the upper relative density of $\mathcal{A}$. A similar conclusion holds for subsets of positive relative density of the Chen primes.

6. Further discussions

6.1. **A word on bounds.** Recall that in the estimate (7), $c(h, \delta)$ has the form

$$c(h, \delta) = \exp\left(-c_1 \delta^{-(k-1)} \log^{\mu}\left(\frac{2}{\delta}\right)\right)$$

while the error term $O(t^{-1/2})$ takes the form $(C/\epsilon)^{(M/\epsilon)^q} t^{-1/2}$, where M, C are constants depending at most on k, c_1 is a constant depending on h, and ϵ is a power of $c(h, \delta)$. Recall that $t \ll \log W \ll_k \log N$. A calculation shows that the error term is dominated by the main term as long as

$$\delta \gg_h \frac{(\log_4 N)^{\mu/(k-1)}}{(\log_3 N)^{1/(k-1)}}$$

(where $\log_i$ denotes the number of times the log has to be taken). Thus we have proved that, inside any subset of size $\gg_h \frac{N}{\log N} \frac{(\log_4 N)^{\mu/(k-1)}}{(\log_3 N)^{1/(k-1)}}$ of the primes in $\{1, \ldots, N\}$, there must exist two distinct elements p_1, p_2 such that $p_1 - p_2 = h(n)$ for some $n \in \mathbf{Z}$. A similar conclution holds for the Chen primes. Such a bound is of course far weaker than Pintz-Steiger-Szemerédi type bounds.

6.2. **On the transference principle.** Our transference principle relies on two properties of the intersective set $H = \{h(n) : n \in \mathbf{Z}\}$, namely Theorem 7 and Proposition 3. Theorem 7 says that the number of solutions to $a - a' = m$ where a, a' are in any given dense set and $m \in H$ is of the expected order of magnitude. Proposition 3 requires that the number of representations of any number as a sum of elements of H be bounded by the expected order of magnitude. We may ask for which other classes of intersective sets these two properties hold. A natural candidate is the set of values of polynomials of prime variables. It is known that the set $\{Q(p) : p \text{ prime}\}$ is intersective, where $Q \in \mathbf{Z}[x]$ is such that $Q(1) = 0$; however there are other examples such as $Q(p) = (p-3)(p-5)$. Other examples of intersective sets include $\{[\alpha n^2] : n \in \mathbf{Z}^+\}$ for irrational α, and more generally the set of values of certain generalized polynomials (whose intersectivity is established in [2]). We may ask the same question for generalized polynomials in prime variables such as $\{[\alpha p^2] : p \text{ prime}\}$ for α irrational (whose intersectivity is not yet established yet but very plausible). However, as we have seen how the W-trick comes into play, we will have to take into account uniform versions of the two properties, which don't seem to be a simple matter.

References

[1] D. Berend, Y.Bilu, *Polynomials with Roots Modulo Every Integer*, Proceedings of the American Mathematical Society, Vol. 124, No. 6 (Jun., 1996), pp. 1663-1671.

[2] V. Bergelson, I. J. Haland, *Sets of recurrence and generalized polynomials*, **Convergence in Ergodic Theory and Probability**, Eds.: Bergelson/March/Rosenblatt, Walter de Gruyter & Co, Berlin, NewYork, 1996, 91-110.

[3] A. Balog, J. Pelikan, J. Pintz, E. Szemerédi, *Difference sets without κ-th powers*, Acta Math Hung 65: 165-187 (1994).

[4] J.-R. Chen, *On the representation of a large even integer as the sum of a prime and a product of at most two primes*, Sci. Sinica 16 (1973), 157176.

[5] G. V. Chudnovsky, **Contributions to the theory of transcendental numbers**, American Mathematical Society, Providence, RI, 1984.

[6] B. Green, *Roth's Theorem in the primes*, Annals of Math. 161 (2005), no. 3, 1609-1636.

[7] B. Green, T. Tao, *The primes contains arbitrarily long arithmetic progressions.*

[8] B. Green, T. Tao, *Restriction theory of the Selberg sieve, with applications*, Jour. Th. Nombres Bordeaux 18 (2006), 147–182.

[9] H. Furstenberg, **Recurrence in Ergodic Theory and Combinatorial Number Theory**, Princeton Univ. Press, 1981.

[10] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. dAnalyse Math, 71 (1977), pp. 204-256.

[11] T. Kamae and M. Mendès France, *Van der Corput's difference theorem*, Israel J. Math. 31 (1978), no. 3-4, 335-342.

[12] I. Laba and M. Hamel, *Arithmetic structures in random sets*, Integers: Electronic Journal of Combinatorial Number Theory 8 (2008), #A4.

[13] H. Iwaniec, **Sieve methods**, Graduate course, Rutgers 1996.

[14] Y-R. Liu, C. Spencer, *A prime analog of Roth's theorem in function fields*, preprint.

[15] H. Li and H. Pan, *Difference sets and Polynomials of prime variables*, Acta Arith, no.1, 138 (2009), 25-52.

[16] J. Lucier, *Intersective sets given by a polynomial*, Acta Arith., 123 (2006), 57-95.

[17] J. Pintz, W. L. Steiger, E. Szemerédi, *On Sets of Natural Numbers Whose Difference Set Contains No Squares*, J. London Math. Soc., 1988; s2-37: 219-231.

[18] A. Sárközy, *On difference sets of sequences of integers, I.*, Acta Math. Acad. Sci. Hungar. 31 (1978), 125-149.

[19] T. Tao, V. Vu, **Additive Combinatorics**, Cambridge Univ. Press, 2006.

[20] T. Tao, T.Ziegler, *The primes contain arbitrarily long polynomial progressions*, Acta Math. 201 (2008), 213305.

[21] Vaughan, **The Hardy-Littlewood method**, 2nd ed., Cambridge Univ. Press, 1997

UCLA Department of Mathematics, Los Angeles, CA 90095-1596.

E-mail address: leth@math.ucla.edu