Equidistribution from the Chinese Remainder Theorem

Author(s):
Kowalski, Emmanuel; Soundararajan, Kannan

Publication date:
2021-07-16

Permanent link:
https://doi.org/10.3929/ethz-b-000483601

Originally published in:
Advances in Mathematics 385, https://doi.org/10.1016/j.aim.2021.107776

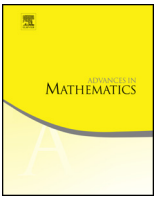
Funding acknowledgement:
175755 - Geometric and Analytic Number Theory (SNF)

Advances in Mathematics 385 (2021) 107776

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim



Equidistribution from the Chinese Remainder Theorem

E. Kowalski [a,*], K. Soundararajan [b]

[a] *ETH Zürich – D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland*
[b] *Department of Mathematics, Stanford University, Stanford, CA 94305, United States of America*

ARTICLE INFO

Article history:
Received 12 June 2020
Received in revised form 1 April 2021
Accepted 16 April 2021
Available online xxxx
Communicated by Kartik Prasanna

Dedicated to the memory of Hédi Daboussi

Keywords:
Chinese Remainder Theorem
Equidistribution
Exponential sums
Roots of congruences

ABSTRACT

We prove the equidistribution of subsets of $(\mathbf{R}/\mathbf{Z})^n$ defined by fractional parts of subsets of $(\mathbf{Z}/q\mathbf{Z})^n$ that are constructed using the Chinese Remainder Theorem.

1. Introduction

Given an irreducible quadratic polynomial $f \in \mathbf{Z}[X]$, the celebrated work of Duke, Friedlander, and Iwaniec [4] (see also Toth [16]) shows that the roots of the congruence $f(x) \equiv 0 \pmod{p}$ become equidistributed when taken over all primes $p \leqslant P$. Precisely,

* Corresponding author.
E-mail addresses: kowalski@math.ethz.ch (E. Kowalski), ksound@stanford.edu (K. Soundararajan).

https://doi.org/10.1016/j.aim.2021.107776

their results establish the equidistribution in $\mathbf{R}/\mathbf{Z}$ of the points x_p/p taken over all $p \leqslant P$ and roots x_p of $f(x_p) \equiv 0 \pmod p$. A similar result is expected for roots of polynomials of higher degree, but this remains an outstanding open problem. In [10], Hooley established that if one considers instead the roots of a polynomial congruence $\pmod n$ over all integer moduli n, then a suitable equidistribution result holds. In this paper we show that Hooley's result may be recast as a general fact concerning the equidistribution of sets arising from the Chinese Remainder Theorem. Our work was partly motivated by the paper [7] of Granville and Kurlberg (who consider the spacing between elements of "large" sets defined by the Chinese Remainder Theorem). Some applications were also suggested by recent work of Hrushovski [11].

For simplicity, we begin by considering equidistribution in $\mathbf{R}/\mathbf{Z}$; later we shall discuss the higher dimensional case of points in $(\mathbf{R}/\mathbf{Z})^n$. Suppose that for each prime power p^v we are given a set A_{p^v} of residue classes modulo p^v (where throughout we include primes among the prime powers, and exclude 1). Let $\varrho(p^v) = |A_{p^v}|$. We allow for the possibility that $\varrho(p^v) = 0$, so that A_{p^v} is empty, for some prime powers p^v, and no assumptions are made concerning the relations between the sets $A_{p^{v_1}}$ and $A_{p^{v_2}}$ corresponding to different powers of the prime p. For a positive integer q, let $A_q \subset \mathbf{Z}/q\mathbf{Z}$ denote the set of residue classes $x \pmod q$ such that $x \pmod{p^v} \in A_{p^v}$ for all prime powers p^v exactly dividing q (that is, $p^v | q$ but $p^{v+1} \nmid q$; we denote this by $p^v \| q$ from now on). These are the "sets defined using the Chinese Remainder Theorem." Let $\varrho(q) = |A_q|$, so that (setting $\varrho(1) = 1$) the function $\varrho(q)$ is multiplicative:

$$\varrho(q) = \prod_{p^v \| q} \varrho(p^v).$$

Let $\mathcal{Q}$ denote the set of all q with $\varrho(q) \geqslant 1$, and for any integer $k \geqslant 1$, let $\mathcal{Q}_k$ denote the elements of $\mathcal{Q}$ with exactly k distinct prime factors. Further, for $x \geqslant 1$, let $\mathcal{Q}(x)$ (resp. $\mathcal{Q}_k(x)$) denote the subset of elements of $\mathcal{Q}$ (resp. of $\mathcal{Q}_k$) that are $\leqslant x$. In order to ensure that the sets $\mathcal{Q}$ and $\mathcal{Q}_k$ are well behaved and have plenty of elements we shall make the following assumption.

Assumption 1.1. There exist constants $\alpha > 0$ and $x_0 \geqslant 2$ such that for all $x \geqslant x_0$

$$\sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 1}} \log p \geqslant \alpha x.$$

Throughout we operate under Assumption 1.1, and the parameter x will be considered to be large in terms of α and x_0, so that for example we would have $\alpha \log\log x \geqslant \sqrt{\log\log x}$.

Given $q \in \mathcal{Q}$, we define a probability measure Δ_q on $\mathbf{R}/\mathbf{Z}$ by

$$\Delta_q = \frac{1}{\varrho(q)} \sum_{a \in A_q} \delta_{\{\frac{a}{q}\}}$$

where δ_t denotes a Dirac mass at the point t, and $\{\cdot\}$ denotes the fractional part of a real number. The limiting behavior of such measures is the object of our study. For example, we are interested in knowing whether Δ_q tends to the uniform measure for most $q \in \mathcal{Q}$. To quantify whether Δ_q is close to uniform, we use the discrepancy

$$\operatorname{disc}(\Delta_q) = \sup_{I \subset \mathbf{R}/\mathbf{Z}} |\Delta_q(I) - |I||,$$

where the supremum is taken over all closed intervals I in $\mathbf{R}/\mathbf{Z}$, and $|I|$ denotes the length of the interval I. By a (closed) interval in $\mathbf{R}/\mathbf{Z}$ we mean the image in $\mathbf{R}/\mathbf{Z}$ of a (closed) interval in $\mathbf{R}$ of length at most 1. One has $0 \leqslant \operatorname{disc}(\Delta_q) \leqslant 1$ for all q, and a small value of $\operatorname{disc}(\Delta_q)$ indicates that Δ_q is close to uniform.

Theorem 1.2. *Suppose that* Assumption 1.1 *holds, and that x is large in terms of α and x_0. Then, there is an absolute constant C such that*

$$\frac{1}{|\mathcal{Q}(x)|} \sum_{q \in \mathcal{Q}(x)} \operatorname{disc}(\Delta_q) \leqslant \frac{C}{\alpha} \exp\Big(- \frac{1}{6} \sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 2}} \frac{1}{p} \Big).$$

Remark 1.3. (1) If we write

$$\sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 2}} \frac{1}{p} = P,$$

then Theorem 1.2 guarantees that apart from at most $C\alpha^{-1}|\mathcal{Q}(x)|e^{-P/12}$ values of q, one has $\operatorname{disc}(\Delta_q) \leqslant e^{-P/12}$. Thus if P is large then for almost all $q \leqslant x$ with $q \in \mathcal{Q}$ one has equidistribution of the sets A_q (by which we mean the equidistribution of the measures Δ_q). Apart from constants, this result is best possible, for we should expect that about $e^{-P}|\mathcal{Q}(x)|$ squarefree elements $q \in \mathcal{Q}(x)$ would be divisible by no prime p with $\varrho(p) \geqslant 2$, and for such q we would have $|A_q| = 1$ and $\operatorname{disc}(\Delta_q) = 1$.

(2) In particular, for almost all $q \in \mathcal{Q}$, the discrepancy bound implies that the smallest element of A_q is $\ll qe^{-P/12}$ (if we identify $\mathbf{Z}/q\mathbf{Z}$ with $\{0, \dots, q-1\}$). In the case of roots of polynomial congruences, such a result was recently proved by Crişan and Pollack [1].

Theorem 1.2 applies to Hooley's result on roots of a polynomial modulo all integers. By the Chebotarev Density Theorem, any irreducible polynomial of degree $d \geqslant 2$ has d roots modulo p for a positive density of primes, so that Assumption 1.1 holds, and further

$$\sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 2}} \frac{1}{p} \geqslant c(d) \log\log x$$

for some constant $c(d) \geqslant \frac{1}{d!}$ (so that the right-hand side of the estimate in Theorem 1.2 is of size $(\log x)^{-c}$ for some $c > 0$). We shall give further applications along these lines in Section 2. Our version is somewhat different from Hooley's, and we shall compare and contrast these in Section 2.2. The generality of Theorem 1.2 indicates that Hooley's equidistribution [10] is a manifestation of the mixing properties of the Chinese Remainder Theorem rather than the arithmetic structure of roots of polynomial congruences.

We shall generalize and strengthen Theorem 1.2 in a few different ways. Firstly, we consider subsets of $(\mathbf{Z}/p^v\mathbf{Z})^n$ for fixed $n \geqslant 1$. Here a key issue is to find the correct generalization of the condition that $\varrho(p) \geqslant 2$ for many primes that arose naturally in the one-dimensional case. Secondly, we shall consider equidistribution of the measures Δ_q when q is restricted to integers in $\mathcal{Q}$ with exactly k distinct prime factors. Under mild hypotheses on $\varrho(p)$, we shall show that in a wide range of k, the discrepancy of the measures Δ_q is typically small. Under more restrictive hypotheses (when $\varrho(p)$ is large for $p \in \mathcal{Q}$) we show that $\mathrm{disc}(\Delta_q)$ is typically small already for numbers with two prime factors.

We begin by introducing the higher dimensional setting, and formulating an analogue of Theorem 1.2. Throughout, the dimension n will be considered fixed, so that implicit constants will be allowed to depend on n, but we shall display the dependencies on all other parameters. For each prime power p^v, let $A_{p^v} \subset (\mathbf{Z}/p^v\mathbf{Z})^n$ be a set of n-tuples of residue classes modulo p^v. As before, we put $\varrho(p^v) = |A_{p^v}|$ and allow A_{p^v} to be the empty set (so that $\varrho(p^v) = 0$) for some prime powers. For a positive integer q, we let $A_q \subset (\mathbf{Z}/q\mathbf{Z})^n$ be the set of residue classes $x \,(\mathrm{mod}\, q)$ such that $x \,(\mathrm{mod}\, p^v) \in A_{p^v}$ for all prime powers $p^v \| q$. Let $\varrho(q)$ denote the size of A_q, which again is a multiplicative function. We let $\mathcal{Q}$, $\mathcal{Q}(x)$, $\mathcal{Q}_k$, and $\mathcal{Q}_k(x)$ have their earlier meanings, and will be working as before under Assumption 1.1.

For $a = (a_1, \dots, a_n) \in \mathbf{R}^n$, we write

$$\{a\} = (\{a_1\}, \dots, \{a_n\}) \in (\mathbf{R}/\mathbf{Z})^n.$$

We define a probability measure Δ_q on $(\mathbf{R}/\mathbf{Z})^n$ by

$$\Delta_q = \frac{1}{\varrho(q)} \sum_{a \in A_q} \delta_{\{\frac{a}{q}\}}.$$

The closeness of Δ_q to the uniform measure is quantified by means of the *box discrepancy*

$$\mathrm{disc}(\Delta_q) = \sup_{B \subset (\mathbf{R}/\mathbf{Z})^n} |\Delta_q(B) - \mathrm{Vol}(B)|$$

where the supremum is taken over all boxes B in $(\mathbf{R}/\mathbf{Z})^n$, and $\mathrm{Vol}(B)$ denotes the usual volume (Lebesgue measure) of the box. Here, by a box in $(\mathbf{R}/\mathbf{Z})^n$, we mean the projection modulo $\mathbf{Z}^n$ of a closed box (that is, a product of closed intervals) in $\mathbf{R}^n$ with all side lengths $\leqslant 1$.

Suppose there is a fixed affine hyperplane H defined over $\mathbf{Z}$ such that the elements in A_{p^v} all lie in the reduction of H modulo p^v for all $p \in \mathcal{Q}$. Then for $q \in \mathcal{Q}$, the elements in A_q would also lie in this hyperplane, so that the measures Δ_q will be supported in a translate of a proper subtorus of $(\mathbf{R}/\mathbf{Z})^n$. This situation prevents equidistribution; it generalizes the case $n = 1$, where an affine hyperplane is a single point, so that concentration in a single hyperplane corresponds to the case when $\varrho(p) \leqslant 1$ for most primes p. Our generalization of Theorem 1.2 establishes that if the sets A_p do not concentrate on hyperplanes for a positive density of primes p, then Δ_q is close to the uniform measure (i.e., has small discrepancy) for most moduli q.

To state this precisely, we need one further definition. Given a prime p in $\mathcal{Q}$, define

$$\lambda(p) = \max_{\substack{H \subset (\mathbf{Z}/p\mathbf{Z})^n \\ H \text{ affine hyperplane}}} |H \cap A_p|,$$

where an affine hyperplane $H \subset (\mathbf{Z}/p\mathbf{Z})^n$ is a subset of the form

$$H = \{x \in (\mathbf{Z}/p\mathbf{Z})^n \mid h_1 x_1 + \cdots + h_n x_n = a\}$$

for some $a \in \mathbf{Z}/p\mathbf{Z}$ and $(h_i) \in (\mathbf{Z}/h\mathbf{Z})^n \setminus \{(0, \dots, 0)\}$.

Theorem 1.4. *Suppose that* Assumption 1.1 *holds, and that x is large in terms of α and x_0. Then, there is a constant $C(n)$ depending only on n such that*

$$\frac{1}{|\mathcal{Q}(x)|} \sum_{q \in \mathcal{Q}(x)} \operatorname{disc}(\Delta_q) \leqslant \frac{C(n)}{\alpha} \exp\Big(-\frac{1}{3} \sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 1}} \Big(1 - \frac{\lambda(p)}{\varrho(p)}\Big) \frac{1}{p} \Big).$$

Remark 1.5. Consider the case $n = 1$. Then we have $\lambda(p) = 1$ whenever $\varrho(p) \geqslant 1$, and thus

$$\sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 1}} \Big(1 - \frac{\lambda(p)}{\varrho(p)}\Big) \frac{1}{p} \geqslant \frac{1}{2} \sum_{\substack{p \leqslant x \\ \varrho(p) \geqslant 2}} \frac{1}{p},$$

and Theorem 1.2 is seen to be a special case of Theorem 1.4.

For any n, given at most n points in $(\mathbf{Z}/p\mathbf{Z})^n$, we may always find an affine hyperplane containing all of them. But given $n + 1$ points we may expect that they are "in general position", in the sense that there is no affine hyperplane that contains all of them. Thus, roughly speaking, Theorem 1.4 says that if there are many primes p with A_p in general position, and containing at least $n + 1$ elements, then for almost all $q \in \mathcal{Q}$, the measures Δ_q are close to equidistribution.

By imposing a stronger (but still mild) hypothesis, we can obtain equidistribution of Δ_q on average, when q is restricted to integers with a given number of prime factors.

Theorem 1.6. *Suppose that* Assumption 1.1 *holds, and that x is large in terms of x_0 and α. Suppose that $0 < \delta \leqslant 1$ is such that*

$$\sum_{\substack{p\leqslant x\\ p\in\Omega}} \Big(1 - \frac{\lambda(p)}{\varrho(p)}\Big)\frac{1}{p} \geqslant \delta \log\log x. \tag{1}$$

Then uniformly in the range

$$\frac{20(7+n)}{\delta}\log\Big(\frac{20(7+n)}{\delta}\Big) \leqslant k \leqslant \exp\Big(\sqrt{\frac{\alpha\delta\log\log x}{20(6+n)}}\Big)$$

we have

$$\frac{1}{|\Omega_k(x)|}\sum_{\substack{q\leqslant x\\ q\in\Omega_k}} \mathrm{disc}(\Delta_q) \ll \frac{1}{\alpha}\Big(e^{-\delta k/18} + (\log x)^{-\alpha\delta/18}\Big).$$

Remark 1.7. (1) If we think of δ as a fixed positive constant, then Theorem 1.6 shows that for most $q \in \Omega_k(x)$ one has equidistribution of Δ_q so long as $k \to \infty$ (arbitrarily slowly with x) and provided $k \leqslant \exp(c\sqrt{\log\log x})$ for some $c > 0$. A condition like $k \to \infty$ is necessary to guarantee that A_q has many points, which is essential for equidistribution.

(2) Although "typical" integers in Ω have on the order of $\log\log x$ prime factors, and larger values of k occur very rarely, it would be interesting to extend the result to larger values of k, especially up to $k \leqslant (\log x)^c$ for some $c > 0$.

Our last result provides equidistribution for Δ_q for most q in Ω_k, for any *fixed* $k \geqslant 2$, provided the sets A_p are known to be large for most $p \in \Omega$.

Theorem 1.8. *Suppose that* Assumption 1.1 *holds, and that x is large in terms of x_0 and α. Let $\delta > 0$ be such that $1/\log\log x \leqslant \delta \leqslant 1/e$ and*

$$\sum_{p\in\Omega(x)} \frac{1}{p}\frac{\lambda(p)}{\varrho(p)} \leqslant \delta \sum_{p\in\Omega(x)} \frac{1}{p}. \tag{2}$$

Then, uniformly in the range $2 \leqslant k \leqslant \alpha\delta\log\log x$,

$$\frac{1}{|\Omega_k(x)|}\sum_{q\in\Omega_k(x)} \mathrm{disc}(\Delta_q) \ll \frac{1}{\alpha}\delta^{(k-1)/10}.$$

The interest in Theorem 1.8 is really for small values of k, since when k is large one may simply use the bounds in Theorem 1.6. If δ in Theorem 1.8 is close to 0, then we get equidistribution for most Δ_q already for integers q with 2 prime factors. For example, this applies, in the case $n = 1$, whenever $\varrho(p)$ tends to infinity for $p \in \Omega$.

The final remark before closing the introduction section is that Assumption 1.1, as well as all the estimates in Theorems 1.2, 1.4, 1.6 and 1.8 only involve the sets A_p and their sizes. In other words, *there is no restriction whatsoever* on the choice of the sets A_{p^v} for $v \geqslant 2$. This should not be surprising because most natural numbers are not divisible by many prime powers p^v with $v \geqslant 2$.

1.1. Outline of the paper

The next section provides a selection of applications of Theorem 1.4, and compares the results with those of [10]. Section 3 discusses some preliminaries, and the proof of Theorem 1.4 (which contains Theorem 1.2 as a special case) is concluded in Section 4. In Section 5 we develop a technical estimate (Proposition 5.1) which is more precise (but more complicated to state) than Theorems 1.6 and 1.8, and in Section 6 we prove them starting from that technical result. Finally, Section 7 discusses briefly another possible generalization of our method, which will be the subject of a later work [13], and an Appendix considers briefly a function field analogue of conjectures about roots of polynomials congruences modulo primes.

2. Examples and counterexamples

In this section, we present some examples of applications of Theorem 1.4, and we discuss the relation of our work with [10].

Applications of Theorem 1.4 are perhaps most interesting when the sets A_q can be described globally without reference to the Chinese Remainder Theorem or the prime factorization of q. For example, A_q could be the set of solutions of certain equations (e.g., roots of a fixed polynomial with integral coefficients), or the set of parameters where a family of equations has a solution (e.g., the set of squares modulo q), or combinations of these. Or, for example, one may restrict the values q to be the norms of ideals in a given number field K.

2.1. Variations on roots of polynomial congruences

We begin with an application of Theorem 1.4 to roots of polynomials. This gives a higher dimensional version of Hooley's result, and is motivated by a question of Hrushovski [11, Conjecture 4.1].

Theorem 2.1. *Let $d \geqslant 1$. Let $f \in \mathbf{Z}[X]$ be a polynomial with d distinct complex roots. For each prime power p^v, let A_{p^v} denote the subset of $(\mathbf{Z}/p^v\mathbf{Z})^{d-1}$ consisting of points $(a, a^2, \dots, a^{d-1})$ where a runs over the roots of $f(x) \equiv 0 \pmod{p^v}$. Then, with the corresponding definitions of $\mathcal{Q}$ and Δ_q, for large x we have*

$$\frac{1}{|\mathcal{Q}(x)|} \sum_{q \in \mathcal{Q}(x)} \operatorname{disc}(\Delta_q) \ll_d (\log x)^{-\frac{1}{(4d)d!}}.$$

Proof. Let K_f denote the splitting field of f over $\mathbf{Q}$, which has degree $[K_f : \mathbf{Q}] \leqslant d!$. If a large prime p splits completely in K_f, then there are d distinct solutions to the congruence $f(x) \equiv 0 \pmod p$, so that $\varrho(p) = d$ for such primes. Further, by the Chebotarev density theorem the proportion of primes that split completely in K_f is $1/[K_f : \mathbf{Q}] \geqslant 1/d!$, so that Assumption 1.1 holds. Finally, any affine hyperplane in $(\mathbf{Z}/p\mathbf{Z})^{d-1}$ can intersect the curve $(t, t^2, \dots, t^{d-1})$ in at most $d-1$ points. Thus $\lambda(p) \leqslant d-1$, and we conclude that

$$\sum_{\substack{p\leqslant x\\ \varrho(p)\geqslant 1}} \Big(1 - \frac{\lambda(p)}{\varrho(p)}\Big)\frac{1}{p} \geqslant \sum_{\substack{p\leqslant x\\ \varrho(p)= d}} \Big(1 - \frac{d-1}{d}\Big)\frac{1}{p} \geqslant \frac{1}{d}\Big(\frac{1}{d!} + o(1)\Big)\log\log x.$$

The result now follows from Theorem 1.4. □

Stated qualitatively, Theorem 2.1 implies that the measures

$$\frac{1}{|\mathcal{Q}(x)|}\sum_{q\in\mathcal{Q}(x)}\frac{1}{\varrho(q)}\sum_{\substack{a \,(\mathrm{mod}\, q)\\ f(a)\equiv 0 \,(\mathrm{mod}\, q)}} \delta_{\{\frac{a}{q}, \frac{a^2}{q}, \dots, \frac{a^{d-1}}{q}\}}$$

converge to the uniform measure as $x \to \infty$. Indeed Theorem 2.1 implies a quantitative "mod q" version of [11, Conjecture 4.1]; this conjecture is related to the axiomatization (in the setting of continuous first-order logic) of the theory of finite prime fields with an additive character. In the remarks below we mention a few other related applications that may be either deduced qualitatively from Theorem 2.1, or established in a quantitative form by adapting the same argument.

Example 1. If $d \geqslant 2$, then by ignoring all but the first coordinate, the equidistribution of $\{\frac{a}{q}, \frac{a^2}{q}, \dots, \frac{a^{d-1}}{q}\}$ implies the equidistribution of the first coordinate $\{\frac{a}{q}\}$. Let $f \in \mathbf{Z}[X]$ be a polynomial with $d \geqslant 2$ distinct complex roots, and let A_{p^v} denote the subset of $\mathbf{Z}/p^v\mathbf{Z}$ consisting of the points a with $f(a) \equiv 0 \pmod{p^v}$. In this 1-dimensional case we may take $\lambda(p) = 1$. Then, with the usual meanings of $\mathcal{Q}$, Δ_q, we have for large x

$$\frac{1}{|\mathcal{Q}(x)|}\sum_{q\in\mathcal{Q}(x)} \mathrm{disc}(\Delta_q) \ll_d (\log x)^{-\frac{1}{8(d!)}}.$$

This is a version of Hooley's result, and we shall discuss the differences from his formulation in the next subsection. Note that f does not have to be irreducible, but should merely have at least two distinct complex roots. The case of reducible quadratic polynomials was discussed earlier by Martin and Sitar [15], and has been studied further by Dartyge and Martin [2].

Example 2. Let $f \in \mathbf{Z}[X]$ have $d \geqslant 2$ distinct complex roots, and let $g \in \mathbf{Z}[X]$ be a non-constant polynomial of degree $< d$. For each prime power p^v, let A_{p^v} denote the set of residue classes $g(a) \pmod{p^v}$ where a is a root of $f(x) \equiv 0 \pmod{p^v}$. Let A_q, $\mathcal{Q}$, Δ_q

have their usual meanings. As we saw in the proof of Theorem 2.1 for a density of primes at least $1/d!$, the congruence $f(x) \equiv 0 \pmod p$ has d roots. Since g is non-constant and has degree $\leqslant d-1$, for such primes p we see that A_p has at least 2 elements. Therefore, we obtain using Theorem 1.2 that

$$\frac{1}{|\mathcal{Q}(x)|} \sum_{q \in \mathcal{Q}(x)} \operatorname{disc}(\Delta_q) \ll_d (\log x)^{-\frac{1}{7(d!)}}.$$

In other words, for most $q \in \mathcal{Q}$, the points $g(a) \pmod q$ get equidistributed.

To give another variant, suppose now that $g \in \mathbf{Z}[X]$ has degree at least 2 but at most $d-1$, and let now A_{p^v} denote the set of points $(a, g(a)) \in (\mathbf{Z}/p^v\mathbf{Z})^2$ where $f(a) \equiv 0 \pmod{p^v}$. The intersection of A_p with any affine hyperplane has at most $d-1$ points, and so an application of Theorem 1.4 shows that

$$\frac{1}{|\mathcal{Q}(x)|} \sum_{q \in \mathcal{Q}(x)} \operatorname{disc}(\Delta_q) \ll_d (\log x)^{-\frac{1}{4d(d!)}}.$$

Example 3. Here is (essentially) a reformulation of the previous example. Let f and g be two polynomials in $\mathbf{Z}[X]$ with degrees d_1 and d_2 respectively. Assume that $f \circ g$ has d distinct complex roots with $d > d_2$. Take A_{p^v} to be the set of residue classes $a \pmod{p^v}$ such that $f(a) \equiv 0 \pmod{p^v}$, and such that $a \equiv g(b) \pmod{p^v}$ is a value of the polynomial g. This fits the framework of Example 2, by noting that b is a root of $f \circ g \pmod{p^v}$ and then a is just the value $g(b)$. Thus, we obtain the equidistribution of $\{a/q\}$ for those roots a of a polynomial f that are constrained to be in the image of a polynomial g.

Example 4. We now consider extensions of Theorem 2.1, where the moduli q are restricted to the integers all of whose prime factors lie in a prescribed set $\mathcal{P}$. That is, given $f \in \mathbf{Z}[X]$ with at least 2 distinct complex roots, we take $A_{p^v} = \emptyset$ if $p \notin \mathcal{P}$ and when $p \in \mathcal{P}$ take A_{p^v} to be the points $(a, a^2, \dots, a^{d-1}) \in (\mathbf{Z}/p^v\mathbf{Z})^{d-1}$ where a is a root of $f \pmod{p^v}$. Or, as in Example 1, we could consider the one dimensional situation of A_{p^v} being the roots of $f \pmod{p^v}$ for $p \in \mathcal{P}$. We now give a couple of examples of such analogues of Theorem 2.1.

Let $K/\mathbf{Q}$ be a Galois extension, and let $\mathcal{P}$ denote the set of primes that are the norm of a principal ideal in K. This means that the primes in $\mathcal{P}$ are those that are completely split in H_K, the Hilbert class field of K. The set $\mathcal{P}'$ of primes that are completely split in the compositum $H_K K_f$ (with K_f the splitting field of f) form a subset of $\mathcal{P}$ and if $p \in \mathcal{P}'$ then $f \equiv 0 \pmod p$ has d roots. The Chebotarev density theorem shows that $\mathcal{P}'$ has positive density. Thus

$$\sum_{\substack{p \in \mathcal{P} \\ \varrho(p) \geqslant 1 \\ p \leqslant x}} \Big(1 - \frac{\lambda(p)}{\varrho(p)}\Big)\frac{1}{p} \geqslant \sum_{\substack{p \in \mathcal{P}' \\ p \leqslant x}} \Big(1 - \frac{d-1}{d}\Big)\frac{1}{p} \geqslant \delta(K, f) \log\log x,$$

for some constant $\delta(K,f)>0$ and all large x. Theorem 1.4 now gives the equidistribution of A_q for most moduli q for which $f\equiv 0 \pmod q$ has a root, and when the prime factors of q are constrained to the set $\mathcal{P}$. For example, if $m\geqslant 1$ is a fixed integer, this applies to $\mathcal{P}$ being the set of primes of the form x^2+my^2.

To give a complementary example, suppose $K/\mathbf{Q}$ is a Galois extension, with $K\neq\mathbf{Q}$, that is linearly disjoint from K_f, and take $\mathcal{P}$ to be the set of primes that are *not* norms of ideals in K. Since K and K_f are linearly disjoint, the Galois group of the compositum KK_f is isomorphic to $G\times G_f$. There is a positive density of primes p such that the Frobenius at p is trivial in G_f, so that $\varrho_f(p)=d\geqslant 2$ (if $p\nmid D$), but non-trivial in G (since $|G|\geqslant 2$). Then p is not the norm of an ideal of $\mathbf{Z}_K$, so $p\in\mathcal{P}$. Now we may apply Theorem 1.4 as usual.

Remark 2.2. D.R. Heath-Brown has informed us of another possible variant of these results. If $F(x,y)$ is an irreducible integral form of degree >1, then one can obtain the equidistribution (for the relevant moduli q) of the fractional parts of solutions (x,y) to $F(x,y)\equiv 0 \pmod q$. Such a result might potentially be used to count the number of points of bounded height on the Châtelet surfaces $Z^2+W^2=F(X,Y)$ where F is a quartic polynomial (see [3]).

2.2. Hooley's measures

We now compare our results with the precise statement of [10]. If f is a fixed primitive irreducible polynomial in $\mathbf{Z}[X]$ with degree at least 2, then Hooley [10] showed that the probability measures

$$\mu_x=\frac{1}{M_x}\sum_{q\in\mathcal{Q}(x)}\varrho_f(q)\Delta_q=\frac{1}{M_x}\sum_{q\in\mathcal{Q}(x)}\sum_{a\in Z_q}\delta_{\{\frac{a}{q}\}}$$

converge, as $x\to+\infty$, to the uniform measure on $\mathbf{R}/\mathbf{Z}$. Here

$$M_x=\sum_{q\leqslant x}\varrho_f(q)$$

denotes a normalizing factor, which is asymptotically $C_f x$ for a positive constant C_f. Hooley's measures are not the same as the measures

$$\frac{1}{|\mathcal{Q}(x)|}\sum_{q\in\mathcal{Q}(x)}\Delta_q=\frac{1}{|\mathcal{Q}(x)|}\sum_{q\in\mathcal{Q}(x)}\frac{1}{\varrho_f(q)}\sum_{a\in Z_q}\delta_{\{\frac{a}{q}\}}$$

that occur implicitly in Theorem 1.2. In the context of equidistribution arising from the Chinese Remainder Theorem, the measures we introduce seem more natural, and an analogue of Theorem 1.2 for the measures μ_x is false in general.

Proposition 2.3. *There exist sets $A_p \subset \mathbf{Z}/p\mathbf{Z}$ defined for all primes p, with $|A_p| \geqslant 2$ for all p large enough, such that the measures*

$$\mu_x = \frac{1}{M_x} \sum_{q \in \mathcal{Q}(x)} \varrho(q) \Delta_q = \frac{1}{M_x} \sum_{q \in \mathcal{Q}(x)} \sum_{a \in Z_q} \delta_{\{\frac{a}{q}\}}, \qquad \text{with} \qquad M_x = \sum_{q \leqslant x} \varrho(q),$$

do not converge to the uniform measure as $x \to +\infty$. Here we take $A_{p^v} = \emptyset$ for all $v \geqslant 2$.

Lemma 2.4. *Let g denote the multiplicative function defined on squarefree integers q by setting $g(p) = 0$ for $p \leqslant e^2$, and $g(p) = \lfloor p/\log p \rfloor$ for $p > e^2$. Then there is an absolute constant C such that for all large x*

$$\sum_{q \leqslant x} g(q) \leqslant C \sum_{p \leqslant x} g(p). \tag{3}$$

Proof. Since $\sum_{p \leqslant x} g(p) \gg x^2/(\log x)^2$, the lemma amounts to proving the bound

$$\sum_{q \leqslant x} g(q) \ll \frac{x^2}{(\log x)^2}. \tag{4}$$

If q is a squarefree integer only divisible by primes $> e^2$, then a simple induction on the number of prime factors of q shows that

$$\prod_{p|q} \log p \geqslant \log q.$$

Consequently, if q can be factored $q = q_1 q_2$ with $q_i > q^{1/10}$, then

$$\prod_{p|q} \log p = \prod_{p|q_1} \log p \prod_{p|q_2} \log p \geqslant (\log q_1)(\log q_2) \geqslant \frac{1}{100} (\log q)^2.$$

Thus the contributions of such integers $q \leqslant x$ to the left-hand side of (4) is

$$\leqslant 100 \sum_{q \leqslant x} \frac{q}{(\log q)^2} \ll \frac{x^2}{(\log x)^2}.$$

The contribution of q with $q \leqslant x^{9/10}$ is also of smaller order of magnitude.

It remains to consider the contribution of integers $x^{9/10} \leqslant q \leqslant x$ that cannot be factored as $q_1 q_2$ with $x^{1/5} \leqslant q_i \leqslant x^{4/5}$. Note that such q must have largest prime factor at least $x^{1/20}$, else a greedy procedure would produce a factorization of q with both factors large. Thus the remaining integers $x^{9/10} \leqslant q \leqslant x$ may be written as pq_1 with $p > x^{1/20}$ and their contribution is

$$\ll \sum_{q_1\leqslant x^{19/20}} g(q_1) \sum_{x^{1/20}\leqslant p\leqslant x/q_1} \frac{p}{\log p} \ll \sum_{q_1\leqslant x^{19/20}} g(q_1)\frac{x^2}{q_1^2(\log x)^2}$$
$$\ll \frac{x^2}{\log x^2}\prod_{p\leqslant x^{19/20}}\Big(1+\frac{g(p)}{p^2}\Big) \ll \frac{x^2}{(\log x)^2},$$

since the Euler product over all primes converges. This concludes the proof of (4), and the lemma. □

Proof of Proposition 2.3. For $p > e^2$ take A_p to be the set of residue classes $k \pmod p$ with $1 \leqslant k \leqslant g(p)$, with g as in Lemma 2.4. Take $A_{p^v} = \emptyset$ for all $v \geqslant 2$. Here $M_x = \sum_{q\leqslant x} g(q)$, and note that for any $\varepsilon > 0$ if $p > e^{1/\varepsilon}$ then all the $g(p)$ points k/p with $k \in A_p$ land in the interval $[0, \varepsilon]$. Therefore, using Lemma 2.4, for large x

$$\mu_x([0,\varepsilon]) \geqslant \frac{1}{M_x}\sum_{e^{1/\varepsilon}<p\leqslant x} g(p) \geqslant \frac{1}{2C}.$$

Choosing $\varepsilon = 1/(4C)$ we see that μ_x does not converge to the uniform measure. □

Remark 2.5. (1) One can prove generalizations of the result of [10] to arbitrary sets defined by the Chinese Remainder Theorem by assuming in addition that the sets A_{p^v} are not too large. For instance, we can show that if the estimates

$$\sum_{\substack{p\leqslant x\\ \varrho(p)\geqslant 2}} \log p \gg x, \qquad \sum_{p^v\leqslant x} \varrho(p^v)^2 \log p^v \ll x$$

hold for x large enough, then the measures

$$\mu_x = \frac{1}{M_x}\sum_{q\in\mathcal{Q}(x)} \varrho(q)\Delta_q, \qquad M_x = \sum_{q\in\mathcal{Q}(x)}\varrho(q),$$

converge to the uniform measure on $\mathbf{R}/\mathbf{Z}$.

Since these conditions hold for the set of roots modulo p of a fixed monic polynomial f (where $\varrho_f(q) \leqslant \deg(f)$), this would recover [10, Th. 2].

(2) For some precise computations of Weyl sums (relative to Hooley's measures) for some reducible polynomials, see the work of Dartyge and Martin [2].

2.3. Equidistribution of Bezout points

Let $n \geqslant 2$ be fixed, and let X_1 and X_2 be two reduced closed subschemes of $\mathbf{A}^n/\mathbf{Z}$. Assume that the generic fiber of X_1 is a geometrically connected curve over $\mathbf{Q}$, of degree d_1, and that the generic fiber of X_2 is a geometrically connected hypersurface of degree d_2.

(Concretely, X_2 is the zero set of an absolutely irreducible integral polynomial with n variables, and X_1 could be given by $n-1$ "generically transverse" such equations.)

Assume that the closures of the generic fibers of X_1 and X_2 in $\mathbf{P}^n/\mathbf{Q}$ intersect transversely. The intersection is then finite by Bezout's Theorem, and has d_1d_2 geometric points (note that we assume transverse intersection also at infinity). Let $k \leqslant d_1d_2$ be the number of geometric intersection points belonging to the hyperplane at infinity.

For any prime power p^v, let $A_{p^v} = (X_1 \cap X_2)(\mathbf{Z}/p^v\mathbf{Z})$ be the set of $\mathbf{Z}/p^v\mathbf{Z}$-rational intersection points of the curve and the hypersurface. Then, for any q, the set A_q is the set of intersection points with coordinates in $\mathbf{Z}/q\mathbf{Z}$.

The generic fiber of the intersection variety $X_1 \cap X_2$ is defined over $\mathbf{Q}$, and has finitely many geometric points. Let γ be the Galois action of the Galois group of $\mathbf{Q}$ on $X_1 \cap X_2$. The fixed field K of the kernel of this action is a finite Galois extension $K/\mathbf{Q}$. If p is totally split in K, then all intersection points are fixed by the Frobenius conjugacy class of K at p, which means that their coordinates belong to $\mathbf{Z}/p\mathbf{Z}$. Combining this with Bezout's Theorem, it follows that there exists a set of primes p of positive density such that $|A_p| = d_1d_2 - k$.

We assume next that $d_2 \geqslant 2$ and that the curve X_1 is not contained in an affine hyperplane H (this implies that $d_1 \geqslant 2$, but is a stronger assumption if $n \geqslant 3$). Then for any affine hyperplane $H \subset (\mathbf{Z}/p\mathbf{Z})^n$, we have

$$|A_p \cap H| \leqslant \min(d_1, d_2)$$

so that $\lambda(p) \leqslant \min(d_1, d_2)$. Hence we conclude from Theorem 1.4 that for most q the fractional parts of the intersection points modulo q become equidistributed in $(\mathbf{R}/\mathbf{Z})^n$, provided $\min(d_1, d_2) < d_1d_2 - k$. As in the case of polynomial congruences, it is natural to ask whether the equidistribution of fractional parts of intersection points holds for prime moduli.

As a concrete example, suppose that X_1 and X_2 are the plane curves given by the equations

$$X_1 \colon X^3 + Y^3 = 1, \qquad X_2 \colon Y^2 = X^3 - 2.$$

These curves intersect transversally (including on the line at infinity in $\mathbf{P}^2$, since they have no common point there), and hence the condition holds since $3 < 9$.

2.4. Pseudo-polynomials

A *pseudo-polynomial*, in the sense of Hall [9], is an arithmetic function $f \colon \mathbf{Z} \to \mathbf{Z}$ such that $m - n$ divides $f(m) - f(n)$ for all integers $m \neq n$. In other words, for each $q \geqslant 1$, the reduction of f modulo q is q-periodic. Examples of such functions are given by polynomials $f \in \mathbf{Z}[X]$, but there are uncountably many pseudo-polynomials that are not polynomials (see [9, Th. 1]). Among the simplest explicit examples are $f_1(n) = \lfloor en! \rfloor$ ([9, Cor. 2]), and

$$f_2(n) = 1 - n + \frac{n(n-1)}{2} + \cdots + (-1)^n \frac{n!}{2} = (-1)^n D(n),$$

where $D(n)$ is the number of *derangements* (permutations without fixed points) in the symmetric group on n letters. The formula for $D(n)$ is a classical application of inclusion–exclusion, and that f_2 is a pseudo-polynomial follows then from [9, Th. 1]).

For a pseudo-polynomial f, and a positive integer q, take A_q to be the zeros of $f \pmod q$; that is, A_q is the set of residue classes $n \pmod q$ with $f(n) \equiv 0 \pmod q$. These sets A_q are built out of the sets A_{p^v} for prime powers p^v using the Chinese Remainder Theorem. As we have discussed, the sets A_q get equidistributed for most q, when f is a genuine polynomial. Does Theorem 1.2 also apply generally to pseudo-polynomials? Vivian Kuperberg [14] pointed out to us that there are pseudo-polynomials whose values are only divisible by a very sparse sequence of primes (indeed, one may make this sequence increase arbitrarily rapidly). Thus there is no hope of applying Theorem 1.2 to a general pseudo-polynomial, but the examples f_1 and f_2 seem well behaved, and we present some numerical experiments concerning these examples. For computations with f_1 and f_2, it is efficient to use the recursive definitions

$$f_1(1) = 2, \qquad f_1(n+1) = 1 + (n+1) f_1(n),$$
$$f_2(0) = 1, \qquad f_2(n+1) = 1 - (n+1) f_2(n).$$

Numerical experiments suggest that the values $f_1(n) = \lfloor en! \rfloor \pmod p$ for $1 \leqslant n \leqslant p$ behave like p independent random residue classes drawn uniformly from $\mathbf{Z}/p\mathbf{Z}$. If so, this suggests that there are k solutions to $f_1(n) \equiv 0 \pmod p$ for a proportion $e^{-1}/k!$ of the primes p below x: that is, for any $k \geqslant 0$

$$\lim_{x \to +\infty} \frac{1}{\pi(x)} |\{p \leqslant x \mid \varrho(p) = k\}| = \frac{1}{e} \frac{1}{k!}.$$

In other words, the quantity $\varrho(p)$ is distributed like a Poisson random variable with parameter 1. If true, this would imply that Theorem 1.2 applies to the zeros of f_1 modulo primes. However, we do not know how to prove that $\varrho(p) \geqslant 2$ for an infinite set of primes.

The following tables give the empirical and theoretical Poisson distribution for the 78498 primes $p \leqslant x = 10^6$ (normalized by multiplying the Poisson probabilities by $\pi(x)$; no empirical value is larger than 8 in that range), as well as the empirical and theoretical moments of order $1 \leqslant n \leqslant 4$.

Empirical and theoretical probability distribution

k	0	1	2	3	4	5	6	7	8
Empirical	29054	28822	14314	4777	1250	236	38	5	2
Poisson	28877.8	28877.8	14438.9	4813	1203.2	240.6	40.17	5.7	0.7

Empirical and theoretical moments

n	1	2	3	4
Empirical	0.99671	1.9964	5.0034	15.054
Poisson	1	2	5	15

Note that if $g \in \mathbf{Z}[X]$ is an irreducible polynomial of degree n with Galois group S_n (the generic case), then the Chebotarev density theorem implies that

$$\lim_{x\to+\infty} \frac{1}{\pi(x)} |\{p \leqslant x \mid \varrho_g(p) = k\}| = \frac{1}{n!} |\{\pi \in S_n \text{ with } k \text{ fixed points}\}|.$$

Now for large n, the number of fixed points of a permutation drawn uniformly at random from S_n is distributed approximately like a Poisson random variable with parameter 1. Thus our guess above on the number of zeros of the pseudo-polynomial $f_1 \pmod p$ is akin to what holds for a generic irreducible polynomial of large degree.

For the function $f_2(n) = (-1)^n D(n)$, numerical experiments also suggest that there is a positive density of primes with $\varrho(p) \geqslant 2$, so that Theorem 1.2 should apply. Once again we are unable to establish such a claim.

But, if we put $f_3(n) = f_2(n) - 1$, then from the recurrence for f_2 given above we may recognize that $f_3(0) = 0$, and $f_3(p-1) \equiv 0 \pmod p$ for each prime p. Thus in this case $\varrho(p) \geqslant 2$ for each prime p, and Theorem 1.2 applies. Note that $|f_3(n)|$ has a combinatorial meaning: it equals the number of permutations in S_n with exactly one fixed point. Since $|f_3|$ and f_3 have the same zeros $\pmod q$ for any q, we see that Theorem 1.2 applies to the combinatorial sequence $|f_3(n)|$.

3. Preliminaries

Throughout we work in the higher dimensional framework of Theorems 1.4, 1.6, 1.8, so that A_q is a subset of $(\mathbf{Z}/q\mathbf{Z})^n$, and $\varrho(q)$ is its cardinality. We keep in place Assumption 1.1, and have in mind that x is large in comparison to α and x_0.

3.1. The sets $\mathcal{Q}$ and $\mathcal{Q}_k$

We begin by gaining an understanding of the size of the sets $\mathcal{Q}(x)$ and $\mathcal{Q}_k(x)$ (of elements in $\mathcal{Q}$ with exactly k distinct prime factors).

Lemma 3.1. *For x large enough in terms of α and x_0*

$$|\mathcal{Q}(x)| \gg \frac{\alpha x}{\log x} \prod_{\substack{p\leqslant x\\ p\in\mathcal{Q}}} \Big(1 + \frac{1}{p}\Big).$$

Proof. Observe that

$$|\mathcal{Q}(x)| \geqslant \frac{1}{\log x} \sum_{q\in\mathcal{Q}(x)} \log q \geqslant \frac{1}{\log x} \sum_{q\in\mathcal{Q}(x)} \sum_{pd=q} \log p \geqslant \frac{1}{\log x} \sum_{\substack{d<x^{1/3}\\ d\in\mathcal{Q}}} \sum_{\substack{x^{1/3}<p\leqslant x/d\\ p\in\mathcal{Q}}} \log p.$$

Using Assumption 1.1, it follows for large x that

$$|\mathcal{Q}(x)| \geqslant \frac{\alpha x}{2\log x} \sum_{\substack{d<x^{1/3}\\ d\in\mathcal{Q}}} \frac{1}{d}.$$

Now put $z = x^{1/9}$ and $\tau = 1/\log z$, and note that (restricting attention to squarefree d)

$$\sum_{\substack{d<x^{1/3}\\ d\in\mathcal{Q}}} \frac{1}{d} \geqslant \sum_{\substack{d<x^{1/3}\\ d\in\mathcal{Q}\\ p|d \Longrightarrow p\leqslant z}} \frac{\mu(d)^2}{d} = \prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}} \Big(1+\frac{1}{p}\Big) - \sum_{\substack{d>x^{1/3}\\ d\in\mathcal{Q}\\ p|d \Longrightarrow p\leqslant z}} \frac{\mu(d)^2}{d},$$

and further

$$\sum_{\substack{d>x^{1/3}\\ d\in\mathcal{Q}\\ p|d \Longrightarrow p\leqslant z}} \frac{\mu(d)^2}{d} \leqslant \sum_{\substack{d\in\mathcal{Q}\\ p|d \Longrightarrow p\leqslant z}} \frac{\mu(d)^2}{d}\Big(\frac{d}{x^{1/3}}\Big)^{\tau} = e^{-3}\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{p^{\tau}}{p}\Big).$$

Therefore

$$\sum_{\substack{d<x^{1/3}\\ d\in\mathcal{Q}}} \frac{1}{d} \geqslant \prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\Big(1-e^{-3}\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\frac{1+p^{\tau}/p}{1+1/p}\Big).$$

Now, for large x (and so large z),

$$\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\frac{1+p^{\tau}/p}{1+1/p} \leqslant \prod_{p\leqslant z}\Big(1+\frac{p^{\tau}-1}{p}\Big) \leqslant \exp\Big(\sum_{p\leqslant z}\frac{p^{\tau}-1}{p}\Big)$$
$$\leqslant \exp\Big(\sum_{p\leqslant z}\frac{(e-1)\tau\log p}{p}\Big) \leqslant e^2.$$

Assembling the above observations together we conclude that

$$|\mathcal{Q}(x)| \geqslant \frac{\alpha x}{2\log x}\Big(1-\frac{1}{e}\Big)\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big).$$

The lemma follows since

$$\prod_{x^{1/9}<p\leqslant x}(1+1/p) \ll 1. \quad \square$$

We can also prove a matching upper bound for $|\mathcal{Q}(x)|$, and in fact will need a such a bound for the smooth (or friable) elements in $\mathcal{Q}(x)$.

Lemma 3.2. *Let x be large, and z be a parameter with $\log x \leqslant z \leqslant x$. Then*

$$\sum_{\substack{q\in\mathcal{Q}(x)\\ p|q\Longrightarrow p\leqslant z}} 1 \ll \frac{x}{\log x}\exp\Big(-\frac{\log x}{\log z}\Big)\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big).$$

Proof. We start by noting that

$$\sum_{\substack{q\in\mathcal{Q}(x)\\ p|q\Longrightarrow p\leqslant z}} 1 \leqslant \sqrt{x}+\frac{2}{\log x}\sum_{\substack{\sqrt{x}<q\leqslant x\\ q\in\mathcal{Q}\\ p|q\Longrightarrow p\leqslant z}}\log q \leqslant \sqrt{x}+\frac{2}{\log x}\sum_{\substack{q\in\mathcal{Q}(x)\\ p|q\Longrightarrow p\leqslant z}}\sum_{\substack{q=d\ell\\ (d,\ell)=1}}\log \ell,$$

where ℓ denotes a prime power. The term $\sqrt{x}$ is much smaller than the estimate we desire, and so we may ignore it and focus on the second term above.

To estimate the second sum, we shall first sum over d (which must be in $\mathcal{Q}$), and then over ℓ. Note that ℓ must be $\leqslant x/d$, and if ℓ is a prime then it is also constrained to be $\leqslant z$. Thus, for a given d, the sum over ℓ is

$$\leqslant \sum_{\substack{p^v\leqslant x/d\\ v\geqslant 2}}\log(p^v)+\sum_{p\leqslant\min(x/d,z)}\log p \ll \frac{\sqrt{x}}{\sqrt{d}}+\min\Big(\frac{x}{d},z\Big)\ll\Big(\frac{x}{d}\Big)^{1-\tau}z^{\tau},$$

for any $\tau\in[0,\frac{1}{2}]$. Using this observation with $\tau=1/\log z$, we obtain

$$\begin{aligned}\sum_{\substack{q\in\mathcal{Q}(x)\\ p|q\Longrightarrow p\leqslant z}}\sum_{\substack{q=d\ell\\ (d,\ell)=1}}\log\ell \ll \sum_{\substack{d\in\mathcal{Q}(x)\\ p|d\Longrightarrow p\leqslant z}}\Big(\frac{x}{d}\Big)^{1-\tau}z^{\tau} &= x\exp\Big(-\frac{\log x}{\log z}\Big)\sum_{\substack{d\in\mathcal{Q}(x)\\ p|d\Longrightarrow p\leqslant z}}\frac{1}{d^{1-1/\log z}}\\ &\leqslant x\exp\Big(-\frac{\log x}{\log z}\Big)\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1-\frac{p^{1/\log z}}{p}\Big)^{-1}\\ &\ll x\exp\Big(-\frac{\log x}{\log z}\Big)\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{p^{1/\log z}}{p}\Big).\end{aligned}$$

The lemma follows upon noting that

$$\begin{aligned}\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{p^{1/\log z}}{p}\Big) &\leqslant \prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\prod_{p\leqslant z}\Big(\frac{1+p^{1/\log z}/p}{1+1/p}\Big)\\ &\leqslant \prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\exp\Big(\sum_{p\leqslant z}\frac{p^{1/\log z}-1}{p}\Big)\ll\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big).\quad\square\end{aligned}$$

The next two lemmas will be analogues of the above for the sets $\mathcal{Q}_k(x)$ for a given integer $k \geqslant 1$. Readers who are mostly interested in Theorems 1.2 and 1.4 may skip at this point to Section 3.2

Define

$$\mathcal{P}(x) = \sum_{\substack{p \leqslant x \\ p \in \mathcal{Q}}} \frac{1}{p} + 3, \tag{5}$$

so that for large x, Assumption 1.1 gives

$$\alpha \log\log x + O(1) \leqslant \mathcal{P}(x) \leqslant \log\log x + O(1). \tag{6}$$

The added constant 3 in (5) is unimportant, but will be convenient later.

Lemma 3.3. *Let x be large, and let k be an integer with $1 \leqslant k \leqslant \exp(\mathcal{P}(x)/4)$. Then*

$$|\mathcal{Q}_k(x)| \gg \frac{\alpha x}{\log x} \frac{\mathcal{P}(x)^{k-1}}{(k-1)!} \exp\Big(-\frac{4k\log k}{\mathcal{P}(x)}\Big),$$

where the implied constant is absolute.

Proof. We obtain a lower bound by counting only those elements of $\mathcal{Q}_k(x)$ that are of the form $p_1 \cdots p_k$, where the primes p_j are in strictly increasing order and satisfy $p_1, \dots, p_{k-1} \leqslant x^{1/(2k)}$. Fixing these primes $p_1, \dots, p_{k-1}$, we see using Assumption 1.1 that there are at least

$$\geqslant \frac{\alpha x}{4 p_1 \cdots p_{k-1} \log x}$$

possible choices for the large prime p_k. Therefore

$$\begin{aligned} |\mathcal{Q}_k(x)| &\geqslant \frac{\alpha x}{4\log x} \sum_{\substack{p_1 < \cdots < p_{k-1} \leqslant x^{1/(2k)} \\ p_j \in \mathcal{Q}}} \frac{1}{p_1 \cdots p_{k-1}} \\ &= \frac{\alpha x}{4\log x} \frac{1}{(k-1)!} \sum_{\substack{p_1, \dots, p_{k-1} \leqslant x^{1/(2k)} \\ p_j \in \mathcal{Q} \\ p_j \text{ distinct}}} \frac{1}{p_1 \cdots p_{k-1}}. \end{aligned}$$

Let $p_1, \dots, p_{k-2}$ be distinct primes in $\mathcal{Q}$ all below $x^{1/(2k)}$. Then

$$\sum_{\substack{p_{k-1} \leqslant x^{1/(2k)} \\ p_{k-1} \neq p_1, \dots, p_{k-2} \\ p_{k-1} \in \mathcal{Q}}} \frac{1}{p_{k-1}} = \big(\mathcal{P}(x^{\frac{1}{2k}}) - 3\big) - \frac{1}{p_1} - \cdots - \frac{1}{p_{k-2}}.$$

The quantity $1/p_1 + \ldots + 1/p_{k-2}$ is at most equal to the corresponding sum when the primes p_i are equal to the first $k-2$ primes, and hence is $\leqslant \log\log(k+1) + O(1)$, so that

$$\sum_{\substack{p_{k-1}\leqslant x^{1/(2k)}\\ p_{k-1}\neq p_1,\ldots,p_{k-2}\\ p_{k-1}\in\mathcal{Q}}} \frac{1}{p_{k-1}} \geqslant \mathcal{P}(x^{\frac{1}{2k}}) - \log\log(k+1) - C$$

for some absolute constant $C \geqslant 0$. Repeating this argument, we find the same lower bound for each of the sums over p_{k-2}, ..., p_1, and therefore we obtain the lower bound

$$|\mathcal{Q}_k(x)| \gg \frac{\alpha x}{\log x} \frac{(\mathcal{P}(x^{\frac{1}{2k}}) - \log\log(k+1) - C)^{k-1}}{(k-1)!}$$

for $x \geqslant x_0$, where the implied constant is absolute. Since

$$\mathcal{P}(x^{\frac{1}{2k}}) \geqslant \mathcal{P}(x) - \sum_{x^{\frac{1}{2k}} < p \leqslant x} \frac{1}{p} = \mathcal{P}(x) - \log k + O(1),$$

and $\log k \leqslant \mathcal{P}(x)/4$, the lemma follows. □

Lemma 3.4. *Let x be large. Let $k \leqslant (\log x)^{\frac{1}{2}}$ be a positive integer, and κ a non-negative integer with $\kappa \leqslant k$. The number of integers in $\mathcal{Q}_k(x)$ having at least κ distinct prime factors that are larger than $x^{1/(4k)}$ is*

$$\ll \frac{kx}{\log x} \frac{\mathcal{P}(x)^{k-1}}{(k-1)!} \exp\Big(\frac{2k\log k}{\mathcal{P}(x)} - \kappa\Big),$$

where the implied constant is absolute.

Proof. Let N denote this number. Write $q \in \mathcal{Q}_k$ as $q = p_1^{v_1} \cdots p_k^{v_k}$ with the primes p_j in strictly ascending order.

First, if $p_k < x^{1/(4k)}$, then $p_1 \cdots p_k \leqslant x^{1/4}$, and the number of choices for the exponents $(v_1, \ldots, v_k)$ is $\ll (\log x)^k \ll x^{\varepsilon}$ for any $\varepsilon > 0$. Therefore in this case (which is only relevant for $\kappa = 0$), we have

$$N \ll x^{1/4+\epsilon} \leqslant x^{1/3}$$

since x is large.

Suppose now that $p_k > x^{1/(4k)}$. Let $p_1^{v_1}$, ..., $p_{k-1}^{v_{k-1}}$ be fixed. Note that $p_1^{v_1} \cdots p_{k-1}^{v_{k-1}} \leqslant x^{1-1/(4k)}$, so by the Brun–Titchmarsh inequality, the number of possible choices for $p_k^{v_k}$ is

$$\leqslant \frac{3x}{p_1^{v_1} \cdots p_{k-1}^{v_{k-1}} \log(x/p_1^{v_1} \cdots p_{k-1}^{v_{k-1}})} \leqslant \frac{12kx}{p_1^{v_1} \cdots p_{k-1}^{v_{k-1}} \log x}.$$

Therefore

$$N \leqslant x^{\frac{1}{3}} + \frac{12kx}{\log x} \sum_{\substack{p_1<\cdots<p_{k-1}\leqslant x\\ p_j^{v_j}\in\mathcal{Q}\\ p_{k-\kappa+1}>x^{1/(4k)}}} \frac{1}{p_1^{v_1}\cdots p_{k-1}^{v_{k-1}}}$$
$$\leqslant x^{\frac{1}{3}} + \frac{12kx}{\log x} \sum_{\kappa-1\leqslant j\leqslant k-1} \frac{1}{j!}\Big(\sum_{\substack{x\geqslant p>x^{1/(4k)}\\ p^v\in\mathcal{Q}}} \frac{1}{p^v}\Big)^j \frac{1}{(k-1-j)!}\Big(\sum_{\substack{p\leqslant x^{1/(4k)}\\ p^v\in\mathcal{Q}}} \frac{1}{p^v}\Big)^{k-1-j},$$

where the variable j represents the number of primes among $p_1, \ldots, p_{k-1}$ that are larger than $x^{1/(4k)}$, and for each p we sum over all v such that $p^v \in \mathcal{Q}$. Now the sum over j above may be bounded by

$$e^{-(\kappa-1)} \sum_{0\leqslant j\leqslant k-1} \frac{1}{j!}\Big(\sum_{\substack{x\geqslant p>x^{1/(4k)}\\ p^v\in\mathcal{Q}}} \frac{e}{p^v}\Big)^j \frac{1}{(k-1-j)!}\Big(\sum_{\substack{p\leqslant x^{1/(4k)}\\ p^v\in\mathcal{Q}}} \frac{1}{p^v}\Big)^{k-1-j}$$
$$= \frac{e^{-(\kappa-1)}}{(k-1)!}\Big(\sum_{\substack{x\geqslant p>x^{1/(4k)}\\ p^v\in\mathcal{Q}}} \frac{e}{p^v} + \sum_{\substack{p\leqslant x^{1/(4k)}\\ p^v\in\mathcal{Q}}} \frac{1}{p^v}\Big)^{k-1}$$
$$\ll \frac{e^{-(\kappa-1)}}{(k-1)!}(\mathcal{P}(x) + (e-1)\log k + O(1))^{k-1},$$

which establishes the lemma. □

3.2. Weyl sums

For a modulus $q \in \mathcal{Q}$ and $h \in \mathbf{Z}^n$, define the normalized Weyl sum

$$W(h;q) = \frac{1}{\varrho(q)} \sum_{x\in A_q} e\Big(\frac{h\cdot x}{q}\Big) \tag{7}$$

where

$$h\cdot x = h_1x_1 + \cdots + h_nx_n.$$

We extend the definition of $\lambda(p)$ (given just before Theorem 1.4) to all positive integers. Given a prime power p^v in $\mathcal{Q}$, we let

$$\lambda(p^v) = \max_{\substack{H\subset(\mathbf{Z}/p^v\mathbf{Z})^n\\ H \text{ affine hyperplane}}} |H\cap A_{p^v}|,$$

and extend λ to $\mathcal{Q}$ by multiplicativity. By the Chinese Remainder Theorem, we have

$$\lambda(q) = \max_{\substack{H\subset(\mathbf{Z}/q\mathbf{Z})^n\\ H \text{ affine hyperplane}}} |H\cap A_q|$$

for $q\in\mathcal{Q}$, where an affine hyperplane $H\subset(\mathbf{Z}/q\mathbf{Z})^n$ is a subset of the form

$$H=\{x\in(\mathbf{Z}/q\mathbf{Z})^n \mid h_1x_1+\cdots+h_nx_n=a\}$$

for some $a\in\mathbf{Z}/q\mathbf{Z}$ and $(h_i)\in(\mathbf{Z}/q\mathbf{Z})^n\setminus\{(0,\dots,0)\}$.

For a given non-zero $h\in\mathbf{Z}^n$ and a prime power p^v, we put

$$\{h,p^v\}=\begin{cases}1 & \text{if } h\equiv 0\,(\mathrm{mod}\,p^v)\\ p^v & \text{otherwise,}\end{cases}$$

and then extend this definition multiplicatively to define $\{h,q\}$.

Lemma 3.5. (1) *If q_1 and q_2 are coprime elements of $\mathcal{Q}$, then*

$$W(h;q_1q_2)=W(\bar{q}_1h;q_2)W(\bar{q}_2h;q_1),$$

where $q_1\bar{q}_1\equiv 1\,(\mathrm{mod}\,q_2)$ and $q_2\bar{q}_2\equiv 1\,(\mathrm{mod}\,q_1)$.

(2) *Let $h\in\mathbf{Z}^n$, with $h\neq(0,\dots,0)$. For $q\in\mathcal{Q}$, we have*

$$\frac{1}{q}\sum_{a\,(\mathrm{mod}\,q)}|W(ah;q)|^2\leqslant\frac{\lambda(\{h,q\})}{\varrho(\{h,q\})}. \tag{8}$$

Proof. These are elementary statements (see [10, Lemmas 1 and 3] for $n=1$).

(1) For $x_1\in\mathbf{Z}^n$ and $x_2\in\mathbf{Z}^n$, the element of $(\mathbf{Z}/q_1q_2\mathbf{Z})^n$ which is congruent to x_i modulo q_i is the residue class of the vector

$$x=q_1\bar{q}_1x_2+q_2\bar{q}_2x_1\in\mathbf{Z}^n.$$

Therefore

$$\begin{aligned}W(h;q_1q_2)&=\frac{1}{\varrho(q_1q_2)}\sum_{x\in A_{q_1q_2}}e\Big(\frac{h\cdot x}{q_1q_2}\Big)\\ &=\frac{1}{\varrho(q_1)\varrho(q_2)}\sum_{x_1\in A_{q_1}}\sum_{x_2\in A_{q_2}}e\Big(\frac{h\cdot(q_1\bar{q}_1x_2+q_2\bar{q}_2x_2)}{q_1q_2}\Big)=W(\bar{q}_1h;q_2)W(\bar{q}_2h;q_1).\end{aligned}$$

(2) Opening the square and interchanging the order of the summations, we find that

$$\sum_{a\,(\mathrm{mod}\,q)}|W(ah;q)|^2=\frac{1}{\varrho(q)^2}\sum_{x,y\in A_q}\sum_{a\,(\mathrm{mod}\,q)}e\Big(\frac{ah\cdot(x-y)}{q}\Big).$$

By orthogonality of characters modulo q, this implies

$$\sum_{a \,(\mathrm{mod}\, q)} |W(ah;q)|^2 = \frac{q}{\varrho(q)^2} \sum_{\substack{x,y\in A_q \\ h\cdot(x-y)=0 \,(\mathrm{mod}\, q)}} 1.$$

Summing over x first, this gives

$$\sum_{a \,(\mathrm{mod}\, q)} |W(ah;q)|^2 \leqslant \frac{q}{\varrho(q)^2} \sum_{x\in A_q} \alpha(x)$$

where $\alpha(x)$ is the number of $y \in A_q$ such that $h \cdot y = h \cdot x \,(\mathrm{mod}\, q)$. By the Chinese Remainder Theorem $\alpha(x)$ is bounded by the product over $p^v \| q$ of the number of solutions to $h \cdot x = h \cdot y \,(\mathrm{mod}\, p^v)$, and this may be bounded by $\varrho(p^v)$ if $h \equiv 0 \,(\mathrm{mod}\, p^v)$ and the resulting hyperplane is degenerate, or by $\lambda(p^v)$ otherwise. Thus

$$\alpha(x) \leqslant \varrho(q/\{h,q\})\lambda(\{h,q\})$$

for all x, and the result follows. □

Remark 3.6. Part (1) is the crucial place where we use the fact that A_q is defined by the Chinese Remainder Theorem, while (2) is the only point where we detect any cancellation in the Weyl sums $W(h;q)$.

3.3. The Erdős–Turán inequality

We recall the n-dimensional Erdős–Turán inequality for the discrepancy of Δ_q (see, e.g., [8, Lemma 2] for references): for any integer $H \geqslant 1$, we have

$$\mathrm{disc}(\Delta_q) \ll \frac{1}{H} + \sum_{0<\|h\|\leqslant H} \frac{1}{M(h)} |W(h;q)|, \tag{9}$$

where $\|h\| = \max(|h_i|)$ and $M(h) = \prod_i \max(1, |h_i|)$ and where the implied constant depends only on n. We now record a consequence of Lemma 3.5 for terms appearing in (9), and then use it to bound certain useful averages of $\mathrm{disc}(\Delta_q)$.

Lemma 3.7. *Let $q \in \mathcal{Q}$ and $H \geqslant 2$ be given. Then*

$$\frac{1}{q} \sum_{a \,(\mathrm{mod}\, q)} \sum_{0<\|h\|\leqslant H} \frac{1}{M(h)} |W(ah;q)| \ll (\log H)^n \prod_{p^v \| q} \Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}} + \frac{1}{p^v} \Big),$$

where the implied constant depends only on n.

Proof. Applying the Cauchy–Schwarz inequality and (8), we have

$$\frac{1}{q}\sum_{a \,(\mathrm{mod}\, q)} \sum_{0<\|h\|\leqslant H} \frac{1}{M(h)}|W(ah;q)| \leqslant \sum_{0<\|h\|\leqslant H} \frac{1}{M(h)}\Big(\frac{\lambda(\{h,q\})}{\varrho(\{h,q\})}\Big)^{\frac{1}{2}}$$
$$= \sum_{\substack{d\mid q\\(d,q/d)=1}} \Big(\frac{\lambda(d)}{\varrho(d)}\Big)^{\frac{1}{2}} \sum_{\substack{0<\|h\|\leqslant H\\ \{q,h\}=d}} \frac{1}{M(h)},$$

since $\{h,q\}=d$ is possible only for those divisors of d that are coprime to q/d. Observe that if $1\leqslant\|h\|\leqslant H$ and $\{h,q\}=d$, then at least one of the coordinates h_i is a non-zero multiple of q/d. Therefore

$$\sum_{\substack{0<\|h\|\leqslant H\\ \{q,h\}=d}} \frac{1}{M(h)} \leqslant \frac{d}{q}\sum_{0<\|h\|\leqslant H}\frac{1}{M(h)} \ll \frac{d}{q}(\log H)^n,$$

and the lemma follows by multiplicativity. □

Lemma 3.8. *Let x be large, and z be a real number in the range $e\leqslant z\leqslant x^{1/3}$. Let $s\leqslant x^{\frac{1}{3}}$ be an integer with $s\in\mathcal{Q}$ and such that all prime factors of s are below z. Then, for any $H\geqslant 2$, we have*

$$\sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p\mid r\Longrightarrow p>z}} \mathrm{disc}(\Delta_{rs}) \ll \frac{x}{\varphi(s)\log z}\Big(\frac{1}{H}+(\log H)^n\prod_{p^v\|s}\Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}}+\frac{1}{p^v}\Big)\Big).$$

Proof. We apply the Erdős-Turán inequality (9). Using the twisted multiplicativity from Lemma 3.5, (1), which applies since r and s are coprime, we obtain

$$\sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p\mid r\Longrightarrow p>z}} \mathrm{disc}(\Delta_{rs}) \ll \sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p\mid r\Longrightarrow p>z}} \Big(\frac{1}{H}+\sum_{0<\|h\|\leqslant H}\frac{1}{M(h)}|W(\overline{r}h;s)W(\overline{s}h;r)|\Big).$$

We bound $|W(\overline{s}h;r)|$ trivially by 1, and split the sum over r into (reduced) residue classes $r\equiv\bar{a}\,(\mathrm{mod}\, s)$. If $r\equiv\bar{a}\,(\mathrm{mod}\, s)$ then $W(\bar{r}h;s)=W(ah;s)$, so that

$$\sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p\mid r\Longrightarrow p>z}} \mathrm{disc}(\Delta_{rs}) \ll \sum_{\substack{a \,(\mathrm{mod}\, s)\\ (a,s)=1}} \Big(\frac{1}{H}+\sum_{0<\|h\|\leqslant H}\frac{1}{M(h)}|W(ah;s)|\Big) \sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p\mid r\Longrightarrow p>z\\ r\equiv\overline{a}\,(\mathrm{mod}\, s)}} 1.$$

Since $s\leqslant x^{\frac{1}{3}}$, it follows that $x/s\geqslant x^{\frac{2}{3}}$. Ignoring the condition that $rs\in\mathcal{Q}$, and using the sieve, we find that

$$\sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p|r\Longrightarrow p>z\\ r\equiv\overline{a}\,(\mathrm{mod}\,s)}} 1\leqslant \sum_{\substack{r\leqslant x/s\\ p|r\Longrightarrow p>z\\ r\equiv\overline{a}\,(\mathrm{mod}\,s)}} 1\ll \frac{x/s}{\varphi(s)\log z}$$

with an absolute implied constant. Therefore

$$\sum_{\substack{r\leqslant x/s\\ rs\in\mathcal{Q}\\ p|r\Longrightarrow p>z}} \mathrm{disc}(\Delta_{rs})\ll \frac{x}{\varphi(s)\log z}\frac{1}{s}\sum_{\substack{a\,(\mathrm{mod}\,s)\\ (a,s)=1}}\Big(\frac{1}{H}+\sum_{0<\|h\|\leqslant H}\frac{1}{M(h)}|W(ah;s)|\Big).$$

Extend the sum over a to all $a\,(\mathrm{mod}\,s)$, and invoke Lemma 3.7 to conclude the proof. □

4. Proof of Theorem 1.4

Our goal is to estimate the sum

$$\sum_{q\in\mathcal{Q}(x)}\mathrm{disc}(\Delta_q),$$

in terms of the quantity

$$P:=\sum_{\substack{p\leqslant x\\ \varrho(p)\geqslant 1}}\Big(1-\frac{\lambda(p)}{\varrho(p)}\Big)\frac{1}{p}.$$

We may assume that $P\geqslant 10$, else there is nothing to prove, and put $z=x^{1/P}$. Below, we will factor any $q\in\mathcal{Q}(x)$ as $q=rs$ where all the prime factors of s are below z, and all the prime factors of r are above z. Here the letters r and s are meant to suggest the "rough" and "smooth" parts of q.[1]

Consider first the contribution of terms with $s\leqslant x^{1/3}$. Applying Lemma 3.8 with $H=e^P$ we obtain

$$\sum_{\substack{q=rs\in\mathcal{Q}(x)\\ s\leqslant x^{1/3}}}\mathrm{disc}(\Delta_{rs})\ll\sum_{\substack{s\leqslant x^{1/3}\\ s\in\mathcal{Q}}}\frac{Px}{\varphi(s)\log x}\Big(e^{-P}+P^n\prod_{p^v\|s}\Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}}+\frac{1}{p^v}\Big)\Big).$$

Note that

$$\sum_{\substack{s\leqslant x^{1/3}\\ s\in\mathcal{Q}}}\frac{1}{\varphi(s)}\leqslant\prod_{p\leqslant z}\Big(1+\sum_{\substack{v\geqslant 1\\ p^v\in\mathcal{Q}}}\frac{1}{p^{v-1}(p-1)}\Big)\ll\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p-1}\Big)\ll\prod_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big).$$

[1] French readers are invited to substitute f for s ("friable") and c for r ("criblé") throughout.

Further note that

$$\sum_{\substack{s\leqslant x^{1/3}\\ s\in\mathcal{Q}}}\frac{1}{\varphi(s)}\prod_{p^v\| s}\Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}}+\frac{1}{p^v}\Big)\leqslant\prod_{p\leqslant z}\Big(1+\sum_{\substack{v\geqslant 1\\ p^v\in\mathcal{Q}}}\frac{1}{p^{v-1}(p-1)}\Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}}+\frac{1}{p^v}\Big)\Big)$$
$$\ll\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p-1}\Big(\frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}}+\frac{1}{p}\Big)\Big)\ll\prod_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}}\Big)$$
$$\ll\prod_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\exp\Big(-\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1-\frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}}\Big)\frac{1}{p}\Big),$$

and that, since $1-\sqrt{t}\geqslant(1-t)/2$ for $0\leqslant t\leqslant 1$,

$$\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1-\frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}}\Big)\frac{1}{p}\geqslant\frac{1}{2}\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1-\frac{\lambda(p)}{\varrho(p)}\Big)\frac{1}{p}=\frac{P}{2}.$$

We conclude that

$$\sum_{\substack{q=rs\in\mathcal{Q}(x)\\ s\leqslant x^{1/3}}}\operatorname{disc}(\Delta_{rs})\ll\frac{x}{\log x}\prod_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\Big(Pe^{-P}+P^{n+1}e^{-P/2}\Big)\ll|\mathcal{Q}(x)|\frac{e^{-P/3}}{\alpha},\quad(10)$$

upon using Lemma 3.1 and recalling that implied constants are allowed to depend on n.

Now consider the contribution of terms $q=rs$ where $s>x^{1/3}$, so that $r\leqslant x^{2/3}$. Using the trivial bound $\operatorname{disc}(\Delta_q)\leqslant 1$, we see that such terms contribute

$$\sum_{\substack{q=rs\in\mathcal{Q}(x)\\ s>x^{1/3}}}\operatorname{disc}(\Delta_q)\leqslant\sum_{\substack{r\leqslant x^{2/3}\\ r\in\mathcal{Q}}}\sum_{\substack{x^{1/3}<s\leqslant x/r\\ s\in\mathcal{Q}}}1.$$

Applying Lemma 3.2, this quantity is

$$\ll\sum_{\substack{r\leqslant x^{2/3}\\ r\in\mathcal{Q}}}\frac{x/r}{\log x}\exp\Big(-\frac{\log(x/r)}{\log z}\Big)\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\ll\frac{x}{\log x}e^{-P/3}\prod_{\substack{p\leqslant z\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\sum_{\substack{r\leqslant x^{2/3}\\ r\in\mathcal{Q}}}\frac{1}{r}$$
$$\ll\frac{x}{\log x}e^{-P/3}\prod_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\Big(1+\frac{1}{p}\Big)\ll|\mathcal{Q}(x)|\frac{e^{-P/3}}{\alpha},$$

where we used Lemma 3.1 in the last step. Combining this bound with (10), we obtain Theorem 1.4, hence also Theorem 1.2.

5. The main technical result

In this section, we establish a general technical estimate, from which the simpler (but less precise) Theorems 1.6 and 1.8 will be deduced in the next section. In addition to $\mathcal{P}(x)$ (defined in (5)), we will use the quantity

$$\widetilde{\mathcal{P}}(x) = \sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}} \frac{1}{p}\Big(\frac{\lambda(p)}{\varrho(p)}\Big)^{1/2} + 3. \tag{11}$$

Since $\lambda(p)\leqslant\varrho(p)$, note that $\widetilde{\mathcal{P}}(x)\leqslant\mathcal{P}(x)$.

Proposition 5.1. *Suppose that Assumption 1.1 holds, and let x be large in terms of α and x_0.*

(1) *In the range $k\leqslant\mathcal{P}(x)$*

$$\frac{1}{|\mathcal{Q}_k(x)|}\sum_{q\in\mathcal{Q}_k(x)}\operatorname{disc}(\Delta_q) \ll \frac{k^{7+n}}{\alpha}\Big(\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k-1}{3}} + e^{-k/2}\Big(\frac{k}{\mathcal{P}(x)}\Big)^{\frac{k-1}{2}}\Big). \tag{12}$$

(2) *In the range $\mathcal{P}(x)<k\leqslant\exp(\sqrt{\log\log x})$*

$$\frac{1}{|\mathcal{Q}_k(x)|}\sum_{q\in\mathcal{Q}_k(x)}\operatorname{disc}(\Delta_q) \ll \frac{1}{\alpha}\exp\Big(\frac{(6+n)k\log k}{\mathcal{P}(x)}\Big)\Big(e^{-k/3} + \Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k}{3(1+\log(k/\mathcal{P}(x)))}}\Big). \tag{13}$$

Put $z=x^{1/(4k)}$ and factor $q\in\mathcal{Q}_k(x)$ uniquely in the form $q=rs$, where all prime factors of s are $\leqslant z$ and all prime factors of r are $>z$. Below, r and s will always be assumed to have this meaning.

We first dispense with a technical case, when $s>x^{\frac{1}{3}}$. Since s has at most k prime factors which are all below $x^{1/(4k)}$ it follows that if we write $s=s_1s_2^2$ with s_1 squarefree, then $s_1\leqslant x^{1/4}$ and $s_2>x^{1/12}$. Since $\operatorname{disc}(\Delta_q)\leqslant 1$ for all q, it follows that

$$\sum_{\substack{q\in Q_k(x)\\ s>x^{1/3}}}\operatorname{disc}(\Delta_q) \ll \sum_{s>x^{1/3}}\frac{x}{s} \ll x^{\frac{11}{12}+\varepsilon} \tag{14}$$

for any $\varepsilon>0$. Thus the contribution of such terms is negligible compared to the bounds we seek, and may be discarded. Henceforth, we restrict attention to terms with $s\leqslant x^{1/3}$.

5.1. When k is small: proof of part (1)

In this case $k\leqslant\mathcal{P}(x)$, so that $k\log k/\mathcal{P}(x)\leqslant\log k$, and Lemma 3.3, together with Stirling's formula, yields

$$|\Omega_k(x)| \gg k^{-4}\frac{\alpha x}{\log x}\frac{\mathcal{P}(x)^{k-1}}{(k-1)!} \gg k^{-5}\frac{\alpha x}{\log x}\Big(\frac{e\mathcal{P}(x)}{k}\Big)^{k-1}. \tag{15}$$

Recall the factorization $q = rs$, that q has exactly k prime factors, and s is assumed to be $\leqslant x^{1/3}$. If $\omega(s) = k$ then r must be 1, and $q = s \leqslant x^{\frac{1}{3}}$. Since $\mathrm{disc}(\Delta_q) \leqslant 1$ always, such terms contribute at most $x^{\frac{1}{3}}$. For the remaining terms when $\omega(s) < k$, we apply for each s the bound arising from Lemma 3.8. Thus, using also (14), for any $H \geqslant 2$,

$$\sum_{q\in\Omega_k(x)} \mathrm{disc}(\Delta_q) \ll x^{\frac{11}{12}+\epsilon} + \frac{kx}{\log x}\sum_{\substack{s\in\Omega(x^{1/3})\\ \omega(s)\leqslant k-1}}\frac{1}{\varphi(s)}\Big(\frac{1}{H} + (\log H)^n\prod_{p^v\| s}\Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}} + \frac{1}{p^v}\Big)\Big). \tag{16}$$

Observe that

$$\sum_{\substack{s\in\Omega(x^{1/3})\\ \omega(s)\leqslant k-1}}\frac{1}{\varphi(s)} \leqslant \sum_{j=0}^{k-1}\frac{1}{j!}\Big(\sum_{\substack{p\in\Omega\\ p\leqslant z}}\frac{1}{p-1} + \sum_{\substack{p\leqslant z\\ v\geqslant 2}}\frac{1}{p^{v-1}(p-1)}\Big)^j \leqslant \sum_{j=0}^{k-1}\frac{1}{j!}\mathcal{P}(x)^j,$$

by summing according to the number j of prime factors of s. Similarly

$$\begin{aligned}\sum_{\substack{s\in\Omega(x^{1/3})\\ \omega(s)\leqslant k-1}}\frac{1}{\varphi(s)}\prod_{p^v\| s}\Big(\frac{\sqrt{\lambda(p^v)}}{\sqrt{\varrho(p^v)}} + \frac{1}{p^v}\Big) &\leqslant \sum_{j=0}^{k-1}\frac{1}{j!}\Big(\sum_{\substack{p\in\Omega\\ p\leqslant z}}\frac{1}{p-1}\Big(\frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}} + \frac{1}{p}\Big)\\ &\quad + \sum_{\substack{p\leqslant z\\ v\geqslant 2}}\frac{1}{\varphi(p^v)}\Big(1+\frac{1}{p^v}\Big)\Big)^j\\ &\leqslant \sum_{j=0}^{k-1}\frac{1}{j!}\widetilde{\mathcal{P}}(x)^j.\end{aligned}$$

Therefore, from (16) it follows that

$$\sum_{q\in\Omega_k(x)}\mathrm{disc}(\Delta_q) \ll x^{\frac{11}{12}+\varepsilon} + \frac{kx}{\log x}\sum_{j=0}^{k-1}\Big(\frac{1}{H}\frac{\mathcal{P}(x)^j}{j!} + (\log H)^n\frac{\widetilde{\mathcal{P}}(x)^j}{j!}\Big)$$

for any $\varepsilon > 0$. We choose here $H = (1 + \mathcal{P}(x)/\widetilde{\mathcal{P}}(x))^k$ so that for all $0 \leqslant j \leqslant k-1$ one has $\mathcal{P}(x)^j/H \leqslant \widetilde{\mathcal{P}}(x)^j$. Noting that

$$(\log H)^n = \Big(k\log\Big(1+\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big)\Big)^n \ll k^n\Big(\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big)^{\frac{1}{10}},$$

we conclude that

$$\sum_{q\in\mathcal{Q}_k(x)}\operatorname{disc}(\Delta_q)\ll\frac{k^{1+n}x}{\log x}\Big(\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big)^{\frac{1}{10}}\sum_{j=0}^{k-1}\frac{\widetilde{\mathcal{P}}(x)^j}{j!},\tag{17}$$

where the term $x^{\frac{11}{12}+\varepsilon}$ has been absorbed into the much larger quantity displayed above (for ε small enough).

Suppose first that $k\leqslant 2\widetilde{\mathcal{P}}(x)-1$. In the range $0\leqslant j\leqslant k-1$, the quantity $\widetilde{\mathcal{P}}(x)^j/j!$ attains its maximum at some j_0 which lies in the range $k-1\geqslant j_0\geqslant (k-1)/2$. Note that, since $k\leqslant\mathcal{P}(x)$

$$\frac{\widetilde{\mathcal{P}}(x)^{j_0}}{j_0!}\frac{(k-1)!}{\mathcal{P}(x)^{k-1}}\leqslant\frac{\widetilde{\mathcal{P}}(x)^{j_0}}{j_0!}\frac{j_0!}{\mathcal{P}(x)^{j_0}}\leqslant\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k-1}{2}}.$$

Combining this with (15) and (17), we conclude that in this range of k,

$$\sum_{q\in\mathcal{Q}_k(x)}\operatorname{disc}(\Delta_q)\ll|\mathcal{Q}_k(x)|\frac{k^{7+n}}{\alpha}\Big(\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big)^{\frac{1}{10}}\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k-1}{2}}\ll|\mathcal{Q}_k(x)|\frac{k^{7+n}}{\alpha}\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k-1}{3}}.\tag{18}$$

Suppose now that $\mathcal{P}(x)\geqslant k\geqslant 2\widetilde{\mathcal{P}}(x)-1$. Here we note that the sum over j in (17) is $\leqslant\exp(\widetilde{\mathcal{P}}(x))\ll e^{(k-1)/2}$. Moreover, since $\widetilde{\mathcal{P}}(x)\geqslant 2$,

$$e^{(k-1)/2}\Big(\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big)^{\frac{1}{10}}\Big(\frac{k}{e\mathcal{P}(x)}\Big)^{k-1}\leqslant k^{\frac{1}{10}}e^{-(k-1)/2}\Big(\frac{k}{\mathcal{P}(x)}\Big)^{k-1-\frac{1}{10}}.$$

Combining these observations with (15) and (17), we find that in this range of k,

$$\sum_{q\in\mathcal{Q}_k(x)}\operatorname{disc}(\Delta_q)\ll|\mathcal{Q}_k(x)|\frac{k^{6+n}}{\alpha}e^{-k/2}\Big(\frac{k}{\mathcal{P}(x)}\Big)^{\frac{k-1}{2}}.\tag{19}$$

The estimates (18) and (19) establish part (1) of Proposition 5.1.

5.2. When k is large: proof of part (2)

Assume that $\mathcal{P}(x)<k\leqslant\exp(\sqrt{\log\log x})$. Let $\kappa\leqslant k/3$ be a parameter to be fixed later. For terms $q=rs$ with $\omega(r)\geqslant\kappa$, note that $\operatorname{disc}(\Delta_q)\leqslant 1$ trivially, and Lemma 3.4 gives a bound on the number of such terms. Thus

$$\begin{aligned}\sum_{\substack{q\in\mathcal{Q}_k(x)\\ \omega(r)\geqslant\kappa}}\operatorname{disc}(\Delta_q)\leqslant\sum_{\substack{q\in\mathcal{Q}_k(x)\\ \omega(r)\geqslant\kappa}}1&\ll\frac{kx}{\log x}\frac{\mathcal{P}(x)^{k-1}}{(k-1)!}\exp\Big(\frac{2k\log k}{\mathcal{P}(x)}-\kappa\Big)\\&\ll|\mathcal{Q}_k(x)|\frac{1}{\alpha}\exp\Big(\frac{7k\log k}{\mathcal{P}(x)}-\kappa\Big),\end{aligned}$$

where we used the lower bound for $|\mathcal{Q}_k(x)|$ arising from Lemma 3.3, and the fact that $k \geqslant \mathcal{P}(x)$.

On the other hand, we estimate the contributions of those q for which $\omega(r) < \kappa$ using Lemma 3.8 exactly as in the argument leading up to (17), with the same choice of H as before. Thus

$$\sum_{\substack{q\in\mathcal{Q}_k(x)\\ \omega(r)<\kappa}} \operatorname{disc}(\Delta_q) \ll \frac{k^{1+n}x}{\log x}\Big(\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big)^{\frac{1}{10}} \sum_{j=k-\kappa}^{k-1} \frac{\widetilde{\mathcal{P}}(x)^j}{j!}.$$

Now for each $k-\kappa \leqslant j \leqslant k-1$ note that, since $\kappa \leqslant k/3$,

$$\frac{\widetilde{\mathcal{P}}(x)^j}{j!}\frac{(k-1)!}{\mathcal{P}(x)^{k-1}} \leqslant \Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^j \Big(\frac{k}{\mathcal{P}(x)}\Big)^{k-1-j} \leqslant \Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{2k}{3}} \Big(\frac{k}{\mathcal{P}(x)}\Big)^{\kappa}.$$

It follows that

$$\begin{aligned}\sum_{\substack{q\in\mathcal{Q}_k(x)\\ \omega(r)<\kappa}} \operatorname{disc}(\Delta_q) &\ll \frac{k^{2+n}x}{\log x}\frac{\mathcal{P}(x)^{k-1}}{(k-1)!}\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k}{2}}\Big(\frac{k}{\mathcal{P}(x)}\Big)^{\kappa}\\ &\ll |\mathcal{Q}_k(x)|\frac{k^{2+n}}{\alpha}\exp\Big(\frac{4k\log k}{\mathcal{P}(x)}\Big)\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k}{2}}\Big(\frac{k}{\mathcal{P}(x)}\Big)^{\kappa}.\end{aligned}$$

Gathering together the bounds in the two cases $\omega(r) \geqslant \kappa$ and $\omega(r) < \kappa$, we conclude that

$$\sum_{q\in\mathcal{Q}_k(x)} \operatorname{disc}(\Delta_q) \ll \frac{|\mathcal{Q}_k(x)|}{\alpha}\exp\Big(\frac{(6+n)k\log k}{\mathcal{P}(x)}\Big)\Big(\exp(-\kappa)+\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k}{2}}\Big(\frac{k}{\mathcal{P}(x)}\Big)^{\kappa}\Big). \quad (20)$$

Choose

$$\kappa = \min\Big(\frac{k}{3}, \frac{k}{3(1+\log(k/\mathcal{P}(x)))}\log\frac{\mathcal{P}(x)}{\widetilde{\mathcal{P}}(x)}\Big).$$

A small calculation then allows us to bound the right side of (20) by

$$\ll \frac{|\mathcal{Q}_k(x)|}{\alpha}\exp\Big(\frac{(6+n)k\log k}{\mathcal{P}(x)}\Big)\Big(e^{-k/3}+\Big(\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\Big)^{\frac{k}{3(1+\log(k/\mathcal{P}(x)))}}\Big).$$

This completes the proof of (13), hence that of Proposition 5.1.

6. Proof of Theorems 1.6 and 1.8

6.1. Proof of Theorem 1.6

From the assumption (1) of Theorem 1.6, and since $1-\sqrt{t} \geqslant (1-t)/2$ for $0 \leqslant t \leqslant 1$, it follows that

$$\mathcal{P}(x) - \widetilde{\mathcal{P}}(x) = \sum_{\substack{p \leqslant x \\ p \in \mathcal{Q}}} \Big(1 - \frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}}\Big)\frac{1}{p} \geqslant \frac{1}{2}\sum_{\substack{p \leqslant x \\ p \in \mathcal{Q}}} \Big(1 - \frac{\lambda(p)}{\varrho(p)}\Big)\frac{1}{p} \geqslant \frac{\delta}{2}\log\log x.$$

Since $\mathcal{P}(x) \leqslant \log\log x + O(1)$, we conclude that

$$\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)} \leqslant 1 - \frac{\delta \log\log x}{2\mathcal{P}(x)} \leqslant 1 - \frac{\delta}{3} \leqslant e^{-\delta/3}.$$

In the range $k \leqslant \mathcal{P}(x)$, part (1) of Proposition 5.1 now gives

$$\frac{1}{|\mathcal{Q}_k(x)|}\sum_{q \in \mathcal{Q}_k(x)} \operatorname{disc}(\Delta_q) \ll \frac{k^{7+n}}{\alpha} e^{-k\delta/9} \ll \frac{1}{\alpha} e^{-k\delta/18},$$

where the last step follows because $k \geqslant 20\delta^{-1}(7+n)\log(20\delta^{-1}(7+n))$.

In the range

$$\mathcal{P}(x) < k \leqslant \exp\Big(\Big(\frac{\alpha\delta\log\log x}{20(6+n)}\Big)^{1/2}\Big),$$

we use part (2) of Proposition 5.1. Since $\mathcal{P}(x) \geqslant \alpha\log\log x + O(1)$, the upper bound on k yields

$$\exp\Big(\frac{(6+n)k\log k}{\mathcal{P}(x)}\Big) \ll \exp\Big(\frac{\delta}{18}\frac{k}{(1+\log(k/\mathcal{P}(x)))}\Big),$$

and so part (2) gives

$$\begin{aligned}\frac{1}{|\mathcal{Q}_k(x)|}\sum_{q \in \mathcal{Q}_k(x)} \operatorname{disc}(\Delta_q) &\ll \frac{1}{\alpha}\exp\Big(\frac{(6+n)k\log k}{\mathcal{P}(x)}\Big)\Big(e^{-k/3} + \big(e^{-\delta/3}\big)^{\frac{k}{3(1+\log(k/\mathcal{P}(x)))}}\Big) \\ &\ll \frac{1}{\alpha}\exp\Big(-\frac{\delta k}{18(1+\log(k/\mathcal{P}(x)))}\Big) \ll \frac{1}{\alpha}(\log x)^{-\alpha\delta/18},\end{aligned}$$

where the last step follows because $k/(1+\log(k/\mathcal{P}(x))) \geqslant \mathcal{P}(x) \geqslant \alpha\log\log x + O(1)$. This completes the proof of Theorem 1.6.

6.2. Proof of Theorem 1.8

By the Cauchy-Schwarz inequality and the assumption (2) in Theorem 1.8, we see that

$$\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\frac{1}{p}\frac{\sqrt{\lambda(p)}}{\sqrt{\varrho(p)}}\leqslant\Big(\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\frac{1}{p}\Big)^{\frac12}\Big(\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\frac{1}{p}\frac{\lambda(p)}{\varrho(p)}\Big)^{\frac12}\leqslant\sqrt{\delta}\sum_{\substack{p\leqslant x\\ p\in\mathcal{Q}}}\frac{1}{p}.$$

Therefore, with the notation of Proposition 5.1

$$\frac{\widetilde{\mathcal{P}}(x)}{\mathcal{P}(x)}\leqslant\sqrt{\delta}+O\Big(\frac{1}{\alpha\log\log x}\Big)\leqslant\delta^{1/3},$$

upon using that $\mathcal{P}(x)\geqslant\alpha\log\log x+O(1)$ and that x is large in terms of α, while $\delta\geqslant 1/\log\log x$ (by assumption again). Now part (1) of Proposition 5.1 implies that for $k\leqslant\alpha\delta\log\log x+O(1)$ one has

$$\frac{1}{|\mathcal{Q}_k(x)|}\sum_{q\in\mathcal{Q}_k(x)}\operatorname{disc}(\Delta_q)\ll\frac{k^{7+n}}{\alpha}\Big(\delta^{(k-1)/9}+e^{-k/2}\delta^{(k-1)/2}\Big)\ll\frac{1}{\alpha}\delta^{(k-1)/10},$$

which establishes Theorem 1.8.

7. Remarks on exponential sums

The method described above may be placed in a more general context as follows. Suppose we are given a function V that associates to each prime p and each reduced residue class $a \pmod p$ a complex number $V(a;p)$. Extend this to a function $V(a;q)$ where q is square-free and $a \pmod q$ is a reduced residue class by "twisted multiplicativity": that is, if $q=q_1q_2$ with $(q_1,q_2)=1$ then

$$V(a;q_1q_2)=V(a\bar{q_1};q_2)V(a\bar{q_2};q_1). \tag{21}$$

Set $V(a;q)=0$ if q is not square-free, or if $(a,q)>1$. For each prime p let $G(p)\geqslant 0$ be such that

$$\max_{(a,p)=1}|V(a,p)|\leqslant G(p). \tag{22}$$

Extend G to all square-free integers using multiplicativity. The problem is then to obtain a bound for

$$\sum_{q\leqslant x}|V(a;q)|$$

(for a fixed integer $a \geqslant 1$) which is better than the trivial bound

$$\sum_{q \leqslant x} |V(a;q)| \leqslant \sum_{q \leqslant x} G(q).$$

Remark 7.1. Our work in Theorem 1.4 fits into this framework by taking $V(a,p)$ to be the normalized Weyl sums $W(ah;p)$ for some fixed non-zero h. The twisted multiplicativity (21) was established in part (1) of Lemma 3.5.

Another very natural class of examples fitting this generalized framework arises from exponential sums. Let f_1 and f_2 be monic integral polynomials, with f_2 non-zero. For any squarefree number q, we put

$$V(a;q) = \frac{1}{\sqrt{q}} \sum_{\substack{n \,(\mathrm{mod}\, q) \\ (f_2(n),q)=1}} e\Big(\frac{a f_1(n)\overline{f_2(n)}}{q}\Big).$$

Note that if for some $p|q$ the function $f_2 \,(\mathrm{mod}\, p)$ reduces to the zero function, then there are no values n with $(f(n),q) = 1$ so that the sum in our definition is empty and $V(a;q) = 0$. These sums $V(a;q)$ satisfy the relation (21). Using the Weil estimates for additive exponential sums modulo primes, one can take $G(p) = c_{f_1,f_2}$ for some integer constant depending only on the degree and number of zeros of f_1 and f_2 (in particular independent of p).

The problem of obtaining non-trivial estimates for

$$\sum_{q \leqslant x} |V(1;q)|$$

in this case has already been addressed in depth by Fouvry and Michel [6], and the special case of Kloosterman sums (namely, $f_1 = X^2 + 1$ and $f_2 = X$) is briefly mentioned by Hooley [10, §3]. One can extend some aspects of the work of Fouvry and Michel, but as this is of a different nature from the present paper, we defer further consideration to another note [13].

Acknowledgments

E.K. was partially supported by a DFG-SNF lead agency program grant (grant number 200020L_175755). K.S. is partially supported through a grant from the National Science Foundation (grant number DMS 1500237), and a Simons Investigator Grant from the Simons Foundation. This work was carried out while K.S. was a senior Fellow at the ETH Institute for Theoretical Studies, whom he thanks for their warm and generous hospitality.

We thank D.R. Heath–Brown and J-P. Serre for useful comments, P. Pollack for pointing out his paper [1] with V. Crişan and V. Kuperberg for sending us the note [14].

Appendix A. Conjectures modulo prime moduli and a function field analogue

As discussed in the introduction, one of the motivating problems is that of the distribution of the roots of polynomial congruences to prime moduli. This can be interpreted in (at least) two ways, depending whether one uses the same measures as in Theorem 1.2, or Hooley's measures as in Section 2.2. For completeness, we state formally the two potential conjectures (which are most likely both correct), and discuss a function field analogue that tends to indicate that, in this case, Hooley's measures are in some sense more natural.

Let $f \in \mathbf{Z}[X]$ be a monic irreducible polynomial of degree $\geqslant 2$, and let $\Pi_f(x)$ be the set of primes $p \leqslant x$ such that the number $\varrho_f(p)$ of roots of f modulo p is at least 1. Let Δ_p be the usual probability measure on the set of roots of f modulo p.

The first conjecture, analogue of the qualitative form of Theorem 1.2, is:

Conjecture A.1. *Let $f \in \mathbf{Z}[X]$ be a monic irreducible polynomial of degree $\geqslant 2$. Then the measures*

$$\frac{1}{|\Pi_f(x)|} \sum_{\substack{p \leqslant x \\ p \in \Omega}} \Delta_p$$

converge to the uniform measure as $x \to +\infty$.

Note that $|\Pi_f(x)| \sim c\pi(x)$ for some constant $c > 0$, namely the proportion of elements of the Galois group of the splitting field of f which have a fixed point, when viewed as permutations of the n roots of f.

Using Hooley's measures, the natural conjecture (which is stated in [4] for instance) is:

Conjecture A.2. *Let $f \in \mathbf{Z}[X]$ be a monic irreducible polynomial of degree $\geqslant 2$. Then the measures*

$$\frac{1}{\pi(x)} \sum_{\substack{p \leqslant x \\ p \in \Omega}} \varrho_f(p) \Delta_p$$

converge to the uniform measure.

Here the normalization by $\pi(x)$ is asymptotically correct, and corresponds to the fact that the average number of fixed points of a transitive permutation group is 1.

Remark A.3. Hrushovski also asked [11, §4.4] if the fractional parts of roots of polynomial congruences are equidistributed modulo primes p restricted to have $\varrho_f(p)$ equal to a fixed integer $r \geqslant 2$, in the case where the Galois group of the splitting field of f is cyclic. The

version modulo all squarefree q follows easily from Theorem 1.4, for all f and all $r \geqslant 2$ such that the Galois group of the splitting contains at least one permutation which has r fixed points when acting on the complex roots of f.

In order to determine which of the two conjectures is more natural, we look at a function field analogue.

Let $f \in \mathbf{Z}[X, Y]$ be a polynomial which is irreducible in $\mathbf{C}[X, Y]$, of degree $\geqslant 2$ with respect to Y and $\geqslant 1$ with respect to X.

For any prime p large enough, the reduction of f modulo p will be absolutely irreducible in $\mathbf{F}_p[X, Y]$; below we only consider such primes.

One analogue of looking at primes $\leqslant x$ is to consider irreducible polynomials π in $\mathbf{F}_p[X]$ of bounded degree. The roots of a polynomial congruence modulo a given prime correspond then to the roots in $k = \mathbf{F}_p[X]/\pi\mathbf{F}_p[X]$ of the polynomial $f \pmod{\pi}$, viewed as an element of $k[Y]$.

To simplify the discussion, we will look at polynomials π of degree 1, i.e., $\pi = X - x$ for $x \in \mathbf{F}_p$, but we will then let $p \to +\infty$ (this is possible since we started with a polynomial $f \in \mathbf{Z}[X, Y]$). Then, for a given $\pi = X - x$, we look at the roots y of $f \pmod{\pi}$ that belong to $\mathbf{F}_p[X]/(X - x)\mathbf{F}_p[X] \simeq \mathbf{F}_p$, i.e., we look at $y \in \mathbf{F}_p$ such that $f(x, y) = 0 \in \mathbf{F}_p$.

Now the Weyl sums to consider for the analogue of Conjecture A.1 are

$$\frac{1}{Z_p} \sum_{\substack{x \in \mathbf{F}_p \\ C_x \neq \emptyset}} \frac{1}{|C_x|} \sum_{\substack{y \in \mathbf{F}_p \\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big), \tag{23}$$

where

$$C_x = \{y \in \mathbf{F}_p \mid f(x, y) = 0\},$$
$$Z_p = |\{x \in \mathbf{F}_p \mid C_x \neq \emptyset\}|,$$

and those for the analogue of Conjecture A.2 are

$$\frac{1}{p} \sum_{x \in \mathbf{F}_p} \sum_{\substack{y \in \mathbf{F}_p \\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big), \tag{24}$$

both for $h \in \mathbf{Z}$ non-zero (it is a consequence of the Riemann Hypothesis for curves over finite fields that p is asymptotically the correct normalization here; this depends on the fact that f is absolutely irreducible).

As it turns out, the sums in (24) converge to 0 as $p \to +\infty$ essentially without further conditions, and those in (23) do so at least in considerable generality, but the argument is less straightforward in that case.

Convergence of (24). It is a standard fact (see e.g. [8]) that if f has degree $\geqslant 2$ with respect to Y, then as $p \to +\infty$, the fractional parts $(\{x/p\}, \{y, p\}) \in (\mathbf{R}/\mathbf{Z})^2$ of the

points $(x, y) \in C(\mathbf{F}_p)$ of the plane algebraic curve defined by the equation $f(x, y) = 0$ become equidistributed with respect to the uniform measure, and moreover, the Riemann Hypothesis for curves implies that

$$|C(\mathbf{F}_p)| = p + O(p^{1/2})$$

as $p \to +\infty$. This implies (more than) the convergence to 0 of the Weyl sums in (24).

Convergence of (23). We split the sum according to the value of $|C_x|$, which is an integer $\leqslant d = \deg_Y(f)$. We get

$$\frac{1}{Z_p} \sum_{\substack{x\in\mathbf{F}_p\\ C_x\neq\emptyset}} \frac{1}{|C_x|} \sum_{\substack{y\in\mathbf{F}_p\\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big) = \frac{1}{Z_p} \sum_{1\leqslant k\leqslant d} \frac{1}{k} \sum_{\substack{x\in\mathbf{F}_p\\ |C_x|=k}} \sum_{\substack{y\in\mathbf{F}_p\\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big).$$

Fix k. The characteristic function φ_k of the set of $x \in \mathbf{F}_p$ such that $|C_x| = k$ can be represented in the form

$$\varphi_k(x) = \sum_{j\in J} \alpha(k, j) t_j(x; p)$$

where J is a finite set and $\alpha(k, j)$ are complex coefficients, both of which are independent of p, and where $t_j(x; p)$ is a trace function modulo p of conductor bounded in terms of f only (more precisely, this formula holds for all x except possibly boundedly many exceptional values where the covering $\pi\colon C \to \mathbf{A}^1$ given by $(x, y) \to x$ is ramified, and it is obtained from Galois theory, the set J being the set of irreducible representations of the Galois group $G_\pi \subset S_d$ of π, and $\alpha(k, j)$ the Fourier coefficients of the characteristic function of those $\sigma \in G_\pi$ with precisely k fixed points; see, e.g., [5, §10.2] for similar computations). Hence

$$\sum_{\substack{x\in\mathbf{F}_p\\ |C_x|=k}} \sum_{\substack{y\in\mathbf{F}_p\\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big) = \sum_{j\in J} \alpha(k, j) \sum_{x\in\mathbf{F}_p} t_j(x; p) \sum_{\substack{y\in\mathbf{F}_p\\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big) + O(1).$$

But the function

$$g(x) = \sum_{\substack{y\in\mathbf{F}_p\\ f(x,y)=0}} e\Big(\frac{hy}{p}\Big)$$

is itself a trace function with conductor bounded in terms of f only, and moreover it is lisse and pure of weight 1 on an open dense subset of $\mathbf{A}^1$.

Now, note that for p large enough, all the trace functions t_j are associated to sheaves that are everywhere tamely ramified (see again [5, §10.2]). On the other hand, if we assume that f is monic with respect to X, then one can check[2] that for p large enough, the

[2] We thank W. Sawin for clarifying this argument.

monodromy representation at infinity of the sheaf underlying g is totally wildly ramified. Consequently, no geometrically irreducible component of g can then be geometrically isomorphic to any of the trace functions t_j. Applying then the Riemann Hypothesis over finite fields (in a form like [12, Prop. 1.8]), we have

$$\sum_{x\in\mathbf{F}_p} t_j(x;p)g(x) \ll p^{1/2},$$

where the implied constant depends only on f (because the conductors of t_j and g are bounded in terms of f).

A similar argument using the Riemann Hypothesis shows that $Z_p \gg p$ as $p \to +\infty$, and hence we deduce (generically at least) that the sums (23) tend to 0 as $p \to +\infty$.

Remark A.4. The condition that f is monic with respect to X is somewhat restrictive, and the convergence of (23) to 0 can be generalized to various other classes of polynomials. Since our goal is to illustrate the difference between the two types of sums, we do not attempt to discuss more general situations here.

References

[1] V. Crişan, P. Pollack, The smallest root of a polynomial congruence, Math. Res. Lett. 27 (1) (2020) 43–67.
[2] C. Dartyge, G. Martin, Exponential sums with reducible polynomials, Discrete Anal. (2019) 15, https://doi.org/10.19086/da.10793.
[3] R. de la Bretèche, G. Tenenbaum, Sur la conjecture de Manin pour certaines surfaces de Châtelet, J. Inst. Math. Jussieu 12 (2013) 759–819.
[4] W. Duke, J. Friedlander, H. Iwaniec, Equidistribution of roots of a quadratic congruence to prime moduli, Ann. Math. 141 (1995) 423–441.
[5] É. Fouvry, E. Kowalski, Ph. Michel, Algebraic twists of modular forms and Hecke orbits, Geom. Funct. Anal. 25 (2015) 580–657, https://doi.org/10.1007/s00039-015-0310-2.
[6] É. Fouvry, Ph. Michel, Sommes de modules de sommes exponentielles, Pac. J. Math. 209 (2003) 261–288.
[7] A. Granville, P. Kurlberg, Poisson statistics via the Chinese Remainder Theorem, Adv. Math. 218 (2008) 2013–2042.
[8] A. Granville, I. Shparlinski, A. Zaharescu, On the distribution of rational functions along a curve over $\mathbf{F}_p$ and residue races, J. Number Theory 112 (2005) 216–237.
[9] R.R. Hall, On pseudo-polynomials, Mathematika 18 (1971) 71–77.
[10] C. Hooley, On the distribution of the roots of polynomial congruences, Mathematika 11 (1964) 39–49.
[11] E. Hrushovski, Ax's theorem with an additive character, https://arxiv.org/abs/1911.01096.
[12] E. Kowalski, Ph. Michel, W. Sawin, Stratification and averaging for exponential sums: bilinear forms with generalized Kloosterman sums, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) XXI (2020) 1453–1530, https://doi.org/10.2422/2036-2145.201805_002.
[13] E. Kowalski, K. Soundararajan, Exponential sums, twisted multiplicativity, and moments, in preparation.
[14] V. Kuperberg, A note on pseudo-polynomials divisible only by a sparse set of primes, https://arxiv.org/abs/2006.02527.
[15] G. Martin, S. Sitar, Erdös–Turán with a moving target, equidistribution of roots of reducible quadratics and Diophantine quadruples, Mathematika 57 (2011) 1–29.
[16] A. Tóth, Roots of quadratic congruences, Int. Math. Res. Not. (2000) 719–739.