Proceedings of the Royal Society of Edinburgh: Section A Mathematics, **156**, 966–1012, *2026*
DOI:10.1017/prm.2024.96

Arithmetic Ramsey theory over the primes

Jonathan Chapman
School of Mathematics, University of Bristol, Bristol BS8 1UG, United Kingdom and
Heilbronn Institute for Mathematical Research, Bristol BS8 1UG, United Kingdom (jonathan.chapman@bristol.ac.uk) (corresponding author)

Sam Chow
Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL, United Kingdom (Sam.Chow@warwick.ac.uk)

(Received 7 February 2024; accepted 12 June 2024)

We study density and partition properties of polynomial equations in prime variables. We consider equations of the form $a_1h(x_1) + \cdots + a_sh(x_s) = b$, where the a_i and b are fixed coefficients and h is an arbitrary integer polynomial of degree d. We establish that the natural necessary conditions for this equation to have a monochromatic non-constant solution with respect to any finite colouring of the prime numbers are also sufficient when the equation has at least $(1 + o(1))d^2$ variables. We similarly characterize when such equations admit solutions over any set of primes with positive relative upper density. In both cases, we obtain lower bounds for the number of monochromatic or dense solutions in primes that are of the correct order of magnitude. Our main new ingredient is a uniform lower bound on the cardinality of a prime polynomial Bohr set.

Keywords: arithmetic combinatorics; arithmetic Ramsey theory; Diophantine equations; Hardy–Littlewood method; partition regularity; restriction theory

2010 Mathematics Subject Classification: 11B30 (primary); 05D10; 11D72; 11L15 (secondary)

1. Introduction

An influential theorem of Szemerédi asserts that sets of positive integers A with positive upper density, meaning that

$$\limsup_{N\to\infty} \frac{|A \cap \{1, 2, \dots, N\}|}{N} > 0,$$

must contain arbitrarily long arithmetic progressions. Green and Tao [7] famously established a version of Szemerédi's theorem for the primes. Specifically, writing

© The Author(s), 2024. Published by Cambridge University Press on behalf of The Royal Society of Edinburgh. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

$\mathcal{P} := \{2, 3, 5, \ldots\}$ for the set of prime numbers, Green and Tao showed that sets $A \subseteq \mathcal{P}$ satisfying

$$\limsup_{N\to\infty} \frac{|A \cap \{1, 2, \ldots, N\}|}{|\mathcal{P} \cap \{1, 2, \ldots, N\}|} > 0 \tag{1.1}$$

contain arbitrarily long arithmetic progressions. In particular, the primes themselves contain arithmetic progressions of any finite length.

One can consider configurations other than arithmetic progressions. We call a system of Diophantine equations *density regular* if it has non-constant solutions over all sets of positive integers with positive upper density. For example, consider a linear homogeneous equation

$$a_1x_1 + \cdots + a_sx_s = 0, \tag{1.2}$$

where $s \geqslant 3$ and $a_1, \ldots, a_s$ are non-zero integers. Roth [20] showed that this equation is density regular if and only if $a_1 + \cdots + a_s = 0$. Green [6] subsequently proved that Roth's theorem holds over the primes; Eq. (1.2) has non-constant solutions over any set of primes $A \subseteq \mathcal{P}$ satisfying (1.1) if and only if $a_1 + \cdots + a_s = 0$.

A related, weaker notion of regularity is that of *partition regularity*, which refers to systems of Diophantine equations, which admit monochromatic non-constant solutions with respect to any finite colouring of the positive integers. By the pigeonhole principle, density regularity implies partition regularity. A foundational result in arithmetic Ramsey theory is *Rado's criterion* [18, Satz IV], which completely characterizes partition regularity for finite systems of linear equations. In particular, Rado's criterion reveals that Eq. (1.2) is partition regular if and only if there exists a non-empty set $I \subseteq \{1, \ldots, s\}$ such that $\sum_{i\in I} a_i = 0$.

Lê [11] observed that Green and Tao's work provides a characterization of partition regularity for systems of linear homogeneous equations in shifted primes. For single equations, Lê proved that Eq. (1.2) admits monochromatic non-constant solutions with respect to any finite colouring of $\mathcal{P}+1 := \{p+1 : p \in \mathcal{P}\}$ or $\mathcal{P}-1$ if and only if there exists a non-empty set $I \subseteq \{1, \ldots, s\}$ such that $\sum_{i\in I} a_i = 0$. Note that there are divisibility obstructions that prevent such a result holding over the set $\mathcal{P} + c$ for integers $c \notin \{-1, 1\}$. For example, the equation $x + y = z$ is partition regular, but if we partition $\mathcal{P} + c$ into residue classes modulo q for any prime q dividing c, then there are no monochromatic solutions to $x + y = z$.

The purpose of this article is to obtain a complete classification of partition and density regularity over primes for equations in sufficiently many variables of the form

$$a_1h(x_1) + \cdots + a_sh(x_s) = b, \tag{1.3}$$

where $a_1, \ldots, a_s$ are non-zero integers, b is an integer, and h is a polynomial with integer coefficients. We say that (1.3) is *partition regular over the primes* if every finite colouring of the prime numbers produces a monochromatic non-constant solution to (1.3). Similarly, we call (1.3) *density regular over the primes* if (1.3) has a non-constant solution over any set of primes $A \subseteq \mathcal{P}$, which satisfies (1.1). For $b = 0$, observe that Lê's result [11] asserts that Rado's condition characterizes partition regularity over primes for (1.3) whenever $h(x) = x \pm 1$.

In our previous work [2], we established necessary and sufficient conditions for partition and density regularity (over $\mathbb{N}$) for all equations (1.3) in sufficiently many

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

variables. We observed that it is necessary for partition regularity that h satisfies a certain 'intersectivity condition' in order to avoid divisibility obstructions, as alluded to above. For partition regularity over primes, we use the following definition, previously introduced in [12, 19]. An integer polynomial h is *intersective of the second kind* if for each positive integer n, there exists an integer x which is coprime to n such that n divides $h(x)$. Observe that any polynomial h satisfying $h(1) = 0$ or $h(-1) = 0$ is intersective of the second kind. However, one can construct numerous polynomials, such as $(x^2 - 13)(x^2 - 17)(x^2 - 221)$ and $(x^3 - 19)(x^2 + x + 1)$, which are intersective of the second kind but do not have rational zeros (see [13, §2]).

Our main result is the following:

THEOREM 1.1 *Let $d \geqslant 2$ be an integer. There exists a positive integer $s_0(d)$ such that the following is true. Let h be an integer polynomial of degree d, and let $s \geqslant s_0(d)$ be an integer. Let $a_1, \dots, a_s$ be non-zero integers, and let b be an integer.*

- *(PR) Equation (1.3) is partition regular over the primes if and only if there exists a non-empty set $I \subseteq \{1, \dots, s\}$ with $\sum_{i \in I} a_i = 0$ and an integer m with $b = (a_1 + \cdots + a_s)m$ such that $h(x) - m$ is an intersective polynomial of the second kind.*
- *(DR) Equation (1.3) is density regular over the primes if and only if $b = a_1 + \cdots + a_s = 0$.*

Furthermore, we have $s_0(2) = 5$, $s_0(3) \leqslant 9$, and

$$s_0(d) \leqslant d^2 - d + 2\lfloor \sqrt{2d+2} \rfloor + 1 \qquad (d \geqslant 4). \tag{1.4}$$

REMARK 1.2. The integer $s_0(d)$, which is defined explicitly in §2, was previously introduced in [2] to study partition and density regularity of equations (1.3).

REMARK 1.3. The "only if" parts of both statements are true without any assumption on the number of variables s (see lemma 2.4). The above theorem asserts that these necessary conditions are sufficient provided $s \geqslant s_0(d)$.

REMARK 1.4. The conditions in the partition regularity statement are trivially satisfied when $b = a_1 + \cdots + a_s = 0$. Indeed, in this case, we can take $I = \{1, \dots, s\}$ and $m = h(1)$. It then follows that $h(x) - m$ vanishes at $x = 1$, whence $h(x) - m$ is trivially intersective of the second kind.

In our previous work [2], we obtained lower bounds for the number of solutions to (1.3) lying in dense or monochromatic subsets of the positive integers, which are sharp up to the leading constant. Our second main result is the following analogous counting version of theorem 1.1 for subsets of $\mathcal{P}_N := \{p \leqslant N : p \text{ prime}\}$.

THEOREM 1.5 *Let $d \geqslant 2$ be an integer, and let $s_0(d)$ be as given in theorem 1.1. Let h be an integer polynomial of degree d. Let $s \geqslant s_0(d)$ be an integer. Let $a_1, \dots, a_s$ be non-zero integers, and let b be an integer. Given a set of integers $\mathcal{A}$, write*

$$S(\mathcal{A}) = \{(x_1, \dots, x_s) \in \mathcal{A}^s : x_i \neq x_j \text{ for all } i \neq j, \text{ and } a_1 h(x_1) + \cdots + a_s h(x_s) = b\}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

(PR) Suppose there exists a non-empty set $I \subseteq \{1,\dots,s\}$ *with* $\sum_{i\in I} a_i = 0$ *and an integer* m *with* $b = (a_1+\cdots+a_s)m$ *such that* $h(x)-m$ *is an intersective polynomial of the second kind. Then, for any positive integer* r*, there exists a positive real number* $c_1 = c_1(h;a_1,\dots,a_s,b;r)$ *and a positive integer* $N_1 = N_1(h;a_1,\dots,a_s,b;r)$ *such that the following is true for any positive integer* $N \geqslant N_1$*. Given any* r*-colouring* $\mathcal{P}_N = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r$*, there exists* $k \in \{1,\dots,r\}$ *such that* $|S(\mathcal{C}_k)| \geqslant c_1 N^{-d}(N/\log N)^s$*.*

(DR) If $a_1 + \cdots + a_s = b = 0$*, then for any positive real number* $\delta > 0$*, there exists a positive real number* $c_2 = c_2(h;a_1,\dots,a_s;\delta)$ *and a positive integer* $N_2 = N_2(h;a_1,\dots,a_s;\delta)$ *such that the following is true for any positive integer* $N \geqslant N_2$*. Given any set* $A \subseteq \mathcal{P}_N$ *satisfying* $|A| \geqslant \delta|\mathcal{P}_N|$*, we have* $|S(A)| \geqslant c_2 N^{-d}(N/\log N)^s$*.*

1.1. Methods

Our goal is to find many monochromatic/dense solutions to

$$L_1(h(\mathbf{x})) = L_2(h(\mathbf{y}))$$

for some linear forms L_1 and L_2, where $L_1(1,\dots,1) = 0$. Here, we have used the slight abuse of notation $h(\mathbf{x}) = (h(x_1),\dots,h(x_s))$. In Fourier space, after normalization and accounting for small prime moduli (the W-trick), the image of h can be shown to behave like $\mathbb{N}$ for our count. The upshot is that it suffices to count solutions to an equation

$$L_1(\mathbf{n}) = L_2(h_D(\mathbf{z})),$$

where h_D is a related polynomial that is intersective of the second kind. This Fourier-analytic transference principle was introduced by Green [6] to show that relatively dense sets of primes contain three-term arithmetic progressions and is based on Fourier decay and restriction (from harmonic analysis). The transference argument is sketched in further detail in the next section and formalized in the five sections afterwards. It can be regarded as a version of the Hardy–Littlewood circle method and estimates for prime Weyl sums feature prominently in our work.

To count monochromatic/dense solutions to our linearized equation, we use an arithmetic regularity lemma. This enables us to decompose the indicator functions of our colour classes into three parts, the first of which exhibits quasi-periodic structure and ultimately dominates the count. Using this quasi-periodicity to obtain a large count requires us to show that polynomials evaluated at primes are dense in Bohr sets, in a suitably uniform sense. The colour class that we choose maximizes this density.

Our main novelty is to uniformly bound from below the density of a 'prime polynomial Bohr set'. This is accomplished by induction on the dimension, beginning with a result of Harman [8] from Diophantine approximation. The latter brings prime Weyl sums into play. The requisite data for these come partly from the analysis in the earlier sections, partly from Lucier's pioneering work on intersective polynomials [15], and partly from the investigations of Lê–Spencer [14] and Rice [19] into polynomials that are intersective of the second kind.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

1.2. Organization

We begin in §2 with some preliminary results. We prove the 'only if' parts of theorem 1.1 by establishing necessary conditions for (1.3) to be partition or density regular over the primes. We synthesize both the density and partition statements presented in theorem 1.5 into a single result, theorem 2.6, on counting solutions to certain linear form equations. We also recall the 'auxiliary intersective polynomials' of Lucier [15] and use them to state theorem 2.8, which is a 'linearized' version of theorem 2.6. Finally, we provide a sketch of the subsequent transference argument we will use to deduce theorem 2.6 from theorem 2.8.

In §3, we introduce formally the 'linearization' procedure, which we will use to infer theorem 2.6 from theorem 2.8. We apply the W-trick and introduce the majorant ν, the latter of which is the focus of our investigations in §5 and §6. To expedite this process, we record in §4 some general results on exponential sums over primes. These will later be used in §5 and §9.

In §5, we study the Fourier transform of our majorant ν by using the Hardy–Littlewood circle method. We follow this in §6 by investigating the restriction properties of ν and a related majorant μ_D. The Fourier decay and restriction estimates we obtain in these sections are then applied in §7 to execute the transference principle. This completes the deduction of theorem 2.6 from theorem 2.8.

The focus of the final two sections is to prove theorem 2.8 using an arithmetic regularity lemma. In §8, we begin this argument by first modifying theorem 2.8 into a new result (theorem 8.1), which is more amenable to arithmetic regularity methods. This reduces matters to counting primes in 'polynomial Bohr sets'. Finally, in §9, we prove theorem 8.1 by establishing density estimates for these prime polynomial Bohr sets.

1.3. Notation

Let $\mathbb{N}$ denote the set of positive integers, and write $\mathcal{P}$ for the set of prime numbers. For each prime p, let $\mathbb{Q}_p$ and $\mathbb{Z}_p$ denote the p-adic numbers and the p-adic integers, respectively. Given a real number $X > 0$, we write $[X] = \{n \in \mathbb{N} : n \leqslant X\}$ and $\mathcal{P}_X = \mathcal{P} \cap [X]$. Set $\mathbb{T} = [0, 1]$. For each $d \in \mathbb{N}$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, we define

$$\|\boldsymbol{\alpha}\| = \max_{1 \leqslant i \leqslant d} \min_{n \in \mathbb{Z}} |\alpha_i - n| = \min_{\mathbf{n} \in \mathbb{Z}^d} \|\boldsymbol{\alpha} - \mathbf{n}\|_\infty.$$

For $q \in \mathbb{N}$ and $x \in \mathbb{R}$, we write $e(x) = e^{2\pi i x}$ and $e_q(x) = e(x/q)$. For $h(x) \in \mathbb{Z}[x]$ and $\mathbf{x} = (x_1, \dots, x_s)$, where $s \in \mathbb{N}$, we abbreviate $h(\mathbf{x}) = (h(x_1), \dots, h(x_s))$. If L is a polynomial with integer coefficients, we write $\gcd(L)$ for the greatest common divisor of its coefficients. The letter ε denotes a small, positive constant, whose value is allowed to differ between separate occurrences.

We employ the Vinogradov and Bachmann–Landau asymptotic notations: for complex-valued functions f and g, we write $f \ll g$ or $g \gg f$ or $f = O(g)$ if there exists a constant C such that $|f(x)| \leqslant C|g(x)|$ holds for all x. We indicate the dependence of the implicit constant C on some parameters $\lambda_1, \dots, \lambda_t$ using subscripts, for example, $f \ll_{\lambda_1, \dots, \lambda_t} g$ or $f = O_{\lambda_1, \dots, \lambda_t}(g)$. We write $f \asymp g$ if $f \ll g$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

and $g \ll f$ both hold. In any statement in which ε appears, we assert that the statement holds for all sufficiently small $\varepsilon > 0$.

For a finitely supported function $f : \mathbb{Z} \to \mathbb{C}$, the *Fourier transform* $\hat{f}$ is defined by

$$\hat{f}(\alpha) := \sum_{n \in \mathbb{Z}} f(n) e(\alpha n) \qquad (\alpha \in \mathbb{R}).$$

Given a real-valued function G, which is bounded over a closed interval $[a, b]$, we write

$$\|G\|_{L^\infty[a,b]} := \sup_{a \leqslant t \leqslant b} |G(t)|.$$

If G is continuously differentiable on an open interval containing $[a, b]$, then we define

$$\|G\|_{\mathcal{S}[a,b]} := \|G\|_{L^\infty[a,b]} + \|(b-a)G'\|_{L^\infty[a,b]}.$$

2. Preliminaries

2.1. Useful ingredients

We will make repeated use of the Siegel–Walfisz theorem [10, lemma 7.14], which we now state for convenience. Recall that the logarithmic integral is given by

$$\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t} \qquad (x \geqslant 2).$$

THEOREM 2.1 (Siegel–Walfisz theorem). *Let $P \geqslant 2$, and write $L = \log P$. Let $A > 0$. Then, there exists $c = c(A) > 0$ such that the following is true. Let $n, q, a \in \mathbb{Z}$ with*

$$1 \leqslant n \leqslant P, \qquad 1 \leqslant q \leqslant L^A.$$

Then,

$$\#\{p \leqslant n : p \equiv a \pmod q\} = \frac{\mathrm{Li}(n)}{\varphi(q)} + O(Pe^{-c\sqrt{\log L}}).$$

We will also make repeated use of [19, lemma 9], which we state below. Owing to an inaccuracy in the published version of Rice's article, we cite a later arXiv version. This is also explained in a remark immediately following [19, lemma 9]. The fact that C only depends on the degree comes from the proof.

LEMMA 2.2. *For any integer $k \geqslant 2$, there exists $C = C(k) > 0$ such that the following holds. Let $g(x) = a_k x^k + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$, and let $W, b \in \mathbb{Z}$. Let $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ be coprime. Let $\omega(q)$ denote the number of distinct prime factors*

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

of q. Let $q = q_1 q_2$, where q_2 is the greatest divisor of q that is coprime to W, and let $\mathrm{cont}(g) = \gcd(a_1, \dots, a_k)$. *Then,*

$$\left| \sum_{\substack{\ell=0 \\ (W\ell+b,q)=1}}^{q-1} e_q(ag(\ell)) \right| \leqslant C^{\omega(q)} \left(\gcd(\mathrm{cont}(g), q_1) \gcd(a_k, q_2)\right)^{1/k} q^{1-1/k}.$$

We note from its proof that the general epsilon-removal lemma [21, lemma 25] holds with $\|\alpha - a/q\|^\kappa$ in place of $\|\alpha - a/q\|$, and for complex-valued f, with the notation therein. For convenience, we state this version below.

LEMMA 2.3. (Epsilon-removal lemma). *Let $\kappa, \epsilon > 0$, $N \in \mathbb{N}$, $K \geqslant 1$, $u > 2/\kappa$, and $v > u + \epsilon$. Let $\phi : [N] \to [0, \infty)$ and $f : [N] \to \mathbb{C}$ with*

$$|f(n)| \leqslant \phi(n) \qquad (n \in [N]).$$

Let C be a large, positive constant, and let

$$Q > C + K^{1+2/(\kappa\epsilon)}, \qquad T > 2Q^2.$$

Let $\mathfrak{M}$ be the union of the sets

$$\mathfrak{M}(q, a) = \{\alpha \in [0, 1] : |\alpha - a/q| \leqslant 1/T\}$$

over coprime integers $0 \leqslant a \leqslant q \leqslant Q$. Assume that

$$\sum_{n \leqslant N} \phi(n) \ll N, \qquad \|\hat{f}\|_u^u \ll KN^{u-1}.$$

Assume, further, that

$$\hat{\phi}(\alpha) \ll \frac{q^{-\kappa} N}{1 + N\|\alpha - a/q\|^\kappa} + o(K^{-2/\epsilon} N) \qquad (\alpha \in \mathfrak{M})$$

and

$$\hat{\phi}(\alpha) = o(K^{-2/\epsilon} N) \qquad (\alpha \notin \mathfrak{M}).$$

Then,

$$\|\hat{f}\|_v^v \ll_v N^{v-1}.$$

Here, $o(K^{-2/\epsilon}N)$ denotes any quantity X such that if $c > 0$ and N is sufficiently large, then

$$X \leqslant cK^{-2/\epsilon} N.$$

This quantity may differ between instances.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

2.2. Necessary conditions

We now provide necessary conditions for Eq. (1.3) to be partition or density regular over the primes. To state our results, we recall that an integer polynomial h is called *intersective* (or *intersective of the first kind)* if, for each $n \in \mathbb{N}$, there exists $x \in \mathbb{Z}$ such that $h(x) \equiv 0 \pmod n$. We call h *intersective of the second kind* if this statement holds under the additional condition such that an x can be found, which is coprime to n. The following lemma demonstrates that intersectivity is a necessary condition for partition regularity of general polynomial equations.

LEMMA 2.4. *Let $s \in \mathbb{N}$ and let $F \in \mathbb{Z}[x_1, \dots, x_s]$. Consider the equation*

$$F(x_1, \dots, x_s) = 0. \tag{2.1}$$

(PR) If (2.1) is partition regular (over the primes), then the single-variable polynomial $F(x, \dots, x) \in \mathbb{Z}[x]$ is intersective (of the second kind).

(DR) If (2.1) is density regular or density regular over the primes, then $F(x, \dots, x)$ is the zero polynomial.

Proof. Suppose (2.1) is partition regular. Let $n \in \mathbb{N}$ and consider the n-colouring of $\mathbb{N}$ defined by partitioning $\mathbb{N}$ into distinct residue classes modulo n. The existence of a monochromatic solution to (2.1) with respect to this colouring implies that $F(t, \dots, t) \equiv 0 \pmod n$ holds for some $t \in [n]$. As n was arbitrary, it follows that $F(x, \dots, x)$ is intersective.

Now suppose (2.1) is partition regular over the primes. Let $n \in \mathbb{N}$. As before, we partition into residue classes modulo n and infer the existence of $t \in [n]$ and primes $p_1, \dots, p_s$, which are not all equal, with $p_1 \equiv \dots \equiv p_s \equiv t \pmod n$ such that $F(p_1, \dots, p_s) = 0$. If we take n to be a prime power, then, since the p_i are not all equal, at least one p_j is coprime to n, whence t and n are coprime. Applying the Chinese remainder theorem, we conclude that $F(x, \dots, x)$ is intersective of the second kind.

Finally, suppose (2.1) is density regular or density regular over the primes. Let $m \in \mathbb{N}$. By the Siegel–Walfisz theorem (in the case of density regularity over the primes), for each prime $p \nmid m$, we can find an integer/prime solution $(x_1, \dots, x_s)$ to (2.1) with $x_1 \equiv \dots \equiv x_s \equiv m \pmod p$. By reducing (2.1) modulo p, we deduce that $F(m, \dots, m)$ is divisible by infinitely many primes, whence $F(m, \dots, m) = 0$. As m was arbitrary, we conclude that $F(x, \dots, x)$ is the zero polynomial. □

We now apply this lemma to (1.3) to establish the 'only if' parts of theorem 1.1. By working modulo $|\mu|n$ for any $n \in \mathbb{N}$, we see that if $\mu \neq 0$ and μh is intersective of the second kind, then so too is h. Note also that the following result does not impose any restriction on the number of variables:

COROLLARY 2.5. *Let $s \in \mathbb{N}$ and let h be an integer polynomial of positive degree. Let $a_1, \dots, a_s$ be non-zero integers, and let b be an integer.*

(PR) If (1.3) is partition regular over the primes, then there exist a non-empty set $I \subseteq \{1, \dots, s\}$ with $\sum_{i \in I} a_i = 0$ and an integer m with $b = (a_1 + \dots + a_s)m$ such that $h(x) - m$ is an intersective polynomial of the second kind.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

(DR) If (1.3) is density regular or density regular over the primes, then $b = a_1 + \cdots + a_s = 0$.

Proof. Throughout this proof, we write $\mu = a_1 + \cdots + a_s$ and $H(x) = \mu h(x) - b$. First suppose (1.3) is partition regular over the primes. Applying lemma 2.4 to the polynomial $F(x_1, \ldots, x_s) = a_1x_1 + \cdots + a_sx_s - b$, we find that $H(x)$ is intersective of the second kind. By considering solutions to $H(x) \equiv 0 \pmod d$ for any $d \mid \mu$, we observe that $\mu \mid b$. If $\mu \neq 0$, then there is a unique $m \in \mathbb{Z}$ with $b = \mu m$ such that $H(x) = \mu(h(x) - m)$, whence $h(x) - m$ is intersective of the second kind. If $\mu = 0$, then $b = 0$, and so, upon taking $m = h(1)$, we have $b = \mu m$, and $h(x) - m$ is trivially intersective of the second kind. In both cases, Eq. (1.3) becomes

$$\sum_{i=1}^{s} a_i(h(x_i) - m) = 0,$$

for some $m \in \mathbb{Z}$ such that $h(x) - m$ is intersective of the second kind. As this new equation is partition regular, we infer from [2, proposition 2.1] the existence of a set $I \subseteq \{1, \ldots, s\}$ with the desired properties.

Finally, suppose that (1.3) is density regular or density regular over the primes. Then, lemma 2.4 implies that $H(x) = \mu h(x) - b$ is the zero polynomial. Since h has positive degree, we conclude that $b = \mu = 0$. □

In view of these necessary conditions, theorem 1.1 is now an immediate consequence of theorem 1.5.

2.3. Linear form equations

Having dispensed with the necessary conditions for partition and density regularity, we focus on finding monochromatic or dense solutions to (1.3). The necessary conditions we have established, therefore, inform us that (1.3) takes the shape

$$\sum_{i \in I} a_i(h(x_i) - m) = -\sum_{j \in [s] \setminus I} a_j(h(x_j) - m),$$

where $I \subseteq [s]$ is non-empty with $\sum_{i \in I} a_i = 0$, and $m \in \mathbb{Z}$ is such that $b = (a_1 + \cdots + a_s)m$ and $h(x) - m$ is intersective of the second kind. Upon replacing $h(x)$ with $h(x) - m$, we can therefore reduce to the case where $b = 0$ and $h(x)$ is intersective of the second kind.

To find monochromatic or dense solutions to (1.3) with $b = 0$, we study equations of the form

$$L_1(h(\mathbf{x})) = L_2(h(\mathbf{y})), \tag{2.2}$$

for some linear forms L_1 and L_2. To avoid trivialities, we only consider non-degenerate linear forms, where $L(\mathbf{x}) = a_1x_1 + \cdots + a_sx_s$ is *non-degenerate* if $a_i \neq 0$ for all $i \in [s]$. For this new equation, the necessary conditions for partition and density regularity become $L_1(1, \ldots, 1) = 0$. Following the recent works [2, 5], we address both density and partition regularity for (2.2) simultaneously by seeking

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

solutions where the x_i variables are sourced in a dense subset of $\mathcal{P}_X$, while the remaining y_j variables come from a colour class $\mathcal{C}_k \subseteq \mathcal{P}_X$.

Before proceeding to our results, we require some notation. We begin by providing an explicit description of the threshold $s_0(d)$ for the number of variables required in our main theorems. Let $T = T(d) \in \mathbb{N}$ be minimal such that if $h(x) \in \mathbb{Z}[x]$ has degree d, then

$$h(x_1) + \cdots + h(x_T) = h(x_{T+1}) + \cdots + h(x_{2T})$$

has $O_{h,\varepsilon}(X^{2T-d+\varepsilon})$ solutions $\mathbf{x} \in [X]^{2T}$. Equivalently, by orthogonality, $T = T(d)$ is the smallest positive integer such that the moment estimate

$$\int_{\mathbb{T}} \left| \sum_{x \leqslant X} e(\alpha h(x)) \right|^{2T} \ll_{h,\varepsilon} X^{2T-d+\varepsilon}$$

holds for any integer polynomial h of degree d. The quantity $s_0(d)$ appearing in theorem 1.1 is now defined to be $s_0(d) := 2T(d) + 1$.

It follows from Hua's lemma [9, equation (1)] that $T(2) \leqslant 2$ and $T(3) \leqslant 4$. In general, the proof of [23, corollary 14.7] delivers

$$T(d) \leqslant \frac{d(d-1)}{2} + \lfloor \sqrt{2d+2} \rfloor.$$

These observations verify the bound (1.4) for $s_0(d)$. Finally, by considering solutions with $x_i = x_{i+T}$ for $i = 1, 2, \ldots, T$, we record the lower bounds

$$T(d) \geqslant d, \qquad s_0(d) \geqslant 2d + 1. \tag{2.3}$$

We can now state our main result on partition and density regularity over primes for linear form equations (2.2).

THEOREM 2.6 *Let r and $d \geqslant 2$ be positive integers, and let $0 < \delta \leqslant 1$. Let h be an integer polynomial of degree d, which is intersective of the second kind. Let $s \geqslant 1$ and $t \geqslant 0$ be integers such that $s + t \geqslant s_0(d)$. Let*

$$L_1(\mathbf{x}) \in \mathbb{Z}[x_1, \ldots, x_s], \qquad L_2(\mathbf{y}) \in \mathbb{Z}[y_1, \ldots, y_t]$$

be non-degenerate linear forms such that $L_1(1, \ldots, 1) = 0$. Then, there exist

$$X_0 = X_0(\delta, h, r, L_1, L_2) \in \mathbb{N}, \qquad \tau_0(\delta) = \tau_0(h, r, L_1, L_2; \delta) \in (0, 1)$$

such that the following is true for all $X \geqslant X_0$. Suppose $\mathcal{P}_X = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r$. Then, there exists $k \in [r]$ with $|\mathcal{C}_k| \geqslant \tau_0(\delta)|\mathcal{P}_X|$ such that if $A \subseteq \mathcal{P}_X$ satisfies $|A| \geqslant \delta|\mathcal{P}_X|$, then

$$\#\{(\mathbf{x}, \mathbf{y}) \in A^s \times \mathcal{C}_k^t : L_1(h(\mathbf{x})) = L_2(h(\mathbf{y}))\} \gg \frac{X^{s+t-d}}{(\log X)^{s+t}}.$$

The implied constant may depend on h, L_1, L_2, r, δ.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

REMARK 2.7. In the case $t=0$, we have a linear form L_2 in zero variables, and we are counting solutions $\mathbf{x} \in A^s$ to the equation

$$L_1(h(\mathbf{x})) = 0.$$

Note that when $t=0$, all linear forms L_2 in t variables are vacuously non-degenerate.

By harnessing a combinatorial 'cleaving' argument of Prendiville [17], we can swiftly deduce theorem 1.5 from theorem 2.6.

Proof of theorem 1.5 given theorem 2.6. Following the argument given at the beginning of this subsection, we may reduce to the case where $b=0$ and h is intersective of the second kind. Combining [2, lemma 3.2] with (2.3), for N sufficiently large, the number of solutions $\mathbf{x} \in [N]^s$ to (1.3) such that $x_i = x_j$ holds for some $i \neq j$ is $O_\varepsilon(N^{s-d+\varepsilon-1/2})$. Therefore, by setting $t=0$, we see that the density regularity statement in theorem 1.5 follows directly from theorem 2.6. Similarly, given a colouring $\mathcal{P}_N = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r$, provided N is sufficiently large, it remains to show that there are at least $c_1 N^{-d}(N/\log N)^s$ monochromatic solutions to (2.2) with

$$L_1(\mathbf{x}) := \sum_{i \in I} a_i x_i, \quad \text{and} \quad L_2(\mathbf{y}) := -\sum_{i \in [s]\setminus I} a_i y_j.$$

For each $\delta > 0$, let $\tau_0(\delta) \in (0,1)$ be as given in the statement of theorem 2.6. By making minor adjustments, we may assume that $\tau_0(\delta) \leqslant \delta$. Set $\delta_0 = 1/r$, and for each $i \in [r]$, let $\delta_i := \tau_0(\delta_{i-1})$, whence $0 < \delta_r \leqslant \ldots \leqslant \delta_0 < 1$. Take $N \geqslant X_0(\delta_r, h, r, L_1, L_2)$ as in theorem 2.6, and suppose $\mathcal{P}_N = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r$. For each $0 \leqslant i \leqslant r$, let $k_i \in [r]$ be the index given by applying theorem 2.6 with $\delta = \delta_i$. By the pigeonhole principle, we can find $k \in [r]$ and $0 \leqslant i < j \leqslant r$ such that $k_i = k_j = k$. Therefore,

$$\frac{|\mathcal{C}_k|}{|\mathcal{P}_N|} \geqslant \tau_0(\delta_i) = \delta_{i+1} \geqslant \delta_j,$$

and

$$\#\{(\mathbf{x},\mathbf{y}) \in A^{|I|} \times \mathcal{C}_k^{s-|I|} : L_1(h(\mathbf{x})) = L_2(h(\mathbf{y}))\} \gg_{\delta_j,h,r,L_1,L_2} \frac{N^{s-d}}{(\log N)^s}$$

holds for any $A \subseteq \mathcal{P}_N$ with $|A| \geqslant \delta_j |\mathcal{P}_N|$. Taking $A = \mathcal{C}_k$ finishes the proof. □

2.4. Auxiliary intersective polynomials

The next step of our argument is to use a version of Green's Fourier-analytic transference principle [6] to obtain solutions to (2.2) by 'transferring' solutions from a 'linearized' equation. To make this precise, we first need to introduce the auxiliary intersective polynomials of Lucier [15], which emerge during the execution of this process.

Let h be an integer polynomial of positive degree d, which is intersective of the second kind. Thus, for each prime p, we can find a p-adic unit $z_p \in \mathbb{Z}_p^\times$ such that

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

$h(z_p) = 0$. Throughout this article, we fix a choice of z_p for each prime p and let m_p be the multiplicity of z_p as a zero of h. For each prime p and positive integer D, let $\mathrm{ord}_p(D)$ denote the largest non-negative integer n such that p^n divides D. We can then define the completely multiplicative function

$$\lambda(D) := \prod_p p^{m_p \mathrm{ord}_p(D)} \qquad (D \in \mathbb{N}). \tag{2.4}$$

By noting that $1 \leqslant m_p \leqslant \deg(h) = d$ for all p, we have

$$D \mid \lambda(D) \mid D^d. \tag{2.5}$$

For each prime p and non-negative integer k, reducing z_p modulo p^k reveals that there exists a unique residue $x \in \mathbb{Z}/p^k\mathbb{Z}$ such that $x \equiv z_p \pmod{p^k\mathbb{Z}_p}$. By the Chinese remainder theorem, we can therefore find a unique integer r_D in the range $(-D, 0]$, which satisfies

$$r_D \equiv z_p \pmod{p^{\mathrm{ord}_p(D)}\mathbb{Z}_p}$$

for all primes p. As h is intersective of the second kind, we have $(r_D, D) = 1$.

Finally, with this notation in place, we define the auxiliary intersective polynomial

$$h_D(x) := \frac{h(r_D + Dx)}{\lambda(D)} \in \mathbb{Z}[x].$$

These polynomials and the surrounding notation were introduced by Lucier [15], who also showed that h_D is indeed a polynomial with integer coefficients [15, lemma 21]. The most important property of these auxiliary polynomials is that the greatest common divisor of the non-constant coefficients of h_D is bounded uniformly in D. Specifically, for all $D \in \mathbb{N}$, [15, lemma 28] states that

$$\gcd(h_D - h_D(0)) \ll_h 1.$$

As in [2, §6], this bound is crucial to our investigation of exponential sums involving intersective polynomials (see §7 and §9).

We can now state our linearized version of theorem 2.6.

THEOREM 2.8 *Let r and $d \geqslant 2$ be positive integers, and let $0 < \delta \leqslant 1$. Let h be an integer polynomial of degree d which is intersective of the second kind. Let $s \geqslant 1$ and $t \geqslant 0$ be integers such that $s + t \geqslant s_0(d)$. Let*

$$L_1(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_s], \qquad L_2(\mathbf{y}) \in \mathbb{Z}[y_1, \dots, y_t]$$

be non-degenerate linear forms such that $L_1(1, \dots, 1) = 0$. Then, there exists

$$Z_0 = Z_0(D, h, r, \delta, L_1, L_2) \in \mathbb{N} \quad \text{and} \quad \eta = \eta(d, \delta, L_1, L_2) \in (0, 1)$$

such that the following is true. Let $D, Z \in \mathbb{N}$ satisfy $Z \geqslant Z_0$, and set $N := h_D(Z)$. Suppose

$$[\eta Z, Z] \cap \{z \in [Z] : r_D + Dz \in \mathcal{P}\} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Then, there exists $k \in [r]$ such that if $\mathcal{A} \subseteq [N]$ satisfies $|\mathcal{A}| \geqslant \delta N$, then

$$\#\{(\mathbf{n}, \mathbf{z}) \in \mathcal{A}^s \times \mathcal{C}_k^t : L_1(\mathbf{n}) = L_2(h_D(\mathbf{z}))\} \gg N^{s-1} \left(\frac{DZ}{\varphi(D)\log Z}\right)^t.$$

The implied constant may depend on h, L_1, L_2, r, δ.

REMARK 2.9. The quantity η is introduced for technical reasons concerning certain weight functions ν_D we employ when applying the transference principle. For further details, see the remarks preceding lemma 8.3.

The proof of theorem 2.8 is deferred to the final two sections of this article. As in [2], we prove this 'linearized' result by applying an arithmetic regularity lemma. To streamline this forthcoming argument, we require the following proposition, which is a minor variation of [2, proposition 3.10] and is proved in the same way.

PROPOSITION 2.10. *Suppose that theorem 2.8 is true in the cases where* $\gcd(L_1) = 1$. *Then, subject to altering the quantities $Z_0(D, r, \delta, L_1, L_2, P)$, η, and the implicit constant in the final bound, theorem 2.8 holds in general.*

2.5. Sketch of the transference argument

In this subsection, we outline how the transference principle allows us to deduce theorem 2.6 from theorem 2.8. Fix an integer polynomial h that is intersective of the second kind, as well as a pair of linear forms L_1 and L_2 as in the statement of theorem 2.6. We begin by recalling that theorem 2.6 concerns the equation

$$L_1(h(\mathbf{x})) = L_2(h(\mathbf{y})), \tag{2.6}$$

while theorem 2.8 considers, for some parameter $D \in \mathbb{N}$, the 'linearized' equation

$$L_1(\mathbf{n}) = L_2(h_D(\mathbf{z})). \tag{2.7}$$

Suppose we have a finite colouring $\mathcal{P}_X = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r$ and a set $A \subseteq \mathcal{P}_X$ with $|A| \geqslant \delta|\mathcal{P}_X|$. For the convenience of this sketch, assume that $X \equiv r_D \pmod D$. Choosing $Z \in \mathbb{N}$ such that $X = r_D + DZ$, we can define an r-colouring

$$\{z \in [Z] : r_D + Dz \in \mathcal{P}\} = \tilde{\mathcal{C}}_1 \cup \cdots \cup \tilde{\mathcal{C}}_r$$

by

$$\tilde{\mathcal{C}}_i := \{z \in [Z] : r_D + Dz \in \mathcal{C}_i\}.$$

Let $N := h_D(Z)$. By pigeonholing, we find a 'dense' set $\mathcal{A} \subseteq [N]$ such that

$$\mathcal{A} \subseteq \left\{\frac{h(x) - h(b)}{\lambda(D)} : x \in A\right\},$$

for some integer b.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Theorem 2.8 now informs us that there are many solutions $(\mathbf{n}, \mathbf{z}) \in \mathcal{A}^s \times \tilde{\mathcal{C}}_k^t$ to (2.7) for some $k \in [r]$. Given such a solution, our construction of $\mathcal{A}$ and $\tilde{\mathcal{C}}_k$ furnishes a solution $(\mathbf{x}, \mathbf{y}) \in A^s \times \mathcal{C}_k^t$ to (2.6) satisfying

$$n_i = \frac{h(x_i) - h(b)}{\lambda(D)}, \quad y_j = r_D + Dz_j \qquad (1 \leqslant i \leqslant s, \quad 1 \leqslant j \leqslant t).$$

Since the map $(\mathbf{n}, \mathbf{z}) \mapsto (\mathbf{x}, \mathbf{y})$ is injective, this argument allows us to obtain many solutions to (2.6). However, observe that the number of solutions to (2.6) given in theorem 2.6 is $X^{s-d+t+o(1)}$, which is far fewer than the number of solutions to (2.7) provided by theorem 2.8, namely $X^{ds-d+t+o(1)}$. This shortfall is handled by instead considering weighted counts of solutions to (2.6). Our task is then to construct an appropriate weight function ν, which is supported on the set

$$[N] \cap \left\{ \frac{h(x) - h(b)}{\lambda(D)} : b < x \leqslant X \right\}.$$

The key utility of the transference principle is that, provided our weight function is suitably 'pseudorandom', we can find a 'dense model' $g : [N] \to [0, 1]$ such that $\hat{\nu} \approx \hat{g}$. Applying theorem 2.8 to a set of the form $\mathcal{A} = \{x \in [N] : g(x) > c\}$, our argument above allows us to prove theorem 2.6.

To ensure our weight ν is sufficiently pseudorandom, we have to contend with the fact that the set $h(\mathcal{P})$ is not equidistributed in residue classes. This issue prevents one from simply taking ν to be a scaled version of the indicator function of $h(\mathcal{P})$. Fortunately, there is a standard technical manoeuvre, known as the *W-trick*, developed by Green [6] to account for equidistribution modulo small primes. In the setting discussed above, this amounts to demanding that our weight ν is supported on a set of the form

$$[N] \cap \left\{ \frac{h(x) - h(b)}{\lambda(D)} : b < x \leqslant X, \quad x \equiv b \,(\mathrm{mod}\, W\kappa) \right\},$$

for some $W, \kappa \in \mathbb{N}$ such that W is divisible by all primes $p \leqslant w$ for some sufficiently large $w \in \mathbb{N}$. If we choose D, W, κ appropriately, then we can ensure that the set

$$\left\{ \frac{h(x) - h(b)}{\lambda(D)} : b < x \leqslant X, \quad x \equiv b \,(\mathrm{mod}\, W\kappa) \right\}$$

equidistributes over congruence classes modulo p for any prime $p \leqslant w$. The contribution of the remaining primes is then subsumed by the error term emerging from the transference of solutions from the 'dense model' g to ν. The appearance of the additional parameter $\kappa \in \mathbb{N}$ here, resulting in a 'double W-trick', was the main innovation of our previous work [2]. Its purpose is to ensure that $\lambda(D)$ precisely accounts for all common divisors of the values of $h(x) - h(b)$, as x ranges over the arithmetic progression b modulo $W\kappa$.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

3. Linearization and the W-trick

In this section, we execute the 'double W-trick' and construct the weight function ν needed for our application of the transference principle. Throughout this section, we fix the parameters

$$\delta, h, r, L_1, L_2,$$

which appear in theorem 2.6.

3.1. The W-trick

Consider a set $A \subseteq \mathcal{P}_X$ with $|A| \geqslant \delta|\mathcal{P}_X|$. Let $C \in \mathbb{N}$ be large with respect to the fixed parameters, and let $w \in \mathbb{N}$ be large in terms of C. Define

$$M = Cd^2 10^{4w}, \qquad W = \left(\prod_{p \leqslant w} p\right)^{100dw}, \qquad V = \sqrt{W}$$

and

$$D = W^2, \qquad Z = \frac{X - r_D}{D}, \qquad N = h_D(Z) = \frac{h(X)}{\lambda(D)}.$$

Henceforth, we take $X \in \mathbb{N}$ sufficiently large in terms of C, w, and the fixed parameters. We also assume that $D \mid (X - r_D)$, whence $Z \in \mathbb{N}$.

For $R \in \mathbb{N}$ and $b \in [R]$, we write

$$A_{b,R} := \{x \in A : x \equiv b \pmod R\}.$$

We denote by $(H, W)_d$ the greatest $m \in \mathbb{N}$ for which $m^d \mid (H, W)$. By [2, lemma A.5] and the Siegel–Walfisz theorem, we have

$$\delta|\mathcal{P}_X| \leqslant |A| \leqslant \sum_{\substack{b \in [W]: \\ (b,W)=1 \\ (h'(b),W)_d \leqslant M}} |A_{b,W}| + O\left(10^w W M^{-1/2} \frac{X}{\varphi(W)\log X}\right).$$

Since w is large relative to C and d, we have $M < 2^{50w}$. Hence, if $(h'(b), W)_d \leqslant M$, then there cannot exist a prime $p \leqslant w$, which divides $(h'(b), W)$ with multiplicity greater than $50dw$. It follows that if $(h'(b), W)_d \leqslant M$, then $(h'(b), W) \mid V$. By incorporating the crude estimate

$$\frac{W}{\varphi(W)} = \prod_{p \leqslant w} \left(1 - \frac{1}{p}\right)^{-1} \leqslant 2^w,$$

we find that

$$\frac{\delta X}{\log X} \ll \sum_{\substack{b \in [W]: \\ (b,W)=1 \\ (h'(b),W)|V}} |A_{b,W}|.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Thus, there exists $b_0 \in [W]$ such that

$$|A_{b_0,W}| \gg \frac{\delta X}{\varphi(W)\log X}, \qquad (b_0, W) = 1, \qquad (h'(b_0), W) \mid V.$$

Define $\kappa \in \mathbb{N}$ by

$$W\kappa(h'(b_0), W) = \lambda(D).$$

Note that (2.4) implies that κ is w-smooth, whence $\varphi(W)\kappa = \varphi(W\kappa)$. By pigeonholing, we can then find $b \in [W\kappa]$ such that

$$b \equiv b_0 \pmod W, \qquad |A_{b,W\kappa}| \gg \frac{\delta X}{\varphi(W)\kappa\log X} = \frac{\delta X}{\varphi(W\kappa)\log X}.$$

Since $(h'(b), W) = (h'(b_0), W) \mid V$, we also have

$$(h'(b), W\kappa) = (h'(b), W).$$

3.2. The weight function

Our next task is to construct an appropriately 'pseudorandom' weight function. Let w, W, and κ be as defined in the previous subsection. Fix some $b \in [W\kappa]$, which satisfies

$$(h'(b), W) \mid V = \sqrt{W}. \tag{3.1}$$

We then define

$$\nu : \mathbb{Z} \to [0,\infty), \qquad \nu(n) := \frac{\varphi(W)}{W(h'(b), W)} \sum_{\substack{b<p\leqslant X\\ p\equiv b \pmod{W\kappa}\\ h(p)-h(b)=n\lambda(D)}} h'(p)\log p. \tag{3.2}$$

Observe that ν is supported on the set

$$\left\{n \in \mathbb{N} : n = \frac{h(p)-h(b)}{\lambda(D)},\ p \in \mathcal{P}_X,\ p \equiv b \pmod{W\kappa}\right\} \subseteq [N].$$

Recall from the previous subsection that we are considering a fixed set $A \subseteq \mathcal{P}_X$ with $|A| \geqslant \delta|\mathcal{P}_X|$, and that we made a judicious choice of $b \in [W\kappa]$ so that $|A_{b,W\kappa}|$ is suitably dense. For this specific choice of b, let

$$\mathcal{A} = \left\{\frac{h(p)-h(b)}{\lambda(D)} : p \in A_{b,W\kappa}\right\}. \tag{3.3}$$

LEMMA 3.1. *Let $\mathcal{A}, \nu$ be as defined above. If X is sufficiently large in terms of h, δ, w, then*

$$\sum_{n\in\mathcal{A}} \nu(n) \gg_\delta N.$$

Proof. Let c be a small, positive constant, and let

$$\Omega_{b,W\kappa,c}(X) = \{p \leqslant c\delta X :\ p \equiv b \pmod{W\kappa}\}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

For X sufficiently large, the Siegel–Walfisz theorem implies that $|A_{b,W\kappa}| \geqslant |\Omega_{b,W\kappa,c}(X)|$. It follows that there exists an injective map $\psi : \Omega_{b,W\kappa,c}(X) \to A_{b,W\kappa}$ such that $p \leqslant \psi(p)$ for all $p \in \Omega_{b,W\kappa,c}(X)$. This implies that

$$\sum_{\substack{p \leqslant c\delta X \\ p \equiv b \pmod{W\kappa}}} p^{d-1} \log p \leqslant \sum_{\substack{p \leqslant c\delta X \\ p \equiv b \pmod{W\kappa}}} \psi(p)^{d-1} \log(\psi(p)) \leqslant \sum_{p \in A_{b,W\kappa}} p^{d-1} \log p.$$

Invoking the bound $h'(x) \ll x^{d-1}$, we deduce that

$$\sum_{p \in A_{b,W\kappa}} h'(p) \log p \gg \sum_{\substack{p \leqslant c\delta X \\ p \equiv b \pmod{W\kappa}}} p^{d-1} \log p.$$

By the Siegel–Walfisz theorem again, we thus have

$$\sum_{p \in A_{b,W\kappa}} h'(p) \log p \gg_\delta \frac{X^d}{\varphi(W\kappa)}.$$

Therefore,

$$\frac{W(h'(b), W)}{\varphi(W)} \sum_{n \in \mathcal{A}} \nu(n) = O((W\kappa)^{d-1} \log W) + \sum_{p \in A_{b,W\kappa}} h'(p) \log p \gg_\delta \frac{X^d}{\varphi(W\kappa)}.$$

Thus, for our choice of κ and b, the desired bound now follows from the equalities

$$\frac{W(h'(b), W)}{\varphi(W)} = \frac{\lambda(D)}{\kappa\varphi(W)} = \frac{\lambda(D)}{\varphi(W\kappa)}.$$

□

4. Exponential sums

In this section, we record some results on exponential sums of the form

$$\sum_{\substack{p \leqslant t \\ p \equiv b \pmod{m}}} e(F(p))G(p), \tag{4.1}$$

where F is a real polynomial, and $G : (1, \infty) \to \mathbb{R}$ is a continuously differentiable function. We apply these results in §5 to study the Fourier transform $\hat{\nu}$ of our weight function ν. The results of this section are also used in §9 to establish density bounds for 'prime polynomial Bohr sets'.

A standard observation in analytic number theory, going back over a century to Hardy and Littlewood, is that such exponential sums can only be large if their phases exhibit 'major arc' behaviour. In the case of (4.1), this means that the leading coefficient of the polynomial F must be very close to a rational number with small denominator. To elucidate this further, we record the following lemma from [10], which considers the situation where the leading coefficient of F is rational.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

In what follows, and throughout this section, for all $k \in \mathbb{N}$, let σ_k be large in terms of k and put $C_k = 2^{8k}\sigma_k$.

LEMMA 4.1. *Let $m \in \mathbb{N}$ and $b \in \mathbb{Z}$ be coprime. Let $F(y) \in \mathbb{R}[y]$ have degree k, and suppose a/q is its leading coefficient, where $a, q \in \mathbb{Z}$ are coprime and*

$$(\log P)^{C_k} < q \leqslant \frac{P^k}{(\log P)^{C_k}}.$$

Assume that P is sufficiently large in terms of m. Then,

$$\sum_{\substack{p \leqslant P \\ p \equiv b \pmod m}} e(F(p)) \ll_k \frac{P}{(\log P)^{\sigma_k+1}}.$$

Proof. This follows immediately from [10, theorem 10]. □

Using this lemma, we can show that (4.1) is small when the leading coefficient of F is 'minor arc', meaning that it is not well-approximated by a rational number with denominator at most polylogarithmic in P.

LEMMA 4.2. *Let $m \in \mathbb{N}$ and $b \in \mathbb{Z}$ be coprime. Let $F(y) \in \mathbb{R}[y]$ have degree k, and let θ be its leading coefficient. Let $G : (1, \infty) \to \mathbb{R}$ be a continuously differentiable function. Assume that P is sufficiently large in terms of m, and that*

$$\max\{q, P^k \|q\theta\|\} > (\log P)^{2C_k} \qquad (q \in \mathbb{N}).$$

Then,

$$\sum_{\substack{p \leqslant P \\ p \equiv b \pmod m}} e(F(p))G(p)\log p \ll_k \frac{P}{(\log P)^{\sigma_k}} \cdot \|G\|_{\mathcal{S}[2,P]}.$$

Proof. By Dirichlet's approximation theorem, there exist coprime $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ such that

$$q \leqslant \frac{P^k}{(\log P)^{2C_k}}, \qquad |q\theta - a| \leqslant \frac{(\log P)^{2C_k}}{P^k}.$$

By our assumption, we also have

$$q > (\log P)^{2C_k}.$$

Thus, $\beta := \theta - a/q$ satisfies

$$|\beta| \leqslant P^{-k}.$$

Let $f(y) = F(y) - \beta y^k$. By partial summation [22, lemma 2.6], we have

$$\sum_{\substack{p \leqslant P \\ p \equiv b \pmod m}} e(F(p))G(p)\log p = A(P)\psi(P) - \int_2^P A(t)\psi'(t)\,\mathrm{d}t,$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

where

$$\psi(t) := e(\beta t^k)G(t)\log t, \qquad A(t) := \sum_{\substack{p\leqslant t\\ p\equiv b \pmod m}} e(f(p)).$$

We deduce from lemma 4.1 and the trivial bound $|A(t)| \leqslant t$ that

$$A(t) \ll \frac{P}{(\log P)^{\sigma_k+1}} \qquad (2 \leqslant t \leqslant P)\,. \tag{4.2}$$

This implies that

$$A(P)\psi(P)(\log P)^{\sigma_k} \ll_k PG(P).$$

It, therefore, remains to estimate

$$\int_2^P A(t)\psi'(t)\,\mathrm{d}t = I_1 + I_2 + I_3 + I_4,$$

where

$$I_1 = \int_2^P A(t)G'(t)e(\beta t^k)\log t\,\mathrm{d}t, \quad I_2 = \beta k\int_2^P t^{k-1}A(t)G(t)e(\beta t^k)\log t\,\mathrm{d}t,$$
$$I_3 = \int_2^{P/(\log P)^{\sigma_k}} (A(t)/t)G(t)e(\beta t^k)\,\mathrm{d}t, \quad I_4 = \int_{P/(\log P)^{\sigma_k}}^P (A(t)/t)G(t)e(\beta t^k)\,\mathrm{d}t.$$

The bound (4.2) gives

$$I_1(\log P)^{\sigma_k} \ll P^2 \max_{2\leqslant t\leqslant P}|G'(t)|.$$

Similarly, since $|\beta| \leqslant P^{-k}$, we see that

$$I_2(\log P)^{\sigma_k} \ll P \max_{2\leqslant t\leqslant P}|G(t)|.$$

Using the trivial bound $|A(t)| \leqslant t$, we have

$$I_3(\log P)^{\sigma_k} \ll P \max_{2\leqslant t\leqslant P/(\log P)^{\sigma_k}}|G(t)|.$$

Similarly, we deduce from (4.2) that

$$I_4(\log P)^{\sigma_k} \ll \int_{P/(\log P)^{\sigma_k}}^P |G(t)|\,\mathrm{d}t \leqslant P \max_{2\leqslant t\leqslant P}|G(t)|.$$

Combining these estimates completes the proof. □

The above two lemmas suffice to handle 'minor arc' behaviour. As is typical in applications of the circle method, we treat the major arcs by establishing asymptotic formulae for the exponential sums (4.1).

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

LEMMA 4.3. (General major arc asymptotic). *Let $f(y) \in \mathbb{Z}[y]$ have degree k, and let $G : (1,\infty) \to \mathbb{R}$ be a continuously differentiable function. Let $b \in \mathbb{Z}$ and $m \in \mathbb{N}$ be coprime, and let $Q \in \mathbb{N}$ with*

$$\frac{f(b+mx)-f(b)}{Q} \in \mathbb{Z}[x]. \tag{4.3}$$

Let $\theta \in \mathbb{R}$ and $P \geqslant 2$, and suppose $(q,a) \in \mathbb{N} \times \mathbb{Z}$ with $(a,q) = 1$ and

$$q \ll (\log P)^{2C_k}, \qquad |q\theta - a| \ll \frac{Q(\log P)^{2C_k}}{P^k}.$$

Let $c > 0$ be constant, small in terms of C_k, and put $\beta = \theta - a/q$. If P is sufficiently large relative to m and Q, then

$$\sum_{\substack{p \leqslant P \\ p \equiv b \pmod m}} e_Q(\theta f(p))G(p)\log p = I_{f,G}(\beta)\frac{S(q,a;m)}{\varphi(mq)} + O_f(Pe^{-c\sqrt{\log P}}\|G\|_{\mathcal{S}[2,P]}),$$

where

$$I_{f,G}(\beta) = \int_2^P e_Q(\beta f(t))G(t)\,\mathrm{d}t, \qquad S(q,a;m) = \sum_{\substack{t \pmod{mq} \\ (t,q)=1 \\ t \equiv b \pmod m}} e_{Qq}(af(t)).$$

REMARK 4.4. The condition (4.3) holds if $f(b+mx)/Q \in \mathbb{Z}[x]$.

Proof. Writing $g(x) \in \mathbb{Z}[x]$ for the integer polynomial appearing in (4.3), if $u \in \mathbb{Z}$, then

$$\frac{f(b+m(u+qv))-f(b)}{Qq} - \frac{g(u)}{q} = \frac{g(u+qv)-g(u)}{q} \in \mathbb{Z}[v].$$

This implies that

$$e_{Qq}(f(b+m(u+qv))) = e_{Qq}(f(b+mu)) \qquad (v \in \mathbb{Z}).$$

Hence, for $n \leqslant P$,

$$S_n := \sum_{\substack{p \leqslant n \\ p \equiv b \pmod m}} e_{Qq}(af(p)) = O(mq) + \sum_{\substack{t \pmod{mq} \\ (t,q)=1 \\ t \equiv b \pmod m}} e_{Qq}(af(t)) \sum_{\substack{p \leqslant n \\ p \equiv t \pmod{mq}}} 1.$$

By the Siegel–Walfisz theorem (theorem 2.1), the inner sum is

$$\sum_{\substack{p \leqslant n \\ p \equiv t \pmod{mq}}} 1 = \frac{\mathrm{Li}(n)}{\varphi(mq)} + O(Pe^{-3c\sqrt{\log P}}),$$

whence

$$S_n = \frac{\mathrm{Li}(n)}{\varphi(mq)}S(q,a;m) + O(Pe^{-2c\sqrt{\log P}}).$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Writing $\psi(t) = e_Q(\beta f(t))G(t)\log t$, summation by parts gives

$$\sum_{\substack{p\leqslant P\\ p\equiv b \pmod m}} e_Q\left(F(p)\right)G(p)\log p = \sum_{n\leqslant P}(S_n - S_{n-1})\psi(n)$$
$$= S_P\psi(P+1) + \sum_{n\leqslant P} S_n(\psi(n)-\psi(n+1)).$$

By hypothesis, for P sufficiently large,

$$|\beta| = |\theta - a/q| \ll \frac{Q(\log P)^{2C_k}}{qP^k} \ll \frac{(\log P)^{2C_k+1}}{P^k}.$$

Hence, for all $x, y \in [2, P]$ with $x < y$, the mean value theorem yields

$$\left|\frac{\psi(y)-\psi(x)}{y-x}\right| \leqslant \sup_{t\in[x,y]}\{|G(t)/t| + |G'(t)\log t| + |\beta f'(t)G(t)\log t|\}$$
$$\ll_f \|G\|_{L^\infty[2,P]}\left(\frac{1}{x} + \frac{(\log P)^{2(C_k+1)}}{P}\right) + \|G'\|_{L^\infty[2,P]}\log P. \tag{4.4}$$

In particular, this shows that

$$\sum_{n\leqslant P}|\psi(n)-\psi(n+1)| \ll \|G\|_{L^\infty[2,P]}\left((\log P)^{2(C_k+1)} + \sum_{n\leqslant P}\frac{1}{n}\right)$$
$$+ \|G'\|_{L^\infty[2,P]}P\log P$$
$$\ll (\log P)^{2(C_k+1)}\|G\|_{\mathcal{S}[2,P]}.$$

As $\mathrm{Li}(t) = \sum_{n=3}^{t}\int_{n-1}^{n}\frac{\mathrm{d}x}{\log x}$, summation by parts now gives

$$\sum_{\substack{p\leqslant P\\ p\equiv b \pmod m}} e_Q\left(F(p)\right)G(p)\log p + O(Pe^{-c\sqrt{\log P}}\|G\|_{\mathcal{S}[2,P]})$$
$$= \frac{S(q,a;m)}{\varphi(mq)}\left(\mathrm{Li}(P)\psi(P+1) + \sum_{n\leqslant P}\mathrm{Li}(n)(\psi(n)-\psi(n+1))\right)$$
$$= \frac{S(q,a;m)}{\varphi(mq)}\sum_{3\leqslant n\leqslant P}\int_{n-1}^{n}\frac{\psi(n)}{\log x}\,\mathrm{d}x.$$

Note that

$$\sum_{3\leqslant n\leqslant P}\int_{n-1}^{n}\frac{\mathrm{d}x}{(n-1)\log x} \ll \sum_{3\leqslant n\leqslant P}\int_{n-1}^{n}\frac{\mathrm{d}x}{x\log x} = \log\log P - \log\log 2.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Thus, using (4.4) to replace each $\psi(n)$ by $\psi(x)$, we obtain

$$\sum_{3\leqslant n\leqslant P}\int_{n-1}^{n}\frac{\psi(n)}{\log x}\,\mathrm{d}x = I_{f,G}(\beta)+O((\log P)^{2(C_k+1)}\|G\|_{\mathcal{S}[2,P]}),$$

which completes the proof. □

5. Fourier decay

Returning to the study of our weight function ν, we need to show that it is suitably 'pseudorandom'. This will then allow us to 'transfer' solutions from the linearized equation (2.7) to our original equation (2.6). As in [2, 4, 5, 17], we accomplish this via a *Fourier decay* estimate (together with the restriction estimates from the next section).

LEMMA 5.1. *Let ν be as defined in (3.2), where $b\in[W\kappa]$ satisfies (3.1), and assume that X is sufficiently large in terms of w. Then, for all $\alpha\in\mathbb{T}$,*

$$|\hat{\nu}(\alpha)-\widehat{1_{[N]}}(\alpha)|\ll_{h,\varepsilon} w^{\varepsilon-1/d}N. \tag{5.1}$$

REMARK 5.2. As in [2, §5], the above lemma does not rely upon nor make any reference to sets $A\subseteq\mathcal{P}_X$ or $\mathcal{A}\subseteq[N]$.

We study the Fourier transform $\hat{\nu}$ using the Hardy–Littlewood circle method and the exponential sum estimates established previously. We define the set of *minor arcs*

$$\mathfrak{m}:=\left\{\alpha\in\mathbb{T}:\max\{q,X^d\|q\alpha\|\}>(\log X)^{2C_d}\text{for all }q\in\mathbb{N}\right\}.$$

The set of *major arcs* $\mathfrak{M}:=\mathbb{T}\setminus\mathfrak{m}$, therefore, consists of all $\alpha\in\mathbb{T}$ for which there exist $a,q\in\mathbb{Z}$ such that

$$1\leqslant q\leqslant(\log X)^{2C_d},\qquad (q,a)=1,\qquad X^d|q\alpha-a|\leqslant(\log X)^{2C_d}. \tag{5.2}$$

For convenience, we recall that

$$\hat{\nu}(\alpha)=\frac{\varphi(W)}{W(h'(b),W)}\sum_{\substack{b<p\leqslant X\\ p\equiv b\ (\mathrm{mod}\,W\kappa)}}h'(p)\log p\cdot e\left(\alpha\frac{h(p)-h(b)}{\lambda(D)}\right)\quad(\alpha\in\mathbb{T}).$$

We, therefore, observe that the results of the previous section may be applied to estimate $\hat{\nu}(\alpha)$ upon taking

$$\theta=\alpha,\quad f(y)=h(y)-h(b),\quad G=h',\quad Q=\lambda(D),\quad m=W\kappa,\quad P=X. \tag{5.3}$$

With this choice of parameters, we compute that

$$\|G\|_{\mathcal{S}[2,P]}\ll_h X^{d-1} \tag{5.4}$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

and

$$I_{f,G}(\beta) = \int_2^X e(\beta h(x)/\lambda(D))h'(x)\,\mathrm{d}x = \lambda(D)\left(O_h(1) + \int_0^N e(\beta y)\,\mathrm{d}y\right). \quad (5.5)$$

Our proof of lemma 5.1 for $\alpha \in \mathfrak{m}$ proceeds by the same strategy as in [4, §4]: we show that $\hat{\nu}(\alpha)$ and $\hat{1}_{[N]}(\alpha)$ are both far smaller than the required upper bound. This is encapsulated in the following corollary of lemma 4.2.

Corollary 5.3. (Minor arc estimate). *If $\alpha \in \mathfrak{m}$, then*

$$\hat{\nu}(\alpha) \ll_h X^d(\log X)^{-\sigma_d} \qquad \text{and} \qquad \widehat{1_{[N]}}(\alpha) \ll X^d(\log X)^{-2C_d}.$$

Proof. In view of (5.3) and (5.4), the first estimate follows immediately from lemma 4.2. For the second estimate, we deduce from the definition of $\mathfrak{m}$ that

$$\widehat{1_{[N]}}(\alpha) = \sum_{n=1}^{N} e(\alpha n) \ll \|\alpha\|^{-1} \leqslant X^d(\log X)^{-2C_d} \qquad (\alpha \in \mathfrak{m}),$$

as required. □

We similarly establish an asymptotic formula for $\hat{\nu}(\alpha)$ on the major arcs by invoking lemma 4.3. Define

$$S(q,a) := \sum_{\substack{t \ (\mathrm{mod}\, W\kappa q)\\ (t,q)=1\\ t\equiv b \ (\mathrm{mod}\, W\kappa)}} e_q\left(a\frac{h(t)-h(b)}{\lambda(D)}\right), \qquad I(\beta) := \int_0^N e(\beta y)\,\mathrm{d}y.$$

Corollary 5.4. (Major arc asymptotic). *Suppose $(\alpha, q, a) \in \mathbb{R}\times\mathbb{N}\times\mathbb{Z}$ with (5.2), and put $\beta = \alpha - a/q$. Let $c > 0$ be constant, small in terms of C_d. Then,*

$$\hat{\nu}(\alpha) = \frac{\varphi(W\kappa)}{\varphi(W\kappa q)}S(q,a)I(\beta) + O_h(Ne^{-c\sqrt{\log X}}).$$

Proof. Recall that $\lambda(D)N = h(X) \asymp_h X^d$. Thus, by combining (5.3) with (5.4) and (5.5), the desired formula is provided by lemma 4.3. □

To elucidate this formula further, we estimate $S(q,a)$.

Lemma 5.5. *Let $a, q \in \mathbb{Z}$ with $q \geqslant 2$ and $(q,a) = 1$. Then,*

$$S(q,a) \ll_{h,\varepsilon} \min\left\{q^{\varepsilon-1/d}\varphi(q),\ \frac{\varphi(W\kappa q)}{\varphi(W\kappa)}w^{\varepsilon-1/d}\right\}.$$

Furthermore, if $(q,W) > 1$, then $S(q,a) = 0$.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Proof. Write $q = q_1 q_2$, where $q_1 \in \mathbb{N}$ is w-smooth and q_2 is w-rough. Observe that

$$S(q,a) = \sum_{\substack{x \pmod q \\ (W\kappa x + b, q) = 1}} e_q(ag(x)),$$

where

$$g(x) = \frac{h(W\kappa x + b) - h(b)}{\lambda(D)} = \frac{h(W\kappa x + b) - h(b)}{W\kappa(h'(b), W\kappa)} \in \mathbb{Z}[x]$$

by Taylor's theorem. As $(q_1, q_2) = 1$, a standard calculation reveals that

$$S(q,a) = S(q_1, A_1) S(q_2, A_2),$$

where

$$\frac{a}{q} = \frac{A_1}{q_1} + \frac{A_2}{q_2},$$

as is noted in the proof of [19, lemma 9]. Observe that $(A_1, q_1) = (A_2, q_2) = 1$.

As q_1 is w-smooth and $(b, W) = 1$, we always have $(W\kappa x + b, q_1) = 1$. Let

$$H = (q_1, W), \qquad q_1 = Hq_1', \qquad W = HW',$$

so that $(q_1', W') = 1$. Writing $x = y + q_1' z$ and

$$g(x) = v_d x^d + \cdots + v_1 x$$

gives

$$S(q_1, A_1) = \sum_{y \pmod{q_1'}} \sum_{z \pmod H} e_{Hq_1'}(A_1 \sum_{j \leqslant d} v_j (y + q_1' z)^j).$$

As

$$(h'(b), W) \mid V = \sqrt{W}, \qquad W \mid \kappa(h'(b), W),$$

we must have $(h'(b), W) \mid V \mid \kappa$. Thus, for $2 \leqslant j \leqslant d$, we have

$$v_j = \frac{h^{(j)}(b)(W\kappa)^{j-1}}{j!(h'(b), W)} \equiv 0 \pmod W.$$

Now

$$S(q_1, A_1) = \sum_{y \leqslant q_1'} e_{q_1}(A_1 g(y)) \sum_{z \leqslant H} e_H(A_1 v_1 z), \qquad v_1 = \frac{h'(b)}{(h'(b), W)}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

For each prime $p \leqslant w$, we have $\mathrm{ord}_p(h'(b)) < \mathrm{ord}_p(W)$, and so $\mathrm{ord}_p(v_1) = 0$. Therefore, $(v_1, W) = 1$, so $(H, A_1 v_1) = 1$, whence

$$S(q_1, A_1) = \begin{cases} 1, & \text{if } q_1 = 1 \\ 0, & \text{if } q_1 \geqslant 2. \end{cases}$$

This completes the proof of the assertion that $S(q, a) = 0$ whenever $(q, W) > 1$.

In view of this result, we may henceforth assume that $q_1 = 1$ and $q_2 = q \geqslant 2$. In particular, we have $q > w$. Let us denote by ℓ_h the leading coefficient of h. Then

$$v_d = \frac{\ell_h (W\kappa)^{d-1}}{(h'(b), W)}, \qquad (q, W) = 1,$$

so $(v_d, q) \ll_h 1$. Now lemma 2.2 provides a constant $C = C(d) > 1$ such that

$$S(q, a) = S(q, A_2) \ll_h C^{\omega(q)} q^{1-1/d} \ll_{d,\varepsilon} q^{\varepsilon - 1/d} \varphi(q) < \frac{\varphi(W\kappa q)}{\varphi(W\kappa)} w^{\varepsilon - 1/d},$$

as required. □

These two results allow us to establish our Fourier decay estimate.

Proof of lemma 5.1. Corollary 5.3 gives (5.1) for all $\alpha \in \mathfrak{m}$. We henceforth assume that $\alpha \in \mathfrak{M}$. We begin by considering the case where (5.2) holds with $q = 1$. As demonstrated in [4, § 4], Euler–Maclaurin summation delivers the bound

$$\widehat{1_{[N]}}(\alpha) - I(\alpha) \ll_h (\log X)^{2C_d}.$$

Applying the triangle inequality and corollary 5.4, therefore, gives

$$|\hat{\nu}(\alpha) - \widehat{1_{[N]}}(\alpha)| \leqslant |\hat{\nu}(\alpha) - I(\alpha)| + |\widehat{1_{[N]}}(\alpha) - I(\alpha)| \ll_h N e^{-c\sqrt{\log X}}.$$

Finally, suppose that (5.2) holds with $q \geqslant 2$. We note from [4, equation (4.3)] that

$$\widehat{1_{[N]}}(\alpha) \ll q \leqslant (\log X)^{2C_d}.$$

Thus, in view of the trivial estimate $|I(\beta)| \leqslant N$, the desired result follows by combining corollary 5.4 with lemma 5.5. □

Before moving on, we record the following consequence of corollary 5.4 and lemma 5.5, which we use in the next section.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

COROLLARY 5.6. (Major arc estimate). *Let $\alpha \in \mathbb{T}$ and $q \in \mathbb{N}$. If (5.2) holds for some $a \in \mathbb{Z}$, then*

$$\hat{\nu}(\alpha) \ll_{h,\varepsilon} q^{\varepsilon-1/d} \min\{N, \|\alpha - a/q\|^{-1}\} + O(Ne^{-c\sqrt{\log X}}).$$

Proof. Integrating by parts delivers the standard estimate

$$I(\beta) \ll \min\{N, \|\beta\|^{-1}\} \qquad (\beta \in \mathbb{T}). \tag{5.6}$$

Incorporating the elementary inequality $\varphi(W\kappa)\varphi(q) \leqslant \varphi(W\kappa q)$ and lemma 5.5 delivers

$$\frac{\varphi(W\kappa)}{\varphi(W\kappa q)} S(q,a) I(\alpha - a/q) \ll_{h,\varepsilon} q^{\varepsilon-1/d} \min\{N, \|\alpha - a/q\|^{-1}\}.$$

The required result now follows from corollary 5.4. □

6. Restriction estimates

Recall from §3 that ν is supported on the set

$$\left\{ n \in [N] : n = \frac{h(p) - h(b)}{\lambda(D)},\ p \in \mathcal{P}_X,\ p \equiv b \ (\mathrm{mod}\, W\kappa) \right\}.$$

After linearizing, we wish to solve (2.7) with the n_i drawn from a dense subset $\mathcal{A}$ of the above set. This leads us to the study of functions $\phi : \mathbb{Z} \to \mathbb{C}$, such as the indicator function of $1_{\mathcal{A}}$, which are majorized by ν, meaning that $|\phi| \leqslant \nu$.

The purpose of this section is to establish two restriction estimates. These will then be used in the next section to execute the transference argument. The first restriction estimate is for the weight ν and is needed to transfer between 'dense variables' x_i and n_i in equations (2.6) and (2.7), respectively. Our second restriction estimate concerns an auxiliary weight function ν_D and is required for interpolation.

Throughout this section, we define ν as in §3.2 for a fixed choice of $b \in [W\kappa]$ satisfying (3.1). We also let $T = T(d)$ be as in §2.

6.1. Restriction I

We begin with the following restriction estimate for ν.

PROPOSITION 6.1. *Let $E > 2T$, and let $\phi : \mathbb{Z} \to \mathbb{C}$ with $|\phi| \leqslant \nu$. Then,*

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^E \, \mathrm{d}\alpha \ll_{h,E} N^{E-1}.$$

This is easily bootstrapped to the following restriction estimate for $\nu + 1_{[N]}$, see the deduction of [2, lemma 6.2].

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

PROPOSITION 6.2. *Let $E > 2T$, and let $\phi : \mathbb{Z} \to \mathbb{C}$ with $|\phi| \leqslant \nu + 1_{[N]}$. Then,*

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^E \, \mathrm{d}\alpha \ll_{h,E} N^{E-1}.$$

To prove proposition 6.1, we proceed in stages. We introduce the auxiliary function

$$\mu : \mathbb{Z} \to [0, \infty), \qquad \mu(n) = \frac{1}{(h'(b), W)} \sum_{\substack{b < x \leqslant X \\ x \equiv b \pmod{W\kappa} \\ h(x) - h(b) = n\lambda(D)}} h'(x),$$

noting that $\nu \leqslant (\log X)\mu$ pointwise. We compute that

$$\begin{aligned} \|\mu\|_1 &= (h'(b), W)^{-1} \sum_{\substack{b < x \leqslant X \\ x \equiv b \pmod{W\kappa}}} h'(x) \\ &\ll (h'(b), W)^{-1} \sum_{y \leqslant X/W\kappa} (W\kappa y)^{d-1} \\ &\ll \frac{X^d}{W\kappa(h'(b), W)} = \frac{X^d}{\lambda(D)} \ll N. \end{aligned}$$

We begin with an epsilon-slack restriction estimate for μ.

LEMMA 6.3. (Epsilon-slack estimate). *Let $\psi : \mathbb{Z} \to \mathbb{C}$ with $|\psi| \leqslant \mu$. Then,*

$$\int_{\mathbb{T}} |\hat{\psi}(\alpha)|^{2T} \, \mathrm{d}\alpha \ll_{h,\varepsilon} N^{2T-1+\varepsilon}.$$

Proof. By orthogonality,

$$\int_{\mathbb{T}} |\hat{\psi}(\alpha)|^{2T} \, \mathrm{d}\alpha = \sum_{n_1 + \cdots + n_T = n_{T+1} + \cdots + n_{2T}} \psi(n_1) \cdots \psi(n_T) \overline{\psi(n_{T+1}) \cdots \psi(n_{2T})}.$$

As $\|\psi\|_\infty \leqslant \|\mu\|_\infty \ll X^{d-1}$, and since ψ is supported on

$$\left\{ \frac{h(x) - h(b)}{\lambda(D)} : x \in [X] \right\},$$

we obtain

$$\begin{aligned} &\int_{\mathbb{T}} |\hat{\psi}(\alpha)|^{2T} \, \mathrm{d}\alpha \\ &\ll (X^{d-1})^{2T} \#\{\mathbf{x} \in [X]^{2T} : h(x_1) + \cdots + h(x_T) = h(x_{T+1}) + \cdots + h(x_{2T})\}. \end{aligned}$$

Finally, using that $T = T(d)$, we find that

$$\int_{\mathbb{T}} |\hat{\psi}(\alpha)|^{2T} \, \mathrm{d}\alpha \ll X^{2T(d-1)+2T-d+\varepsilon} \ll N^{2T-1+2\varepsilon}.$$

□

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Our next goal is to largely remove ε from the exponent in this restriction estimate for μ, obtaining a log-slack estimate by passing to a slightly higher moment. To accomplish this, we require some bounds on

$$\hat{\mu}(\theta) = \frac{1}{(h'(b), W)} \sum_{\substack{b<x\leqslant X \\ x\equiv b \ (\mathrm{mod}\, W\kappa)}} h'(x) e\left(\theta \frac{h(x)-h(b)}{\lambda(D)}\right).$$

The triangle inequality and partial summation yield

$$\hat{\mu}(\theta) \ll \frac{X^{d-1}}{(h'(b), W)} \left(1 + \max_{X^{1/2}\leqslant P\leqslant X} |g(\theta; P)|\right),$$

where

$$g(\theta; P) = \frac{1}{(h'(b), W)} \sum_{\substack{b<x\leqslant X+b \\ x\equiv b \ (\mathrm{mod}\, W\kappa)}} e\left(\theta \frac{h(x)-h(b)}{\lambda(D)}\right).$$

Writing $x = W\kappa y + b$ gives

$$\sum_{\substack{x\leqslant X \\ x\equiv b \ (\mathrm{mod}\, W\kappa)}} e\left(\theta \frac{h(x)-h(b)}{\lambda(D)}\right) = \sum_{y\leqslant X/(W\kappa)} e\left(\frac{c_d (W\kappa)^d}{\lambda(D)} \theta f_d(y)\right),$$

where c_d is the leading coefficient of h and $f_d(y) \in \mathbb{Z}[y]$ is monic of degree d.

Suppose $X^{1/2} \leqslant P \leqslant X$. The exponential sum

$$g_1(\alpha; P) := \sum_{y\leqslant P/(W\kappa)} e(\alpha f_d(y))$$

can be treated using Roger Baker's estimates [1]. We apply the formulation [3, lemma 2.3], noting from its proof that the quantity $\sigma(d)$ therein can be replaced by 2^{1-d}. This delivers the following conclusion.

Lemma 6.4. *If $|g_1(\alpha; P)| > (P/(W\kappa))^{1-2^{1-d}+\varepsilon}$, then there exist coprime $r \in \mathbb{N}$ and $b \in \mathbb{Z}$ such that*

$$g_1(\alpha; P) \ll_{d,\varepsilon} r^{\varepsilon-1/d} P(W\kappa)^{-1} (1 + (P/(W\kappa))^d |\alpha - b/r|)^{-1/d}.$$

Lemma 6.5. *Let*

$$\mathfrak{n} = \{\theta \in \mathbb{T} : |\hat{\mu}(\theta)| \leqslant X^{d-2^{-d}}\}.$$

If $\theta \in \mathbb{T} \setminus \mathfrak{n}$, then there exist coprime $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ such that

$$\hat{\mu}(\theta) \ll_{d,\varepsilon} N(\log X) q^{\varepsilon-1/d} (1 + N|\theta - a/q|)^{-1/d}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Proof. Let $P \in [X^{1/2}, X]$ maximize

$$\left| g_1\left(\frac{c_d(W\kappa)^d}{\lambda(D)}\theta; P\right)\right|.$$

By (6.1),

$$\left| g_1\left(\frac{c_d(W\kappa)^d}{\lambda(D)}\theta; P\right)\right| > (P/(W\kappa))^{1-2^{1-d}+\varepsilon}.$$

By lemma 6.4, there exist coprime $r \in \mathbb{N}$ and $b \in \mathbb{Z}$ such that

$$\begin{aligned} g_1\left(\frac{c_d(W\kappa)^d}{\lambda(D)}\theta; P\right) &\ll r^{\varepsilon-1/d}P(W\kappa)^{-1}\left(1+(P/(W\kappa))^d\left|\frac{c_d(W\kappa)^d}{\lambda(D)}\theta - b/r\right|\right)^{-1/d} \\ &\ll r^{\varepsilon-1/d}X(W\kappa)^{-1}\left(1+(X/(W\kappa))^d\left|\frac{c_d(W\kappa)^d}{\lambda(D)}\theta - b/r\right|\right)^{-1/d}. \end{aligned}$$

With

$$a = \frac{b}{(b, c_d(W\kappa)^d/\lambda(D))}, \qquad q = \frac{rc_d(W\kappa)^d/\lambda(D)}{(b, c_d(W\kappa)^d/\lambda(D))},$$

we now have

$$\begin{aligned} g_1\left(\frac{c_d(W\kappa)^d}{\lambda(D)}\theta; P\right) &\ll r^{\varepsilon-1/d}\frac{X}{W\kappa}\left(1+\frac{X^d}{\lambda(D)}|\theta - a/q|\right)^{-1/d} \\ &\ll q^{\varepsilon-1/d}\frac{X\log X}{W\kappa}\left(1+\frac{X^d}{\lambda(D)}|\theta - a/q|\right)^{-1/d}, \end{aligned}$$

since X is large in terms of w. Finally, recall that

$$N \asymp \frac{X^d}{\lambda(D)} = \frac{X^d}{W\kappa(h'(b), W)}$$

and

$$X^{d-2^{-d}} < |\hat{\mu}(\theta)| \ll X^{d-1} + \frac{X^{d-1}}{(h'(b), W)}g_1\left(\frac{c_d(W\kappa)^d}{\lambda(D)}\theta; P\right).$$

□

By passing to a slightly higher moment, we are now able to obtain a log-slack analogue of lemma 6.3.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

LEMMA 6.6. (Log-slack estimate). *Let $v > 2T$ be real, and let $\psi : \mathbb{Z} \to \mathbb{C}$ with $|\psi| \leqslant \mu$. Then,*

$$\int_{\mathbb{T}} |\hat{\psi}(\alpha)|^v \, \mathrm{d}\alpha \ll_{h,v} N^{v-1} (\log X)^v.$$

Proof. Inserting lemmas 6.3 and 6.5 into lemma 2.3 gives

$$\int_{\mathbb{T}} (|\hat{\psi}(\alpha)|/\log X)^v \, \mathrm{d}\alpha \ll N^{v-1}.$$

□

Proof of proposition 6.1. Let $v \in (2T, E)$. Applying lemma 6.6 to $\psi = (\log X)^{-1}\phi$ gives the log-slack restriction estimate

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^v \, \mathrm{d}\alpha \ll_{h,v} N^{v-1} (\log X)^{2v}.$$

We can choose σ_d to be large in terms of E and v. Thus, in view of corollaries 5.3 and 5.6, the desired result follows from lemma 2.3. □

6.2. Restriction II

Define the auxiliary weight function $\nu_D : \mathbb{Z} \to [0, \infty)$ by

$$\begin{aligned} \nu_D(n) &= \frac{\varphi(D)}{\lambda(D)} \sum_{\substack{p \leqslant X \\ p \equiv r_D \pmod D \\ h(p) = n\lambda(D)}} h'(p) \log p \\ &= \frac{\varphi(D)}{\lambda(D)} \sum_{\substack{z \leqslant Z \\ (Dz + r_D) \in \mathcal{P}_X \\ h_D(z) = n}} h'(Dz + r_D) \log(Dz + r_D). \end{aligned}$$

Observe that ν_D is supported on $h_D([Z]) \subseteq [N]$. By the Siegel–Walfisz theorem, we have

$$\begin{aligned} \|\nu_D\|_1 &= \frac{\varphi(D)}{\lambda(D)} \sum_{\substack{p \leqslant X \\ p \equiv r_D \pmod D}} h'(p) \log p \\ &\leqslant \frac{\varphi(D)}{\lambda(D)} \sum_{\substack{p \leqslant X \\ p \equiv r_D \pmod D}} h'(X) \log X \ll N. \end{aligned}$$

One can be more precise using partial summation, but we do not need to.

In this subsection, we establish the following restriction estimate for ν_D.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Proposition 6.7. *Let $E > 2T$ be real, and let $\phi : \mathbb{Z} \to \mathbb{C}$ with $|\phi| \leqslant \nu_D$. Then,*

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^E \, \mathrm{d}\alpha \ll_{h,E} N^{E-1}.$$

Our approach is similar to that of proposition 6.1, so we will not repeat all of the details. We introduce the auxiliary function

$$\mu_D : \mathbb{Z} \to [0, \infty), \qquad \mu_D(n) = \frac{ND}{X} \sum_{\substack{z \leqslant Z \\ h_D(z) = n}} 1,$$

noting that $\nu_D \ll (\log X)\mu_D$ pointwise. As a special case of [2, lemma 6.3], we have the following sharp restriction estimate for μ_D.

Lemma 6.8. *Let $E > 2T$, and let $\psi : \mathbb{Z} \to \mathbb{C}$ with $|\psi| \leqslant \mu_D$. Then,*

$$\int_{\mathbb{T}} |\hat{\psi}(\alpha)|^E \, \mathrm{d}\alpha \ll_{h,E} N^{E-1}.$$

The upshot is that if $v > 2T$ and $|\phi| \leqslant \nu_D$, then

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^v \, \mathrm{d}\alpha \ll_v N^{v-1} (\log X)^v. \tag{6.1}$$

To apply the general epsilon-removal lemma, we require major and minor arc bounds for

$$\hat{\nu}_D(\alpha) = \frac{\varphi(D)}{\lambda(D)} \sum_{\substack{p \leqslant X \\ p \equiv r_D \pmod D}} h'(p) \log p \cdot e(\alpha h(p)/\lambda(D)).$$

On the minor arcs, we infer the following analogue of corollary 5.3, by essentially the same argument:

$$\hat{\nu}_D(\alpha) \ll X^d (\log X)^{-\sigma_d} \qquad (\alpha \in \mathfrak{m}). \tag{6.2}$$

On the major arcs, we have (5.2), for some $q, a \in \mathbb{Z}$. Define

$$S_D(q, a) = \sum_{\substack{t \pmod{Dq} \\ (t,q)=1 \\ t \equiv r_D \pmod D}} e_q(ah(t)/\lambda(D)).$$

We infer the following analogue of corollary 5.4 by essentially the same proof.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

LEMMA 6.9. *Let c be a small, positive constant, small in terms of C_d. Suppose $(\alpha, q, a) \in \mathbb{R} \times \mathbb{N} \times \mathbb{Z}$ with (5.2), and put $\beta = \alpha - a/q$. Then,*

$$\hat{\nu}_D(\alpha) = \frac{\varphi(D)}{\varphi(Dq)} S_D(q,a) I(\beta) + O(Ne^{-c\sqrt{\log X}}).$$

It follows from lemma 2.2 that

$$S_D(q,a) \ll_{h,\varepsilon} q^{\varepsilon - 1/d} \varphi(q).$$

Thus, by incorporating (5.6), we arrive at the following variant of corollary 5.6:

$$\hat{\nu}_D(\alpha) \ll_{h,\varepsilon} q^{\varepsilon-1/d} \min\{N, \|\alpha - a/q\|^{-1}\} + O(Ne^{-c\sqrt{\log X}}). \tag{6.3}$$

Equipped with the bounds (6.1), (6.2), and (6.3), proposition 6.7 now follows from the general epsilon-removal lemma (lemma 2.3).

7. The transference principle

In this section, we are finally ready to use transference to deduce theorem 2.6 from theorem 2.8. We start with some notation and a preparatory lemma.

For finitely-supported $f_1, \dots, f_s, g_1, \dots, g_t : \mathbb{Z} \to \mathbb{R}$, define

$$\Phi(f_1, \dots, f_s; g_1, \dots, g_t) = \sum_{L_1(\mathbf{n}) = L_2(\mathbf{m})} f_1(n_1) \cdots f_s(n_s) g_1(m_1) \cdots g_t(m_t).$$

We frequently make use of the abbreviations

$$\Phi(f_1, \dots, f_s; g) = \Phi(f_1, \dots, f_s; g, \dots, g), \qquad \Phi(f; g) = \Phi(f, \dots, f; g, \dots, g).$$

Given finite sets of integers A and B, we also write $\Phi(A; g) = \Phi(1_A; g)$, and similarly for the expressions $\Phi(f; B)$ and $\Phi(A; B)$.

LEMMA 7.1. (Fourier control). *Let $f_1, \dots, f_s, g : \mathbb{Z} \to \mathbb{R}$ be finitely supported. If*

$$|f_j| \leqslant \nu + 1_{[N]} \quad (1 \leqslant j \leqslant s), \qquad |g| \leqslant \nu_D,$$

then

$$\Phi(f_1, \dots, f_s; g) \ll N^{s+t-1} \prod_{j \leqslant s} (\|\hat{f}_j\|_\infty / N)^{1/(2s+2t)}.$$

Proof. Following the proof of [2, lemma 7.1] yields

$$|\Phi(f_1, \dots, f_s; g)| \leqslant \left(\int_{\mathbb{T}} |\hat{g}(\alpha)|^{s+t} \, \mathrm{d}\alpha \right)^{t/(s+t)} \cdot \prod_{j \leqslant s} \left(\|\hat{f}_j\|_\infty^{1/2} \int_{\mathbb{T}} |\hat{f}_j(\alpha)|^{s+t-1/2} \, \mathrm{d}\alpha \right)^{1/(s+t)}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Propositions 6.2 and 6.7 now give

$$\Phi(f_1,\dots,f_s;g) \ll N^{(s+t-1)t/(s+t)} \prod_{j\leqslant s} \left(\|\hat{f}_j\|_\infty^{1/(2s+2t)} N^{(s+t-3/2)/(s+t)} \right)$$
$$= N^{s+t-1} \prod_{j\leqslant s} (\|\hat{f}_j\|_\infty/N)^{1/(2s+2t)}.$$

□

Proof of theorem 2.6 given theorem 2.8. Fix δ, h, r, L_1, and L_2. The implied constants are henceforth allowed to depend on all of these parameters. Let $\tilde{\delta}$ be small in terms of the fixed parameters, and let $w \in \mathbb{N}$ be large in terms of them. We insist that $\tilde{\delta}$ is an integer power of 2, and that $\tilde{\delta} \gg 1$, so that dependence on $\tilde{\delta}$ is subsumed by dependence on the fixed parameters.

Given sufficiently large $X \in \mathbb{N}$, we define D, N, W, Z as in §3. For the purposes of proving theorem 2.6, we may assume that $Z \in \mathbb{N}$. Indeed, assuming X is sufficiently large relative to D, any $A \subseteq \mathcal{P}_X$ with $|A| \geqslant \delta|\mathcal{P}_X|$ must satisfy $|A \cap [X-D]| > (\delta/2)|\mathcal{P}_X|$. Thus, by replacing (δ, A, X) with $(\delta/2, A \cap [X-a], X-a)$ for some $a \in [D]$ such that $D \mid (X - a - r_D)$, we can assume that $D \mid (X - r_D)$, whence $Z \in \mathbb{N}$.

Set

$$\tilde{\mathcal{C}}_i = \{z \in [Z] : r_D + Dz \in \mathcal{C}_i\} \qquad (1 \leqslant j \leqslant r).$$

By theorem 2.8, there exists $k \in [r]$ such that every $\tilde{\mathcal{A}} \subseteq [N]$ with $|\tilde{\mathcal{A}}| \geqslant \tilde{\delta}N$ satisfies

$$\#\{(\mathbf{n},\mathbf{z}) \in \tilde{\mathcal{A}}^s \times \tilde{\mathcal{C}}_k^t : L_1(\mathbf{n}) = L_2(h_D(\mathbf{z}))\} \gg N^{s-1} \left(\frac{DZ}{\varphi(D)\log Z} \right)^t.$$

By a simple counting argument, the number of solutions counted here for any given value of z_t is $O(|\tilde{C}_k| N^{s-1}(Z/\log Z)^{t-1})$, whence $|\tilde{\mathcal{C}}_k| \gg Z/\log Z$. Thus,

$$|\mathcal{C}_k| \geqslant |\tilde{\mathcal{C}}_k| \gg \frac{Z}{\log Z} \gg_w \frac{X}{\log X}.$$

Define $\mathcal{A}$ by (3.3), and define

$$f = \nu 1_{\mathcal{A}}, \qquad g_i(n) = \frac{N\varphi(D)\log Z}{DZ} \sum_{\substack{z \in \tilde{C}_i \\ h_D(z)=n}} 1 \quad (1 \leqslant i \leqslant r).$$

By lemma 5.1 and the dense model lemma [16, theorem 5.1], there exists a function f_0 such that

$$0 \leqslant f_0 \leqslant 1_{[N]}, \qquad \|\hat{f} - \hat{f}_0\|_\infty \ll (\log w)^{-3/2} N.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

For $\ell \in [s]$, we write $\mathbf{u}^{(\ell)} = (u_1^{(\ell)}, \dots, u_s^{(\ell)})$, where

$$u_j^{(\ell)} = \begin{cases} f_0, & \text{if } j < \ell \\ f - f_0, & \text{if } j = \ell \\ f, & \text{if } j > \ell. \end{cases}$$

The telescoping identity and lemma 7.1 now give

$$\Phi(f; g_i) - \Phi(f_0; g_i) = \sum_{\ell \leqslant s} \Phi(\mathbf{u}^{(\ell)}; g_i) \ll (\log w)^{-3/(4s+4t)} N^{s+t-1} \quad (1 \leqslant i \leqslant r). \tag{7.1}$$

By lemma 3.1, we have

$$\sum_{n \in \mathbb{Z}} f(n) \gg N.$$

Since $\hat{f}(0) - \hat{f}_0(0) \ll (\log w)^{-3/2}$, and w is large, we also have

$$\sum_{n \in \mathbb{Z}} f_0(n) \gg N.$$

Let c be a small, positive constant that depends only on the fixed parameters. Setting

$$\tilde{\mathcal{A}} = \{n \in \mathbb{Z} : f_0(n) \geqslant c\} \subseteq [N],$$

we observe that

$$N \ll \sum_{n \in \tilde{\mathcal{A}}} f_0(n) + \sum_{n \in [N] \setminus \tilde{\mathcal{A}}} f_0(n) \leqslant |\tilde{\mathcal{A}}| + cN.$$

Taking c sufficiently small, therefore, allows us to extract the lower bound $|\tilde{\mathcal{A}}| \geqslant \tilde{\delta} N$. Now theorem 2.8 gives $\Phi(\tilde{\mathcal{A}}; g_k) \gg N^{s+t-1}$. Since $0 \leqslant c1_{\tilde{A}} \leqslant f_0$, it follows that $\Phi(f_0; g_k) \gg N^{s+t-1}$. Taking w sufficiently large, we infer from (7.1) that

$$\Phi(f; g_k) \gg N^{s+t-1}.$$

Finally, since $N \leqslant h(X) \asymp X^d$, we conclude that

$$\begin{aligned} &\#\{(\mathbf{x}, \mathbf{y}) \in A^s \times \mathcal{C}_k^t : L_1(h(\mathbf{x})) = L_2(h(\mathbf{y}))\} \\ &\geqslant \|f\|_\infty^{-s} \|g_k\|_\infty^{-t} \Phi(f; g_k) \gg_w \left(\frac{N \log X}{X}\right)^{-s} \left(\frac{N \log X}{X}\right)^{-t} N^{s+t-1} \\ &\gg_w \frac{X^{s+t-d}}{(\log X)^{s+t}}. \end{aligned}$$

By specifying that $w = O_{\delta,h,r,L_1,L_2}(1)$, this completes the proof. □

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

8. Arithmetic regularity

Our only remaining task is to prove theorem 2.8. We are, therefore, interested in counting solutions to the linearized equation (2.7). We seek a colour class $\mathcal{C}_k$ such that there are many solutions $(\mathbf{n}, \mathbf{z})$ to (2.7) with $\mathbf{n} \in \mathcal{C}_k^t$ and $\mathbf{z} \in \mathcal{A}^s$ for some arbitrary dense set $\mathcal{A} \subseteq [N]$.

Following [17] and [2], we begin by weakening the statement of theorem 2.8. Rather than assert the existence of a colour class $\mathcal{C}_k$ which gives many solutions with respect to all dense sets $\mathcal{A}$, we instead consider a finite collection of dense sets $\mathcal{A}_1, \dots, \mathcal{A}_r \subseteq [N]$ and seek a colour class $\mathcal{C}_k$ such that, for all $i \in [r]$, there are many solutions to (2.7) with $(\mathbf{n}, \mathbf{z}) \in \mathcal{A}_i^s \times \mathcal{C}_k^t$. This leads to the following version of theorem 2.8.

THEOREM 8.1 *Let r and $d \geqslant 2$ be positive integers, and let $0 < \delta \leqslant 1$. Let h be an integer polynomial of degree d, which is intersective of the second kind. Let $s \geqslant 1$ and $t \geqslant 0$ be integers such that $s + t \geqslant s_0(d)$. Let*

$$L_1(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_s], \qquad L_2(\mathbf{y}) \in \mathbb{Z}[y_1, \dots, y_t]$$

be non-degenerate linear forms such that $L_1(1, \dots, 1) = 0$ and $\gcd(L_1) = 1$. Then, there exists $\eta = \eta(d, \delta, L_1, L_2) \in (0, 1)$ such that the following is true. Let $D, Z \in \mathbb{N}$ satisfy $Z \geqslant Z_0(D, h, r, \delta, L_1, L_2)$, and set $N := h_D(Z)$. Suppose $\mathcal{A}_1, \dots, \mathcal{A}_r \subseteq [N]$ satisfy $|\mathcal{A}_i| \geqslant \delta N$ for all $i \in [r]$. If

$$[\eta Z, Z] \cap \{z \in [Z] : r_D + Dz \in \mathcal{P}\} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_r,$$

then there exists $k \in [r]$ such that

$$\#\{(\mathbf{n}, \mathbf{z}) \in \mathcal{A}_i^s \times \mathcal{C}_k^t : L_1(\mathbf{n}) = L_2(h_D(\mathbf{z}))\} \gg N^{s-1} \left(\frac{DZ}{\varphi(D) \log Z}\right)^t \qquad (1 \leqslant i \leqslant r).$$

The implied constant may depend on h, L_1, L_2, r, δ.

Proof of theorem 2.8 given theorem 8.1. In view of proposition 2.10, it is enough to prove theorem 2.8 under the assumption that $\gcd(L_1) = 1$. We claim that theorem 2.8 holds with the same quantities $Z_0(D, h, r, \delta, L_1, L_2)$, $\eta(d, \delta, L_1, L_2)$, and the same implicit constant $C = C(h, L_1, L_2, r, \delta)$ appearing in the final bound. Suppose for a contradiction that this is false. For each $k \in [r]$, we can then find $\mathcal{A}_k \subseteq [N]$ with $|\mathcal{A}_k| \geqslant \delta N$ such that

$$\#\{(\mathbf{n}, \mathbf{z}) \in \mathcal{A}_k^s \times \mathcal{C}_k^t : L_1(\mathbf{n}) = L_2(h_D(\mathbf{z}))\} < CN^{s-1} \left(\frac{DZ}{\varphi(D) \log Z}\right)^t.$$

Applying theorem 8.1 to the collection of dense sets $\mathcal{A}_1, \dots, \mathcal{A}_r$ delivers a contradiction. □

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

By taking Z sufficiently large relative to η in theorem 8.1, we may assume that h_D is positive and strictly increasing on the real interval $[\eta Z, Z]$. We can then define a function $\mathcal{Q}_Z = \mathcal{Q}_{Z;\eta,h_D} : \mathbb{Z} \to \mathbb{R}$ by

$$\mathcal{Q}_Z(t) := \begin{cases} z, & \text{if } z \in [\eta Z, Z] \text{ satisfies } t = h_D(z) \\ 0, & \text{otherwise.} \end{cases}$$

Notice that $\mathcal{Q}_Z$ is supported on $h_D([\eta Z, Z]) \subseteq [N]$, and the restriction of $\mathcal{Q}_Z$ to $h_D([\eta Z, Z])$ defines a bijection from $h_D([\eta Z, Z])$ to $[\eta Z, Z]$. Hence, given functions $f_1, \dots, f_s : \mathbb{Z} \to \mathbb{R}$ supported on $[N]$, and $g : \mathbb{Z} \to \mathbb{R}$ supported on $[\eta Z, Z]$, we have

$$\Phi(f_1, \dots, f_s; g \circ \mathcal{Q}_Z) = \sum_{L_1(\mathbf{n}) = L_2(h_D(\mathbf{z}))} f_1(n_1) \cdots f_s(n_s) g(z_1) \cdots g(z_t).$$

LEMMA 8.2. (Arithmetic regularity lemma). *Let $r \in \mathbb{N}$, $\sigma > 0$, and let $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ be a monotone increasing function. Then, there exists a positive integer $K_0(r; \sigma, \mathcal{F}) \in \mathbb{N}$ such that the following is true. Let $N \in \mathbb{N}$ and $f_1, \dots, f_r : [N] \to [0, 1]$. Then, there is a positive integer $K \leqslant K_0(r; \sigma, \mathcal{F})$ and a phase $\boldsymbol{\theta} \in \mathbb{T}^K$ such that, for every $i \in [r]$, there is a decomposition*

$$f_i = f_{\mathrm{str}}^{(i)} + f_{\mathrm{sml}}^{(i)} + f_{\mathrm{unf}}^{(i)}$$

of f_i into functions $f_{\mathrm{str}}^{(i)}, f_{\mathrm{sml}}^{(i)}, f_{\mathrm{unf}}^{(i)} : [N] \to [-1, 1]$ with the following stipulations.

(I) The functions $f_{\mathrm{str}}^{(i)}$ and $f_{\mathrm{str}}^{(i)} + f_{\mathrm{sml}}^{(i)}$ take values in $[0, 1]$.
(II) The function $f_{\mathrm{sml}}^{(i)}$ obeys the bound $\|f_{\mathrm{sml}}^{(i)}\|_{L^2(\mathbb{Z})} \leqslant \sigma \|1_{[N]}\|_{L^2(\mathbb{Z})}$.
(III) The function $f_{\mathrm{unf}}^{(i)}$ obeys the bound $\|\hat{f}_{\mathrm{unf}}^{(i)}\|_\infty \leqslant \|\hat{1}_{[N]}\|_\infty / \mathcal{F}(K)$.
(IV) The function $f_{\mathrm{str}}^{(i)}$ satisfies $\sum_{m=1}^{N} (f_i - f_{\mathrm{str}}^{(i)})(m) = 0$.
(V) There exists a K-Lipschitz function $F_i : \mathbb{T}^K \to [0, 1]$ such that $F_i(x\boldsymbol{\theta}) = f_{\mathrm{str}}^{(i)}(x)$ for all $x \in [N]$.

Proof. This is [2, lemma 8.3]. □

Applying this to a given function f allows us to write $\Phi(f; g)$ as the sum of $\Phi(f_{\mathrm{str}} + f_{\mathrm{sml}}; g)$ and $2^s - 1$ terms $\Phi(f_1, \dots, f_s; g)$, where at least one of the f_i equals f_{unf} and the rest are equal to $f_{\mathrm{str}} + f_{\mathrm{sml}}$. As is typical in applications of the arithmetic regularity lemma, we expect the term $\Phi(f_{\mathrm{str}} + f_{\mathrm{sml}}; g)$ to provide the main contribution, while the remaining terms should be asymptotically negligible. This prediction is verified by combining Property (III) of lemma 8.2 with our Fourier control result (lemma 7.1).

To carry out this strategy of removing the contribution of f_{unf}, we need to perform a minor technical manoeuvre. Applying lemma 7.1 requires us to bound the function g appearing in $\Phi(f; g)$ in terms of ν_D. To achieve a strong asymptotic lower bound for the number of solutions, we desire a bound of the form $g \|\nu_D\|_\infty \ll \nu_D$. This is the method we used in [2, lemma 8.4], only with μ_D in place of ν_D. However,

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

this relied on the fact that μ_D is constant on its support, while ν_D is not. In particular, if g is the indicator function of a colour class, then $g(z)\|\nu_D\|_\infty$ could be asymptotically larger than $\nu_D(z)$ for small z.

To overcome this issue, we restrict attention from $[Z]$ to $[\eta Z, Z]$, for some sufficiently small $\eta > 0$. On this latter interval, the function ν_D does not vary too much. This is made precise by the following lemma, which is a variation of [2, lemma 8.12].

LEMMA 8.3. *Let P be a real polynomial of degree $d \in \mathbb{N}$ with positive leading coefficient. Then, there exists a positive integer $M_0(P)$ such that the following is true. For all $\eta \in (0,1)$, if $x \in \mathbb{R}$ satisfies $x \geqslant \eta^{-1}M_0(P)$, then*

$$\eta^d P(x) \leqslant 3P(\eta x) \leqslant 9\eta^d P(x).$$

Proof. Let $\ell_P > 0$ be the leading coefficient of P. We can then find $M_0(P) \in \mathbb{N}$ such that

$$\ell_P x^d \leqslant 2P(x) \leqslant 3\ell_P x^d$$

holds for all real $x \geqslant M_0(P)$. In particular, if $x \geqslant \eta^{-1}M_0(P)$, then

$$\frac{\eta^d}{3} \leqslant \frac{P(\eta x)}{P(x)} \leqslant 3\eta^d.$$

□

LEMMA 8.4. (Removing f_{unf}). *Let $f : \mathbb{Z} \to [0,1]$ be supported on $[N]$. Let $\eta, \sigma > 0$, and let $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ be a monotone increasing function. Let $f_{\mathrm{str}}, f_{\mathrm{sml}}, f_{\mathrm{unf}}$ be the functions obtained upon applying lemma 8.2 to f. Then, for any $g : \mathbb{Z} \to [0,1]$ supported on the set*

$$\{z \in [Z] : r_D + Dz \in \mathcal{P}\} \cap [\eta Z, Z],$$

we have

$$|\Phi(f; g \circ \mathcal{Q}_Z) - \Phi(f_{\mathrm{str}} + f_{\mathrm{sml}}; g \circ \mathcal{Q}_Z)| \ll_{h,\eta,D} N^{s-1}\left(\frac{DZ}{\varphi(D)\log Z}\right)^t \mathcal{F}(K)^{-1/(2s+2t)}.$$

Proof. Note that $|f| \leqslant 1_{[N]}$ and $\|\hat{f}\|_\infty \leqslant N$. Thus, by using a telescoping identity, as in the derivation of (7.1), lemma 7.1 informs us that

$$|\Phi(f; G) - \Phi(f_{\mathrm{str}} + f_{\mathrm{sml}}; G)| \ll N^{s+t-1}\mathcal{F}(K)^{-1/(2s+2t)}$$

holds for any $G : \mathbb{Z} \to \mathbb{R}$ such that $|G| \leqslant \nu_D$. Taking $G = \xi(g \circ \mathcal{Q}_Z)$ for some $\xi > 0$, we deduce that

$$|\Phi(f; g \circ \mathcal{Q}_Z) - \Phi(f_{\mathrm{str}} + f_{\mathrm{sml}}; g \circ \mathcal{Q}_Z)| \ll N^{s-1}(N/\xi)^t \mathcal{F}(K)^{-1/(2s+2t)}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

To complete the proof, it remains to find $\xi > 0$ with $\xi|g \circ \mathcal{Q}_Z| \leqslant \nu_D$ such that

$$N/\xi \ll \frac{DZ}{\varphi(D)\log Z}.$$

Set

$$B = \{n \in h_D([\eta Z, Z] \cap \mathbb{N}) : n = h_D(z), \quad r_D + Dz \in \mathcal{P}\}.$$

Observe that, for all $n = h_D(z) \in B$, we have

$$\lambda(D)\nu_D(n) \geqslant \varphi(D)h'(D\eta Z + r_D)\log(D\eta Z + r_D).$$

By lemma 8.3, if Z is sufficiently large relative to h, η, and D, then

$$DZh'(D\eta Z + r_D)\log(D\eta Z + r_D) \gg h(X)\log Z = \lambda(D)N\log Z.$$

Since $g \circ \mathcal{Q}_Z$ is supported on B and takes values in $[0,1]$, we conclude that there exists $c \gg 1$ such that $\xi := c(DZ)^{-1}\varphi(D)N\log Z$ has all the required properties. □

9. Prime polynomial Bohr sets

The final step of the proof of theorem 8.1 is to obtain a lower bound for the main term $\Phi(f_{\mathrm{str}} + f_{\mathrm{sml}}; g \circ \mathcal{Q}_Z)$. Using our assumption that the coefficients of L_1 are coprime, Bézout's lemma provides us with some $\mathbf{v} \in \mathbb{Z}^s$, which depends only on L_1, such that $L_1(\mathbf{v}) = 1$. We can therefore write

$$\Phi(f_1, \dots, f_s; g \circ \mathcal{Q}_Z) = \sum_{\mathbf{z} \in [\eta Z, Z]^t} g(z_1)\cdots g(z_t)\Psi_{\mathbf{z}}(f_1, \dots, f_s), \tag{9.1}$$

where we have introduced the auxiliary counting operator

$$\Psi_{\mathbf{z}}(f_1, \dots, f_s) := \sum_{L_1(\mathbf{n})=0} \prod_{i=1}^{s} f_i(n_i + v_i L_2(h_D(\mathbf{z}))).$$

Our goal is to show that there is a large supply of $\mathbf{z} \in \mathbb{Z}^t$ for which $\Psi_{\mathbf{z}}(f_{\mathrm{str}} + f_{\mathrm{sml}})$ is asymptotically as large as possible. This is accomplished by choosing the z_i to lie in a set of 'almost-periods' for f_{str}. These are known as *polynomial Bohr sets* and take the form

$$\{n \in \mathbb{N} : \|Q(n)\boldsymbol{\alpha}\| < \rho\} = \bigcap_{i=1}^{K} \{n \in \mathbb{N} : \|Q(n)\alpha_i\| < \rho\},$$

for some $\rho > 0$, $K \in \mathbb{N}$, $\boldsymbol{\alpha} \in \mathbb{T}^K$, and $Q(x) \in \mathbb{Z}[x]$.

LEMMA 9.1. (Lower bound for $\Psi_{\mathbf{z}}(f_{\mathrm{str}} + f_{\mathrm{sml}})$). *For all $\delta > 0$, there exist positive constants $c_1(\delta) = c_1(L_1, L_2; \delta)$ and $\eta_0 = \eta_0(d, L_1, L_2, \delta) < 1$ such that the following is true. Suppose $f : \mathbb{Z} \to [0,1]$ is supported on $[N]$ and satisfies $\|f\|_1 \geqslant \delta N$. Given*

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

$\sigma \in (0,1]$ and a monotone increasing function $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, let f_{str}, f_{sml}, K, and $\boldsymbol{\theta}$ be as given by applying lemma 8.2 to f. If $\mathbf{z} \in [\eta_0 Z]^t$ satisfies

$$\|h_D(z_i)\theta_j\| < \sigma/K \qquad (1 \leqslant i \leqslant t, \quad 1 \leqslant j \leqslant K),$$

then

$$\Psi_{\mathbf{z}}(f_{\mathrm{str}} + f_{\mathrm{sml}}) \geqslant (c_1(\delta) - O_{L_1,L_2}(\sigma))\, N^{s-1}.$$

In particular, if σ is sufficiently small relative to (d, L_1, L_2, δ), then

$$\Psi_{\mathbf{z}}(f_{\mathrm{str}} + f_{\mathrm{sml}}) \gg_{L_1,L_2,\delta} N^{s-1}.$$

Proof. This is [2, lemma 8.13] with $\rho = \sigma/K$. □

Recall that we seek solutions to the linearized equation (2.7) with $r_D + Dz_j$ prime for all j. In view of lemma 9.1, we are therefore interested in sets of the form

$$\mathcal{B}_h(\boldsymbol{\alpha}, \rho) := \{p \in \mathcal{P} : p \equiv r_D \pmod D,\ \|\boldsymbol{\alpha} h(p)/\lambda(D)\| < \rho\}. \tag{9.2}$$

The main purpose of this section is to establish the following density bounds for these *prime polynomial Bohr sets*.

THEOREM 9.2 *Let $K, D, d \in \mathbb{N}$, $0 < \rho \leqslant 1$, and let h be an integer polynomial of degree d which is intersective of the second kind. Then, there exists a positive integer $P_1 = P_1(D, h, K, \rho)$ and a positive real number $\Delta(\rho) = \Delta(h, K; \rho) \leqslant 1$ such that the following is true for all $P \geqslant P_1$. If $\boldsymbol{\alpha} \in \mathbb{T}^K$, then*

$$\sum_{p \in \mathcal{B}} \log p \geqslant \frac{\Delta(\rho) P}{\varphi(D)},$$

where $\mathcal{B} = [P] \cap \mathcal{B}_h(\boldsymbol{\alpha}, \rho)$. Moreover, we may take

$$\Delta(h, 1; \rho) \gg_{h,\varepsilon} \rho^{d+3+\varepsilon},$$
$$\Delta(h, K; \rho) \gg_{h,K,\varepsilon} \rho^{3+K(d+\varepsilon)} \Delta\left(h, K-1; \frac{\rho^2}{2K^2}\right) \quad (K > 1).$$

We demonstrate the utility of this result by using it to complete the proof of theorem 8.1.

Proof of theorem 8.1 given theorem 9.2. As usual, we fix the parameters

$$\delta, h, r, L_1, L_2$$

appearing in the statement of theorem 8.1. Unless specified otherwise, we allow all forthcoming implicit constants to depend implicitly on these parameters. Let $\sigma, \eta \in (0,1)$ and let $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ be a monotone increasing function, all three of which depend only on the fixed parameters. Let $D \in \mathbb{N}$, and assume throughout this proof that $Z \in \mathbb{N}$ and $N := h_D(Z)$ are sufficiently large with respect to all of these quantities.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

For each $i \in [r]$, let $\mathcal{A}_i \subseteq [N]$ with $|\mathcal{A}_i| \geqslant \delta N$. Suppose we have an r-colouring

$$\{z \in [\eta Z, Z] : r_D + Dz \in \mathcal{P}\} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r.$$

In the notation of the previous section, our goal is to find $k \in \mathbb{N}$ such that

$$\Phi(1_{\mathcal{A}_i}; 1_{\mathcal{C}_k} \circ \mathcal{Q}_Z) \gg N^{s-1} \left(\frac{DZ}{\varphi(D)\log Z}\right)^t \qquad (1 \leqslant i \leqslant r).$$

Lemma 8.2 provides us with decompositions

$$1_{\mathcal{A}_i} = f_{\mathrm{str}}^{(i)} + f_{\mathrm{sml}}^{(i)} + f_{\mathrm{unf}}^{(i)} \qquad (1 \leqslant i \leqslant r),$$

along with a positive integer $K \ll_{\sigma,\mathcal{F}} 1$ and a phase $\boldsymbol{\theta} \in \mathbb{T}^K$ with the properties described therein. Let $\eta_0 = \eta_0(d, L_1, L_2, \delta)$ be as defined in lemma 9.1, and let

$$\Omega := \{z \in [\eta Z, \eta_0 Z] : r_D + Dz \in \mathcal{P},\ \|h_D(z)\boldsymbol{\theta}\| < \sigma/K\} \subseteq \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r.$$

By choosing σ sufficiently small, lemma 9.1 and (9.1) inform us that

$$\Phi(f_{\mathrm{str}}^{(i)} + f_{\mathrm{sml}}^{(i)}; 1_{\mathcal{C}_j} \circ \mathcal{Q}_Z) \gg N^{s-1}|\Omega \cap \mathcal{C}_j|^t \qquad (1 \leqslant i, j \leqslant r).$$

We now claim that, for an appropriate choice of $\eta < \eta_0$, the set Ω satisfies

$$|\Omega| \gg_K \frac{DZ}{\varphi(D)\log Z}. \tag{9.3}$$

Assume for the moment that this is true. By the pigeonhole principle, we can choose $k \in [r]$ such that $r|\Omega \cap \mathcal{C}_k| \geqslant |\Omega|$, whence

$$\Phi(f_{\mathrm{str}}^{(i)} + f_{\mathrm{sml}}^{(i)}; 1_{\mathcal{C}_k} \circ \mathcal{Q}_Z) \gg_K N^{s-1} \left(\frac{DZ}{\varphi(D)\log Z}\right)^t \qquad (1 \leqslant i \leqslant r).$$

Incorporating lemma 8.4 furnishes the bound

$$\Phi(1_{\mathcal{A}_i}; 1_{\mathcal{C}_k} \circ \mathcal{Q}_Z) \gg N^{s-1} \left(\frac{DZ}{\varphi(D)\log Z}\right)^t \left(c(K) - \mathcal{F}(K)^{-1/(2s+2t)}\right),$$

for some positively-valued function $c(K) > 0$ whose value depends only on the fixed parameters and K. Specifying $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ to be a monotone increasing function which obeys

$$2\mathcal{F}(y)^{-1/(2s+2t)} \leqslant c(y) \qquad (y \in \mathbb{N})$$

then finishes the proof of theorem 8.1, subject to our claim.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

It remains to establish (9.3). Let

$$P = \eta_0 DZ + r_D.$$

Let $\rho = \sigma/K$, and for each $\xi \in (0,1)$ put

$$\begin{aligned}\mathcal{D}_\xi &= \mathcal{P} \cap \{p \in [\xi P, P] : p \equiv r_D \pmod D,\ \|\boldsymbol{\theta} h(p)/\lambda(D)\| < \rho\} \\ &= [\xi P, P] \cap \mathcal{B}_h(\boldsymbol{\theta}, \rho).\end{aligned}$$

By theorem 9.2 and the Siegel-Walfisz theorem, there exists $\xi \in (0,1)$ with $\xi \gg \Delta(\rho)$ such that

$$\sum_{p \in \mathcal{D}_\xi} \log p \geqslant \frac{\Delta(\rho) P}{2\varphi(D)} \gg_\rho \frac{P}{\varphi(D)}.$$

Choosing $\eta = \eta_0 \xi/2$, we ensure that the injective function $y \mapsto (y - r_D)/D$ maps $[\xi P, P]$ into $[\eta Z, \eta_0 Z] \subseteq [\eta Z, Z]$ and maps $\mathcal{D}_\xi$ into Ω. Since $\rho = O_K(1)$, we therefore conclude that

$$|\Omega| \geqslant \sum_{p \in \mathcal{D}_\xi} \frac{\log p}{\log P} \gg_K \frac{DZ}{\varphi(D) \log Z},$$

as claimed. □

9.1. Exponential sums

To study prime polynomial Bohr sets, we are interested in exponential sums over primes of the form

$$\sum_{\substack{p \leqslant P \\ p \equiv r_D \pmod D}} e\left(\frac{h(p)\theta}{\lambda(D)}\right) \log p,$$

where $\theta \in \mathbb{T}$ and h is intersective of the second kind. Lê and Spencer [14] analysed properties of sums of this form to obtain estimates for the smallest element of a prime polynomial Bohr set (9.2), showing in particular that these sets are always non-empty. Our goal is to obtain a lower bound for the densities of these Bohr sets which does not depend on the choice of phase $\boldsymbol{\alpha}$.

Following [14], our argument begins with the observation that if the prime polynomial Bohr set (9.2) has few elements, then we can construct a corresponding exponential sum which is particularly large. This is elucidated by the following lemma, which is a consequence of a much more general result of Harman [8].

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

LEMMA 9.3. *Let $D, K, P \in \mathbb{N}$. Let h be an integer polynomial of degree $d \in \mathbb{N}$ which is intersective of the second kind. Define r_D and $\lambda(D)$ as in §2.4. Let $\rho \in (0,1)$, $\boldsymbol{\alpha} \in \mathbb{T}^K$, and*

$$\mathcal{C} = \mathcal{C}_h(\boldsymbol{\alpha}, \rho) := \mathcal{P} \cap \left\{ p \leqslant P : p \equiv r_D \pmod{D}, \left\| \frac{h(p)}{\lambda(D)} \boldsymbol{\alpha} \right\| \geqslant \rho \right\}.$$

Then, there exists $\mathbf{m} \in \mathbb{Z}^K$ with $0 < \|\mathbf{m}\|_\infty \leqslant K\rho^{-1}$ such that

$$(2K+1)^K \left| \sum_{p \in \mathcal{C}} e\left(\frac{h(p)\mathbf{m} \cdot \boldsymbol{\alpha}}{\lambda(D)} \right) \log p \right| \geqslant \frac{\rho^K}{4K^2 - 1} \sum_{p \in \mathcal{C}} \log p.$$

Proof. If $\rho > 1/2$, then $\mathcal{C}$ is empty and both sides of the desired inequality equal zero. If $\rho \leqslant 1/2$, then the result follows from the contrapositive of [8, corollary to lemma 5] and the pigeonhole principle. □

For the purpose of proving theorem 9.2, we may assume that the Bohr set $\mathcal{B}_h(\boldsymbol{\alpha}, \rho)$ has too few elements, in a manner that will be clarified in due course. Then we can apply lemma 9.3 to find some $L \ll_K \rho^{-K}$ and $\mathbf{m} \in \mathbb{Z}^K$ of bounded size such that $\theta = \mathbf{m} \cdot \boldsymbol{\alpha}$ satisfies

$$\left| \sum_{\substack{p \leqslant P \\ p \equiv r_D \pmod{D}}} e\left(\frac{h(p)\theta}{\lambda(D)} \right) \log p \right| \geqslant \frac{P}{L\varphi(D)}. \tag{9.4}$$

Our next task is to investigate the consequences of (9.4). As discussed in §4, an exponential sum being large is indicative of the phase θ exhibiting 'major arc' behaviour, meaning that θ is well-approximated by a rational number with small denominator. This is made precise in the following lemma.

LEMMA 9.4. (Low major arc). *Suppose $L, P \in \mathbb{N}$ and $\theta \in \mathbb{R}$ satisfy (9.4). Assume that P is sufficiently large relative to D, h and L. Then, there exists $q \in \mathbb{N}$ such that*

$$q \ll_{h,\varepsilon} L^{d+\varepsilon}, \qquad \|q\theta\| \ll_h qL^d\lambda(D)/P^d.$$

Proof. By (9.4) and lemma 4.2, we can find $r \in \mathbb{N}$ and $b \in \mathbb{Z}$ such that

$$\max\left\{ r, P^d \left| r\frac{\theta}{\lambda(D)} - b \right| \right\} \ll (\log P)^{2C_d}.$$

Let $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ with

$$(a,q) = 1, \qquad \frac{\lambda(D)b}{r} = \frac{a}{q}.$$

Then

$$q \leqslant r \ll (\log P)^{2C_d}, \qquad |q\theta - a| \leqslant |r\theta - \lambda(D)b| \ll \frac{\lambda(D)(\log P)^{2C_d}}{P^d}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Put $\beta = \theta - a/q$. Applying lemma 4.3 with $(f, G, b, m, Q) = (h, 1, r_D, D, \lambda(D))$ gives

$$\sum_{\substack{p \leqslant P \\ p \equiv r_D \pmod D}} e\left(\frac{\theta h(p)}{\lambda(D)}\right) \log p = I(\beta)\frac{S(q,a;D)}{\varphi(Dq)} + O_h(Pe^{-c\sqrt{\log P}}),$$

where

$$I(\beta) = \int_2^P e\left(\frac{\beta h(t)}{\lambda(D)}\right) \mathrm{d}t, \qquad S(q,a;D) = \sum_{\substack{t \pmod{Dq} \\ (t,q)=1 \\ t \equiv r_D \pmod D}} e\left(\frac{ah(t)}{q\lambda(D)}\right).$$

Thus, by (9.4), we have

$$|S(q,a;D)I(\beta)| \gg \frac{P\varphi(Dq)}{L\varphi(D)}. \tag{9.5}$$

In light of the trivial bound $|I(\beta)| \leqslant P$, we now have

$$|S(q,a;D)| \gg \frac{\varphi(Dq)}{L\varphi(D)} \geqslant \frac{\varphi(q)}{L}.$$

By [15, lemma 28], the GCD of the non-constant coefficients of

$$h_D(x) := \frac{h(r_D + Dx)}{\lambda(D)} \in \mathbb{Z}[x]$$

is $O_h(1)$. Now lemma 2.2 yields

$$S(q,a;D) \ll_{h,\varepsilon} q^{1+\varepsilon-1/d},$$

and we conclude that

$$q \ll_{h,\varepsilon} L^{d+\varepsilon}.$$

It remains to bound $\|q\theta\|$.

Observe that

$$\begin{aligned} \psi : (\mathbb{Z}/Dq\mathbb{Z})^\times &\to (\mathbb{Z}/D\mathbb{Z})^\times \\ [t] &\mapsto [t] \end{aligned}$$

defines a surjective group homomorphism, and that $\psi^{-1}(r_D)$ is a coset of $\ker(\psi) \leqslant (\mathbb{Z}/Dq\mathbb{Z})^\times$. Therefore,

$$|S(q,a;D)| \leqslant |\psi^{-1}(r_D)| = |\ker(\psi)| = \frac{\varphi(Dq)}{\varphi(D)}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Pairing this with (9.5) yields

$$|I(\beta)| \gg P/L.$$

A change of variables gives

$$I(\beta) = D\int_0^{P/D} e(\beta h_D(z))\,\mathrm{d}z + O(D),$$

and now [22, theorem 7.3] yields

$$I(\beta) \ll_h (|\beta|/\lambda(D))^{-1/d} + D.$$

Consequently,

$$\beta \ll (L/P)^d\lambda(D),$$

and finally,

$$\|q\theta\| \ll qL^d\lambda(D)/P^d.$$

□

9.2. Proof of theorem 9.2

Let $K \in \mathbb{N}$. Writing $\mathcal{B} = \mathcal{B}_h(\boldsymbol{\alpha}, \rho)$ and $\mathcal{C} = \mathcal{C}_h(\boldsymbol{\alpha}, \rho)$, we deduce from the Siegel–Walfisz theorem (theorem 2.1) that

$$\sum_{p\in\mathcal{B}} \log p + \sum_{p\in\mathcal{C}} \log p = \sum_{\substack{p\leqslant P\\ p\equiv r_D \pmod D}} \log p \asymp \frac{P}{\varphi(D)}.$$

Suppose $\boldsymbol{\alpha} \in \mathbb{T}^K$ is such that

$$\sum_{p\in\mathcal{B}} \log p \leqslant \frac{c(K)\rho^K P}{\varphi(D)},$$

for some suitably small $c(K) > 0$. If no such $\boldsymbol{\alpha}$ were to exist, then theorem 9.2 would hold with $\Delta(h, K, \rho) \gg_K \rho^K$ which, since we demand that $\Delta(h, K, \rho) \leqslant 1$, is stronger than the bound we require. For this choice of $\boldsymbol{\alpha}$, we infer from lemma 9.3 the existence of some $\mathbf{m} \in \mathbb{Z}^K$ with $0 < \|\mathbf{m}\|_\infty \leqslant K\rho^{-1}$ such that

$$\left|\sum_{p\in\mathcal{C}} e\left(\frac{h(p)\mathbf{m}\cdot\boldsymbol{\alpha}}{\lambda(D)}\right)\log p\right| \gg_K \frac{\rho^K P}{\varphi(D)}.$$

We also have

$$\left|\sum_{p\in\mathcal{B}} e\left(\frac{h(p)\mathbf{m}\cdot\boldsymbol{\alpha}}{\lambda(D)}\right)\log p\right| \leqslant \sum_{p\in\mathcal{B}} \log p \ll \frac{c(K)\rho^K P}{\varphi(D)}.$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

Choosing $c(K)$ sufficiently small, the triangle inequality furnishes some $L \ll_K \rho^{-K}$ such that (9.4) holds with $\theta = \mathbf{m} \cdot \boldsymbol{\alpha}$. By increasing the value of L if necessary—which only weakens the bound (9.4) that we have obtained—we may assume that $L = C\rho^{-K}$ for some suitably large constant $C = C(h, K) > K$ to be specified later. Applying lemma 9.4 supplies us with some $q \in \mathbb{N}$ such that

$$q \ll_{h,\varepsilon} L^{d+\varepsilon}, \qquad \|q\mathbf{m} \cdot \boldsymbol{\alpha}\| \ll \frac{qL^d\lambda(D)}{P^d}. \tag{9.6}$$

To complete the proof, beginning from the above deductions, we proceed by induction on K. First, suppose that $K = 1$ and write $m \in \mathbb{Z}$ in place of $\mathbf{m} \in \mathbb{Z}^K$. As h is intersective of the second kind, any integer $n \equiv r_{qmD} \pmod{qmD}$ satisfies

$$h(n) \equiv 0 \pmod{\lambda(qmD)}, \qquad (n, qmD) = 1.$$

Further, by the Siegel–Walfisz theorem and (9.6), we have

$$\sum_{\substack{p \leqslant P/L^2 \\ p \equiv r_{qmD} \pmod{qmD}}} \log p \asymp \frac{P/L^2}{\varphi(qmD)} \gg_{h,\varepsilon} \frac{P}{L^{d+3+\varepsilon}\varphi(D)}.$$

Since $r_{qmD} \equiv r_D \pmod{D}$, every prime p appearing in the sum on the left is congruent to r_D modulo D. Using (2.5), we have

$$qm\lambda(D) \mid \lambda(qm)\lambda(D) = \lambda(qmD) \mid h(p)$$

for each prime p in the sum, whence

$$\left\| \frac{\alpha h(p)}{\lambda(D)} \right\| \leqslant \left| \frac{h(p)}{qm\lambda(D)} \right| \|qm\alpha\| \ll_h \frac{(P/L^2)^d}{q\lambda(D)} \frac{qL^d\lambda(D)}{P^d} = L^{-d} = (\rho/C)^d.$$

Taking C sufficiently large, we deduce that $p \in \mathcal{B}$. We conclude that

$$\sum_{p \in \mathcal{B}} \log p \geqslant \sum_{\substack{p \leqslant P/L^2 \\ p \equiv r_{qmD} \pmod{qmD}}} \log p \gg_{h,\varepsilon} \frac{P}{L^{d+3+\varepsilon}\varphi(D)} \gg_h \frac{\rho^{d+3+\varepsilon}P}{\varphi(D)},$$

as required.

Now suppose $K \geqslant 2$ and assume the induction hypothesis that theorem 9.2 holds with $K - 1$ in place of K. Write $\mathbf{m} = (\mathbf{m}', m_K)$ and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}', \alpha_K)$. The induction hypothesis, applied to

$$\frac{\lambda(qm_K)}{m_K}\boldsymbol{\alpha}',$$

informs us that the set

$$\mathcal{A} = \left\{ p \leqslant P/L^2 : p \equiv r_{qm_KD} \pmod{qm_KD},\ \left\| \frac{h(p)}{m_K\lambda(D)}\boldsymbol{\alpha}' \right\| < \frac{\rho^2}{2K^2} \right\}$$

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

satisfies

$$\sum_{p\in\mathcal{A}}\log p \geqslant \frac{P}{L^2\varphi(qm_KD)}\Delta\left(h, K-1; \frac{\rho^2}{2K^2}\right)$$
$$\gg_{h,K,\varepsilon} \frac{P}{\varphi(D)}\rho^{3+K(d+\varepsilon)}\Delta\left(h, K-1; \frac{\rho^2}{2K^2}\right).$$

Let $a\in\mathbb{Z}$ be such that

$$\|q\mathbf{m}\cdot\boldsymbol{\alpha}\| = |q\mathbf{m}\cdot\boldsymbol{\alpha} - a|.$$

For each $p\in\mathcal{A}$, we have $h(p)\ll_h (P/L^2)^d$ so, by (9.6),

$$\left|\alpha_K\frac{h(p)}{\lambda(D)} - \frac{a/q-\mathbf{m}'\cdot\boldsymbol{\alpha}'}{m_K}\frac{h(p)}{\lambda(D)}\right| = \left|\frac{h(p)}{m_K\lambda(D)}\right|\cdot|\mathbf{m}\cdot\boldsymbol{\alpha}-a/q| \ll_h L^{-d}.$$

As in the previous case, we have $qm_K\lambda(D)\mid h(p)$, whence

$$\left\|\frac{a/q-\mathbf{m}'\cdot\boldsymbol{\alpha}'}{m_K}\frac{h(p)}{\lambda(D)}\right\| = \left\|\frac{h(p)\mathbf{m}'\cdot\boldsymbol{\alpha}'}{m_K\lambda(D)}\right\| < \frac{\rho}{2}.$$

Thus, by the triangle inequality,

$$\left\|\alpha_K\frac{h(p)}{\lambda(D)}\right\| < \frac{\rho}{2} + O_h(L^{-d}) = \frac{\rho}{2} + O_h((C\rho^{-K})^{-d}).$$

By taking C sufficiently large, we find that $p\in\mathcal{B}$. Therefore

$$\sum_{p\in\mathcal{B}}\log p \geqslant \sum_{p\in\mathcal{A}}\log p \gg_{h,K,\varepsilon} \frac{P}{\varphi(D)}\rho^{3+K(d+\varepsilon)}\Delta\left(h, K-1; \frac{\rho^2}{2K^2}\right),$$

as required.

Acknowledgements

J.C. is supported by the Heilbronn Institute for Mathematical Research. S.C. thanks the University of Bristol for their kind hospitality on various occasions when this work was being discussed. We are grateful to the anonymous referee for their thorough report and numerous constructive comments.

References

[1] R. C. Baker. *Diophantine Inequalities*, London Math. Soc. Monographs (N.S.). Vol.**1**, (Clarendon Press, Oxford, 1986).

[2] J. Chapman and S. Chow. Generalised Rado and Roth criteria. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* 38.

[3] S. Chow. Waring's problem with shifts. *Mathematika*. **62** (2016), 13–46.

[4] S. Chow. Roth–Waring–Goldbach. *Int. Math. Res. Not. IMRN*. **2018** (2018), 2341–2374.

[5] S. Chow, S. Lindqvist and S. Prendiville. Rado's criterion over squares and higher powers. *J. Eur. Math. Soc.* **23** (2021), 1925–1997.

[6] B. J. Green. Roth's theorem in the primes. *Ann. of Math.* **161** (2005), 1609–1636.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press

[7] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math*. **167** (2008), 481–547.
[8] G. Harman. Small fractional parts of additive forms. *Phil. Trans. R. Soc. Lond. A*. **345** (1993), 327–338.
[9] L. -K. Hua. On Waring's problem. *Quart. J. Math. Oxford*. **9** (1938), 199–202.
[10] L. -K. Hua. *Additive Theory of Prime numbers*, Translations of Mathematical Monographs. Vol.**13**, (American Mathematical Society, Providence, R.I, 1965).
[11] T. H. Lê. Partition regularity and the primes. *C. R. Math. Acad. Sci. Paris*. **350** (2012), 439–441.
[12] T. H. Lê. Problems and results on intersective sets. *Combinatorial and Additive Number theory-CANT 2011 and 2012*, Springer Proc. Math. Stat., Vol.**101**, pp. 115–128 (Springer, New York, 2014).
[13] T. H. Lê and C. V. Spencer. Intersective polynomials and diophantine approximation, II. *Monatsh. Math*. **177** (2015), 79–99.
[14] T. H. Lê and C. V. Spencer. Intersective polynomials and diophantine approximation. *Int. Math. Res. Not. IMRN*. **2014** (2014), 1153–1173.
[15] J. Lucier. Intersective sets given by a polynomial. *Acta Arith*. **123** (2006), 57–95.
[16] S. Prendiville. Four variants of the Fourier-analytic transference principle. *Online J. Anal. Comb*. **12** (2017), 25, Paper No. 5.
[17] S. Prendiville. Counting monochromatic solutions to diagonal Diophantine equations. *Discrete Anal*. (2021), 47, Paper No. 14
[18] R. Rado. Studien zur Kombinatorik. *Math. Z*. **36** (1933), 242–280.
[19] A. Rice. Sárközy's theorem for $\mathcal{P}$-intersective polynomials arXiv:1111.6559v5
[20] K. F. Roth. On certain sets of integers (II). *J. London Math. Soc*. **29** (1954), 20–26.
[21] J. Salmensuu. On the Waring–Goldbach problem with almost equal summands. *Mathematika*. **66** (2020), 255–296.
[22] R. C. Vaughan. *The Hardy–Littlewood method*. second edition, (Cambridge University Press, Cambridge, 1997).
[23] T. D. Wooley. Nested efficient congruencing and relatives of Vinogradov's mean value theorem. *Proc. London Math. Soc*. **118** (2019), 942–1016.

https://doi.org/10.1017/prm.2024.96 Published online by Cambridge University Press