# POISSON SPACING STATISTICS FOR VALUE SETS OF POLYNOMIALS

PÄR KURLBERG

Abstract. If $f$ is a non-constant polynomial with integer coefficients and $q$ is an integer, we may regard $f$ as a map from $\mathbf{Z}/q\mathbf{Z}$ to $\mathbf{Z}/q\mathbf{Z}$. We show that the distribution of the (normalized) spacings between consecutive elements in the image of these maps becomes *Poissonian* as $q$ tends to infinity along any sequence of square free integers such that the mean spacing modulo $q$ tends to infinity.

## 1. Introduction

Let $f$ be a non-constant polynomial with integer coefficients. Given an integer $q$, we may regard $f$ as a map from $\mathbf{Z}/q\mathbf{Z}$ to $\mathbf{Z}/q\mathbf{Z}$, and the image of this map will be denoted the *image of $f$ modulo $q$*. The purpose of this paper is to investigate the distribution of spacings between consecutive elements in the image of $f$ modulo $q$, as $q$ tends to infinity along *square free* integers. The main emphasis will be placed on the highly composite case, i.e., by letting $q$ tend to infinity in such a way that the number of prime factors of $q$ also tends to infinity, but we will also present some results for $q$ prime that might be of independent interest.

The case $f(x) = x^2$ and $q$ prime was investigated by Davenport. In [6, 7] he proved that the probability of two consecutive squares being spaced $h$ units apart tends to $2^{-h}$ as $q \to \infty$. We may interpret this as if spacings between squares modulo prime $q$ behave like gaps between heads in a sequence of fair coin flips.

The case $f(x) = x^2$ and $q$ highly composite was studied by Rudnick and the author in [16, 15]. If we let $\omega(q)$ be the number of distinct prime factors of $q$, then the number of squares modulo $q$ equals $\prod_{p|q} \frac{p+1}{2}$, and

*Date*: Dec 3, 2007.
Author supported in part by the Göran Gustafsson Foundation, the National Science Foundation (DMS 0071503), the Royal Swedish Academy of Sciences, and the Swedish Research Council.

the mean spacing between *squares* modulo $q$ is given by

$$s_q = \frac{q}{\prod_{p|q} \frac{p+1}{2}} = 2^{\omega(q)} \prod_{p|q} \frac{p}{p+1}.$$

Hence $s_q \to \infty$ as $\omega(q) \to \infty$, so we would expect that the probability of two squares being 1 unit apart vanishes as $\omega(q) \to \infty$, and it is thus natural to normalize so that the mean spacing is one. A natural statistical model for the spacings is then given by looking at random points in $\mathbf{R}/\mathbf{Z}$; for independent uniformly distributed numbers in $\mathbf{R}/\mathbf{Z}$, the *normalized* spacings are said to be Poissonian. In particular, the distribution $P(s)$ of spacings between consecutive points is that of a Poisson arrival process, i.e., $P(s) = e^{-s}$, and the joint distribution of $l$ consecutive spacings is a product of $l$ independent exponential random variables (see [8]). Using Davenport's result together with the heuristic that "primes are independent", it seems reasonable to expect that the distribution of the normalized spacings between squares modulo $q$ becomes Poissonian in the limit $s_q \to \infty$, and the main result of [16] is that this is indeed the case for squarefree $q$ (the general case is treated in [15].)

What can be said about *general* polynomials $f \in \mathbf{Z}[x]$? For $p$ *prime*, let

$$\Omega_p := \{t \in \mathbf{Z}/p\mathbf{Z} : t \equiv \bar{f}(x_0) \mod p \text{ for some } x_0 \in \mathbf{Z}/p\mathbf{Z}\}$$

be the *image of $f$ modulo $p$*, where $\bar{f}$ denotes the reduction of $f$ modulo $p$. Given an integer $k \geq 2$ and integers $h_1, h_2, \ldots, h_{k-1}$, let

$$N_k((h_1, h_2, \ldots, h_{k-1}), p) := |\{t \in \Omega_p : t + \overline{h_1}, \ldots, t + \overline{h_{k-1}} \in \Omega_p\}|$$

be the counting function for the number of $k$-tuples of elements in the image of the form $t, t + \overline{h_1}, \ldots, t + \overline{h_{k-1}}$, where $\overline{h_i} \in \mathbf{Z}/p\mathbf{Z}$ denotes the reduction of $h_i$ modulo $p$. The average gap between the elements in $\Omega_p$, or the *mean spacing modulo $p$* is then, for *general* $f$, given by

$$s_p := p/|\Omega_p|,$$

and the "probability" of an element being in the image is $1/s_p$. Thus, if the conditions $t \in \Omega_p, t+\overline{h_1} \in \Omega_p, \ldots, t+\overline{h_{k-1}} \in \Omega_p$ are independent, we would expect $N_k((h_1, h_2, \ldots, h_{k-1}), p)$ to be of size $p/s_p^k$, and a natural analogue of Davenport's result is then that

$$N_k((h_1, h_2, \ldots, h_{k-1}), p) = p/s_p^k + o(p), \tag{1}$$

as $p \to \infty$ provided that $0, h_1, \ldots, h_{k-1}$ are distinct modulo $p$. In [11] Granville and the author proved that

$$N_k((h_1, h_2, \ldots, h_{k-1}), p) = p/s_p^k + O_{f,k}(\sqrt{p}) \tag{2}$$

holds if $f$ is a Morse polynomial and $0, h_1, \ldots, h_{k-1}$ are distinct modulo $p$. Using this, Poisson spacings for the image of Morse polynomials in the highly composite case follows from the following criteria (see [11], Theorem 1): *Assume that there exists* $\epsilon > 0$ *such that for each integer* $k \geq 2$,

$$N_k((h_1, h_2, \ldots, h_{k-1}), p) = \frac{p}{s_p^k}\left(1 + O_k((1 - s_p^{-1})p^{-\epsilon})\right) \tag{3}$$

*provided that* $0, h_1, h_2, \ldots, h_{k-1}$ *are distinct mod* $p$*. If* $s_p = p^{o(1)}$ *for all primes* $p$*, then the spacings modulo* $q$ *become Poisson distributed as the mean spacing modulo* $q$ *tends to infinity*[a].

What about non-Morse polynomials? Rather surprisingly, it turns out that (1) does not hold for all polynomials — that is, there are polynomials such that the spacing distribution of the image modulo $p$ is *not consistent with the coin flip model!* (That is, independent coin flips where the probability of heads is given by $|\Omega_p|/p$.) For example, in [11] it was shown that for $f(x) = x^4 - 2x^2$,

$$N_2((h_1), p) = \begin{cases} 2/3 \cdot \frac{p}{s_p^2} + O(\sqrt{p}) & \text{if } h_1 \equiv \pm 1 \mod p,\ p \equiv 1 \mod 4 \\ 4/3 \cdot \frac{p}{s_p^2} + O(\sqrt{p}) & \text{if } h_1 \equiv \pm 1 \mod p,\ p \equiv 3 \mod 4 \\ \frac{p}{s_p^2} + O(\sqrt{p}) & \text{if } h_1 \not\equiv \pm 1, 0 \mod p \end{cases}$$

Hence the assumptions in (3) are violated. However, we can prove that (2) holds for most values of $(h_1, \ldots, h_{k-1})$:

**Theorem 1.** *Let* $f \in \mathbf{Z}[x]$ *be a non-constant polynomial. Given a prime* $p$*, let*

$$R_p := \{\bar{f}(\xi) : \bar{f}'(\xi) = 0, \xi \in \overline{\mathbb{F}_p}\} \tag{4}$$

*be the set of critical values modulo* $p$*. If the sets*[b] $R_p, R_p - \overline{h_1}, R_p - \overline{h_2}, \ldots, R_p - \overline{h_{k-1}}$ *are pairwise disjoint*[c]*, then*

$$N_k((h_1, h_2, \ldots, h_{k-1}), p) = p/s_p^k + O_{f,k}(\sqrt{p}). \tag{5}$$

In other words, the analogue of Davenport's result holds for all but $O(p^{k-2})$ elements in $(\mathbf{Z}/p\mathbf{Z})^{k-1}$. Allowing for overlap between two translates of the set of critical values, we also have the following weaker upper bound on $N_k((h_1, h_2, \ldots, h_{k-1}), p)$:

---

[a] In [11] it was also shown that for the image of a Morse polynomial, the mean spacing modulo $q$ tends to infinity as $\omega(q) \to \infty$.

[b] By $R_p - \overline{h_j}$ we mean the set $\{r - \overline{h_j} : r \in R_p\}$.

[c] In the case $f(x) = x^2$ this condition is equivalent to $0, h_1, \ldots, h_{k-1}$ being distinct modulo $p$. However, for general polynomials (including the case of Morse polynomials), the two conditions are *not* equivalent.

**Proposition 2.** *Let $p$ be a prime. There exists a constant $C_0 < 1$, only depending on $f$, with the following property: if the sets*

$$(R_p \cup R_p - \overline{h_1}), R_p - \overline{h_2}, \dots, R_p - \overline{h_{k-1}}$$

*are pairwise disjoint and $h_1 \not\equiv 0 \mod p$, then*

$$N_k((h_1, h_2, \dots, h_{k-1}), p) \leq \frac{C_0}{s_p^{k-1}} \cdot p + O_{f,k}(\sqrt{p})$$

*unless $f$ is a permutation polynomial* [d] *modulo $p$.*

It turns out that these two results are enough to obtain Poisson spacings in the highly composite case. However, rather than studying the spacings directly, we proceed by determining the *$k$-level correlation functions*. Given a square free integer $q$ and a *general* polynomial $f \in \mathbf{Z}[x]$, let

$$\Omega_q := \{t \in \mathbf{Z}/q\mathbf{Z} : t \equiv \bar{f}(x_0) \mod q \text{ for some } x_0 \in \mathbf{Z}/q\mathbf{Z}\}$$

be the *image of $f$ modulo $q$* (here $\bar{f}$ denotes the reduction of $f$ modulo $q$), and let

$$s_q := q/|\Omega_q|$$

be the *mean spacing modulo $q$*. By the Chinese Remainder Theorem (since $q$ is square free), $|\Omega_q| = \prod_{p|q} |\Omega_p|$, where $p$ ranges over all prime divisors of $q$, and thus $s_q = \prod_{p|q} s_p$. Given $\mathbf{h} = (h_1, h_2, \dots, h_{k-1}) \in \mathbf{Z}^{k-1}$, put

$$N_k(\mathbf{h}, q) := |\{t \in \Omega_q : t + \overline{h_1}, t + \overline{h_2}, \dots, t + \overline{h_{k-1}} \in \Omega_q\}|$$

For $X \subset \mathbf{R}^{k-1}$, the *$k$-level correlation function* is then given by

$$R_k(X, q) := \frac{1}{|\Omega_q|} \sum_{\mathbf{h} \in s_q X \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q)$$

The main result of this paper is then the following:

**Theorem 3.** *Let $q$ be square free, $k \geq 2$ an integer, and let $X \subset \mathbf{R}^{k-1}$ be a convex set with the property that $(x_0, x_1, \dots, x_{k-1}) \in X$ implies that $x_i \neq x_j$ if $i \neq j$. Then the $k$-level correlation function of the image of $f$ modulo $q$ satisfies*

$$R_k(X, q) = \mathrm{vol}(X) + O_{f,k}\left(s_q^{-1/2+o(1)} + C_0^{\omega(q)(1-o(1))}\right)$$

*as $s_q \to \infty$, where $C_0 < 1$ is the constant given in Proposition 2.*

[d] $f$ is said to be a permutation polynomial modulo $p$ if $|\Omega_p| = p$.

Using a standard inclusion-exclusion argument (see [16], appendix A for details), this implies that the spacing statistics are Poissonian. In particular we have the following:

**Theorem 4.** *For $q$ tending to infinity along a sequence of square free integers such that $s_q \to \infty$, the limiting (normalized) spacing distribution*[e] *of the image of $f$ modulo $q$ is given by $P(t) = \exp(-t)$. Moreover, for any integer $k \geq 2$, the limiting joint distribution of $k$ consecutive spacings is a product $\prod_{i=1}^{k} \exp(-t_i)$ of $k$ independent exponential variables.*

1.1. **Some remarks on the mean spacing.** We note that the only way for which $s_p = 1$ for all primes $p$ is if $f(x)$ is of degree one. However, there are nonlinear polynomials $f$ such that $s_p = 1$ for infinitely many primes. For example, if $f(x) = x^3$ and we take $q$ to be a product of primes $p \equiv 2 \mod 3$, then $s_p = 1$ for all $p|q$, and $s_q = \prod_{p|q} s_p = 1$ clearly does not tend to infinity. On the other hand, if $\deg(f) > 1$, there is always a positive density set of primes $p$ such that $s_p > 1$. Moreover, if $f$ is not a permutation polynomial modulo $p$, Wan has shown [17] that

$$|\Omega_p| \leq p - \frac{p-1}{\deg(f)}. \tag{6}$$

Thus, for primes $p$ such that $s_p > 1$, $s_p$ is in fact uniformly bounded away from 1.

It is also worth noting that Birch and Swinnerton-Dyer have shown [1] that for $f$ Morse, $|\Omega_p| = c_f \cdot p + O_f(\sqrt{p})$ where $c_f < 1$ only depends on the degree of $f$, hence $s_p = 1/c_f + O(p^{-1/2})$ for all $p$, and thus $s_q \to \infty$ as $\omega(q) \to \infty$.

1.2. **Related results.** There are only a few other cases for which Poisson spacings have been proven. Notable examples are Hooley's result [12, 13] on invertible elements modulo $q$ under the assumption that the average gap $s_q = q/\phi(q)$ tends to infinity, and the work by Cobeli and Zaharescu [3] on spacings between primitive roots modulo $p$, again under the assumption that the average gap $s_p = (p-1)/\phi(p-1)$ tends to infinity. Recently, Cobeli,Vâjâitu, and Zaharescu [2] extended Hooley's results and showed that subsets of the form $\{x \mod q : x \in I_q, x^{-1} \in J_q\}$ have limiting Poisson spacings if the intervals $I_q, J_q$ have

[e] By normalized spacings we mean the following: with $0 \leq x_1 < x_2 < \cdots < x_{|\Omega_q|} < q$ being integer representatives of the image of $f$ modulo $q$, the spacings between consecutive elements are defined to be $\Delta_i = x_{i+1} - x_i$ for $1 \leq i < |\Omega_q|$, and $\Delta_{|\Omega_q|} = x_1 - x_{|\Omega_q|} + q$. The normalized spacings are then given by $\widetilde{\Delta_i} := \Delta_i / s_q$.

large lengths (more precisely, that $|I_q| \in [q^{1-(2/9(\log\log q)^{1/2})}, q]$, and $|J_q| \in [q^{1-1/(\log\log q)^2}, q]$) as $q$ tends to infinity along a subsequence of integers such that $q/\phi(q) \to \infty$.

1.3. **Acknowledgements.** The author would like to thank Juliusz Brzeziński, Andrew Granville, Dan Haran, Moshe Jarden, Zeév Rudnick, and Thomas J. Tucker for helpful discussions, and the anonymous referee for valuable suggestions on improving the exposition. The author is also very grateful to Peter Müller for a major simplification of the proof of Lemma 11.

## 2. Proof of Theorem 1

Allowing for a worse constant in the error term, we may assume that $p$ is large enough so that $p > \deg(f)$ and that $f(x)$ is not constant modulo $p$. We note that $N_k((h_1, \ldots, h_{k-1}), p)$ only depends on the reduction of $h_1, \ldots, h_{k-1}$ modulo $p$, so if $\mathbf{h} \in \mathbb{F}_p^{k-1}$ then $N_k(\mathbf{h}, p)$ is in fact well defined. To simplify the notation, the reduction of $h_1, \ldots, h_{k-1}$ modulo $p$ will also be denoted by $h_1, \ldots, h_{k-1}$ in this Section. Thus, given a non-constant polynomial $\bar{f} \in \mathbb{F}_p[x]$ and $k$ distinct elements $h_0 = 0, h_1, h_2, \ldots, h_{k-1} \in \mathbb{F}_p$, we wish to count the number of $t \in \mathbb{F}_p$ for which there exist $x_0, \ldots, x_{k-1} \in \mathbb{F}_p$ such that

$$\bar{f}(x_0) = t + h_0, \bar{f}(x_1) = t + h_1, \ldots, \bar{f}(x_{k-1}) = t + h_{k-1}.$$

For ease of notation, put

$$\mathbf{h} := (h_1, h_2, \ldots, h_{k-1}) \in \mathbb{F}_p^{k-1}.$$

Given $h \in \mathbb{F}_p$, define a polynomial $F_h \in \mathbb{F}_p[T][X]$ by

$$F_h(X, T) := \bar{f}(X) - (T + h).$$

Since the $T$-degree of $F_h$ is one, $F_h$ is irreducible, and thus

$$K_h = \mathbb{F}_p(T)[X]/(F_h(X, T))$$

is a field. Fix once and for all a separable closure $\overline{\mathbb{F}_p(T)}$ of $\mathbb{F}_p(T)$, and for $i = 0, \ldots, k-1$, choose embeddings of $K_{h_i}$ into $\overline{\mathbb{F}_p(T)}$, as well as an embedding of $\overline{\mathbb{F}_p}$ in $\overline{\mathbb{F}_p(T)}$. Further, let $L_h$ be the *Galois closure* of $K_h$ in $\overline{\mathbb{F}_p(T)}$, and let

$$G_h := \mathrm{Gal}(L_h/\mathbb{F}_p(T)) \tag{7}$$

be the Galois group of the field extension $L_h/\mathbb{F}_p(T)$. Since we assume that $p > \deg(f)$, all field extensions $L_h$ are separable, and no wild ramification can occur.

The following Lemma shows that $G_h$ and $L_h \cap \overline{\mathbb{F}_p}$ are independent of $h$.

**Lemma 5.** *Let $h \in \mathbb{F}_p$. Then $G_h \cong G_0$ and $L_h \cap \overline{\mathbb{F}_p} = L_0 \cap \overline{\mathbb{F}_p}$.*

*Proof.* Define a $\mathbb{F}_p$-linear automorphism $\sigma : \mathbb{F}_p[T] \to \mathbb{F}_p[T]$ by $\sigma(T) = T + h$. Since $\sigma(F_0) = F_h$ we may extend $\sigma$ to an isomorphism $\sigma' : L_0 \to L_h$. Moreover, given $\tau \in G_0$, and $\sigma'\tau(\sigma')^{-1} \in G_h$, the map $\tau \to \sigma'\tau(\sigma')^{-1}$ gives an isomorphism between $G_0$ and $G_h$.

Let $l_0 = L_0 \cap \overline{\mathbb{F}_p}$ and let $l_h = L_h \cap \overline{\mathbb{F}_p}$. Since $l_0/\mathbb{F}_p$ is normal, $l_0 = \sigma'(l_0) \subset L_h \cap \overline{\mathbb{F}_p} = l_h$, and the same argument for $(\sigma')^{-1}$ gives that $l_h \subset l_0$, hence $l_h = l_0$. □

Thus

$$l := L_0 \cap \overline{\mathbb{F}_p}$$

is the field of constants for $L_h$ for any $h \in \mathbb{F}_p$. Arguing as in the proof of Lemma 5 we obtain:

**Lemma 6.** *For $h \in \mathbb{F}_p$, let*

$$H_h := \mathrm{Gal}(L_h/l(T)).$$

*Then $H_h \cong H_0$.*

Our next goal is to obtain a criterion for linear disjointness for the field extensions $L_h/l(T)$ as $h$ varies.

**Lemma 7.** *Let $E_1, E_2$ be finite Galois extensions of $\mathbb{F}_p(T)$, both having the same constant field $l$, and degree smaller than $p$. If $E_1/l(T)$ and $E_2/l(T)$ have disjoint finite ramification, then $E_1 \cap E_2 = l(T)$ and hence $E_1$ and $E_2$ are linearly disjoint over $l(T)$. Furthermore, $l$ is the field of constants in the compositum $E_1E_2$.*

*Proof.* Let $E = E_1 \cap E_2$. By the assumption, $E/l(T)$ can only ramify at infinity. Moreover, the ramification must be tame. With $g_E$ denoting the genus of $E$, the Riemann-Hurwitz genus formula now gives

$$-2 \leq 2(g_E - 1) = [E : l(T)]2(0-1) + \sum_{\mathfrak{P}|\infty}(e(\mathfrak{P}/\infty) - 1)\deg(\mathfrak{P})$$

$$= -2[E : l(T)] + [E : l(T)] - \sum_{\mathfrak{P}|\infty}\deg(\mathfrak{P}) < -[E : l(T)]$$

and thus $[E : l(T)] < 2$.

As for the final assertion, we argue as follows: Let $m$ be the constant field of $E_1E_2$. The degree $[mE_1 : m(T)]$ is then equal to $[E_1 : l(T)]$, and similarly $[mE_2 : m(T)] = [E_2 : l(T)]$, and $m$ is the constant field of both $mE_1$ and $mE_2$. Applying the the first part of the Lemma to $mE_1$ and

$mE_2$, we find that $mE_1$ and $mE_2$ are linearly disjoint over $m(T)$, hence $[E_1E_2 : m(T)] = [mE_1 : m(T)] \cdot [mE_2 : m(T)] = [E_1 : l(T)] \cdot [E_2 : l(T)]$, which in turn equals $[E_1E_2 : l(T)]$. Hence $m(T) = l(T)$ and $m = l$. □

For $k \geq 2$, denote by

$$L^k := L_{h_0} L_{h_1} \dots L_{h_{k-1}}$$

the compositum of the fields $L_{h_0}, \dots, L_{h_{k-1}}$, and let

$$L^1 := L_{h_0} = L_0.$$

We now easily obtain the desired linear disjointedness criteria, and can also determine the field of constants in $L^k$.

**Proposition 8.** *If the sets $R_p, R_p - h_1, R_p - h_2, \dots, R_p - h_k$ are pairwise disjoint, then the field extensions $L_0/l(T), L_{h_1}/l(T), \dots, L_{h_{k-1}}/l(T)$ are linearly disjoint. Moreover, $l$ is the field of constants of $L^k$.*

*Proof.* Since $L_h$ is the Galois closure of $K_h$, both extensions, relative to $\mathbb{F}_p(T)$, ramify over the same primes. The assumption of pairwise disjointness of $R_p, R_p - h_1, \dots, R_p - h_{k-1}$ means that there is no common finite ramification among the fields $L_0, L_{h_1}, \dots L_{h_{k-1}}$. Hence by using Lemma 5 and applying Lemma 7 inductively, we find that $L_0, L_{h_1}, \dots L_{h_{k-1}}$ are linearly disjoint, and that $l$ is the field of constants in $L^k$. □

If $G = \mathrm{Gal}(E/\mathbb{F}_p(T))$ is the Galois group of a normal separable extension $E/\mathbb{F}_p(T)$ with constant field $l$, define (following Cohen, e.g., see [4, Section 1] or [5, Section 2])

$$G^* := \{\sigma \in G : \sigma|_{l(T)} = \mathrm{Frob}(l(T)/\mathbb{F}_p(T))\} \tag{8}$$

where $\mathrm{Frob}(l(T)/\mathbb{F}_p(T))$ is the canonical generator of $\mathrm{Gal}(l(T)/\mathbb{F}_p(T))$ given by $T \to T$ and $\alpha \to \alpha^p$ for all $\alpha \in l$. For $k \geq 2$, define a conjugacy class $\mathrm{Fix}_{k,\mathbf{h}} \subset \mathrm{Gal}(L^k/\mathbb{F}_p(T))^*$ by

$$\mathrm{Fix}_{k,\mathbf{h}} := \{\sigma \in \mathrm{Gal}(L^k/\mathbb{F}_p(T))^* : \\ \sigma \text{ fixes at least one root of } F_{h_i} \text{ for } i = 0, 1, \dots, k-1\}.$$

For $k = 1$ we define a conjugacy class $\mathrm{Fix}_1 \subset \mathrm{Gal}(L^1/\mathbb{F}_p(T))^*$ (note that there is no dependence on $\mathbf{h}$ and also recall that $L^1 = L_{h_0}$) by

$$\mathrm{Fix}_1 := \{\sigma \in \mathrm{Gal}(L^1/\mathbb{F}_p(T))^* : \sigma \text{ fixes at least one root of } F_{h_0}\}.$$

Given a finite separable extension $E$ of $\mathbb{F}_p(T)$, let $\mathfrak{O}_E$ denote the *integral closure* of $\mathbb{F}_p[T]$ in $E$. If $E/\mathbb{F}_p(T)$ is a Galois extension and $\mathfrak{M} \subset \mathfrak{O}_E$ is an unramified prime ideal lying above $\mathfrak{m} \subset \mathbb{F}_p[T]$, let $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m}) \in$

$\mathrm{Gal}(E/\mathbb{F}_p(T))$ denote the Frobenius automorphism. (In what follows, the use of $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m})$ implicitly signifies that $\mathfrak{M}/\mathfrak{m}$ is unramified.)

We can now relate $N_k(\mathbf{h}, p)$ to the number of degree one prime ideals in $\mathbb{F}_p[T]$ having a certain type of Frobenius action.

**Proposition 9.** *We have*

$$(9) \quad N_k(\mathbf{h}, p) = \\ = |\{\mathfrak{m} \subset \mathbb{F}_p[T] : \deg(\mathfrak{m}) = 1,\ \exists \mathfrak{M}|\mathfrak{m},\ \mathfrak{M} \subset \mathfrak{O}_{L^k},\ \mathrm{Frob}(\mathfrak{M}|\mathfrak{m}) \in \mathrm{Fix}_{k,\mathbf{h}}\}| + O_{k,f}(1).$$

*Proof.* Since the coordinate ring $\mathbb{F}_p[X_i, T]/(F_{h_i}(X_i, T))$ is easily seen to be isomorphic to $\mathbb{F}_p[X_i]$, we find that $\mathbb{F}_p[X_i, T]/(F_{h_i}(X_i, T))$ equals $\mathfrak{O}_{K_{h_i}}$, the integral closure of $\mathbb{F}_p[T]$ in $K_{h_i}$. Further, the condition that $t + h_i = \bar{f}(x_i)$ for $t, x_i \in \mathbb{F}_p$ is equivalent to a maximal ideal $\mathfrak{m}_i' = (T - t, X_i - x_i) \subset \mathfrak{O}_{K_{h_i}}$, of degree one, lying above the maximal ideal $\mathfrak{m} = (T - t) \subset \mathbb{F}_p[T]$. In terms of the Frobenius automorphism, assuming that $\mathfrak{m}$ does not ramify in $L^k$, this is equivalent to the existence of a prime ideal $\mathfrak{M} \subset \mathfrak{O}_{L^k}$ such that $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m})$ restricted to $L_{h_i}$ fixes one or more roots of $F_{h_i}$. Moreover, $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m})$ *must* take values in $\mathrm{Gal}(L^k/\mathbb{F}_p(t))^*$ since the action of $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m})$ restricted to $l(T)$ is given by $T \to T$ and $\alpha \to \alpha^p$ for all $\alpha \in l$. More generally, if $t = \bar{f}(x_0), t + h_1 = \bar{f}(x_1), \ldots, t + h_{k-1} = \bar{f}(x_{k-1})$ for $t, x_0, \ldots, x_{k-1} \in \mathbb{F}_p$ and $\mathfrak{m}$ does not ramify in $L^k$, this is equivalent to the restriction of $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m})$ to $L_{h_i}$ fixing at least one root of $F_{h_i}$ for all $i \in \{0, \ldots, k-1\}$, i.e., $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m}) \in \mathrm{Fix}_{k,\mathbf{h}}$. Since there are at most $O_{k,f}(1)$ ramified primes, the result follows.

□

Applying the Chebotarev density theorem (e.g., see [10], Proposition 5.16), we obtain

$$(10) \qquad N_k(\mathbf{h}, p) = \frac{|\mathrm{Fix}_{k,\mathbf{h}}|}{|\mathrm{Gal}(L^k/l(T))|} \cdot p + O_{k,f}(\sqrt{p})$$

Our next goal is to determine $|\mathrm{Fix}_{k,\mathbf{h}}| / |\mathrm{Gal}(L^k/l(T))|$.

**Lemma 10.** *Given $k \geq 2$, define*

$$C_k(\mathbf{h}, p) := \frac{|\mathrm{Fix}_{k,\mathbf{h}}|}{|\mathrm{Gal}(L^k/l(T))|},$$

*and*

$$C_1(p) := \frac{|\mathrm{Fix}_1|}{|\mathrm{Gal}(L^1/l(T))|}.$$

*Assume that $R_p, R_p - h_1, \ldots, R_p - h_{k-1}$ are pairwise disjoint. Then $C_k(\mathbf{h}, p) = C_1(p)^k$ where $C_1(p) = 1/s_p + O_f(p^{-1/2})$, and in particular*

$$C_k(\mathbf{h}, p) = 1/s_p^k + O_{f,k}(p^{-1/2}). \tag{11}$$

*Proof.* For simplicity, we consider only the case $k = 2$, and for ease of notation, let $\mathbf{h} = (h_1) = (h)$. The action of $\mathrm{Gal}(L^2/\mathbb{F}_p(T))$ on the roots of $F_0$ and $F_h$ allows us to identify $\mathrm{Gal}(L^2/\mathbb{F}_p(T))$ and $\mathrm{Gal}(L^2/l(T))$ with subgroups of $S_n \times S_n$, where $n = \deg(f)$. Moreover, since $L_0$ and $L_h$ are linearly disjoint over $l(T)$ and have isomorphic Galois groups, we may identify $\mathrm{Gal}(L^2/l(T)) \cong H_0 \times H_h$ with a subgroup of $S_n \times S_n$ in such a way that

$$H_0 \cong H' \times 1 \subset S_n \times 1 \subset S_n \times S_n$$

and

$$H_h \cong 1 \times H' \subset 1 \times S_n \subset S_n \times S_n$$

where $H' \cong H_0 \cong H_h$ and $H'$ is a subgroup of $S_n$.

Define a $\mathbb{F}_p$-linear map $\tau : \mathbb{F}_p(T) \to \mathbb{F}_p(T)$ by $\tau(T) = T + h$, and extend it to a map from $L_0$ to $L_h$. Given $\mu_1 \in G_0^*$ (recall (7) and (8) for the definition of $G_0$ and $G_0^*$) let $\mu_2 = \tau\mu_1\tau^{-1}$. Clearly $\mu_2 \in G_h$, and since $\mathrm{Gal}(l(T)/\mathbb{F}_p(T)) \cong \mathrm{Gal}(l/\mathbb{F}_p)$ is abelian, $\mu_1|_{l(T)} = \mu_2|_{l(T)}$ and hence $\mu_2 \in G_h^*$. Since $\tau$ gives a bijection between the roots of $F_0$ and $F_h$, we may label the roots in such a way that $\mu_1$ and $\mu_2$ correspond to the *same* element in $S_n$. Let us consider the possible extensions of $\mu_1, \mu_2$ to $L^2$. After making a fixed, but arbitrary choice, of extensions $\tilde{\mu_1}, \tilde{\mu_2}$ we find that all pairs extensions are of the form $(\delta\tilde{\mu_1}, \gamma\tilde{\mu_2})$ where $\delta \in H_h$ and $\gamma \in H_0$. Now, for any such pair of extensions, we have

$$\delta\tilde{\mu_1}(\gamma\tilde{\mu_2})^{-1} = \delta\tilde{\mu_1}\tilde{\mu_2}^{-1}\gamma^{-1} \in \mathrm{Gal}(L^2/l(T))$$

Since $\mathrm{Gal}(L^2/l(T)) \cong H_0 \times H_h$ we may choose $\gamma$ and $\delta$ in such a way that $\delta\tilde{\mu_1}\tilde{\mu_2}^{-1}\gamma^{-1} = 1$. In other words, it is possible to choose $\tilde{\mu_1}, \tilde{\mu_2}$ so that $\tilde{\mu_1} = \tilde{\mu_2}$.

Thus, there is an extension of $\mu \in G_0^*$ to an element $\tilde{\mu}$ of $\mathrm{Gal}(L^2/\mathbb{F}_p(T))^*$ in such a way that $\tilde{\mu}$ embeds diagonally when regarded as an element of $S_n \times S_n$, i.e., there exists $\sigma \in S_n$ such that $\tilde{\mu}$ corresponds to $(\sigma, \sigma) \in S_n \times S_n$. Now, all elements of $\mathrm{Gal}(L^2/\mathbb{F}_p(T))^*$, regarded as elements of $S_n \times S_n$, must be of the form $(\delta\sigma, \gamma\sigma) \in S_n \times S_n$ where $\delta, \gamma \in H'$. In particular, if we let $H'' \subset H'$ be the set of elements $\delta$ such that $\delta\sigma$ has at least one fix point, we find that

$$C_2(\mathbf{h}, p) = \frac{|H''|^2}{|\,\mathrm{Gal}(L^2/l(T))|} = \frac{|H''|^2}{|\,\mathrm{Gal}(L^1/l(T))|^2} = C_1(p)^2$$

since $\mathrm{Gal}(L^2/l(T)) \cong H_0 \times H_h$ and $H_h \cong H_0 = \mathrm{Gal}(L^1/l(T))$.

To determine $C_1(p)$, we note that

$$|\Omega_p| = p/s_p = |\{t \in \mathbb{F}_p \text{ for which there exists } x_0 \in \mathbb{F}_p \text{ such that } \bar{f}(x_0) = t\}|.$$

Arguing as in Proposition 9, we note that $\bar{f}(x_0) = t$ for $x_0, t \in \mathbb{F}_p$ means that for some $\mathfrak{M} \subset \mathfrak{O}_{L^1}$ lying above $\mathfrak{m} = (T-t) \subset \mathbb{F}_p[T]$, $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m}) \in \mathrm{Gal}(L^1/\mathbb{F}_p(t))$ will fix one or more roots of $F_{h_0}(X,T) = \bar{f}(X) - T$, i.e., $\mathrm{Frob}(\mathfrak{M}|\mathfrak{m}) \in \mathrm{Fix}_1$. Thus, after taking $O_f(1)$ ramified primes into account, we find that

$$p/s_p = |\{\mathfrak{m} \subset \mathbb{F}_p[T] : \\ \deg(\mathfrak{m}) = 1, \exists \mathfrak{M}|\mathfrak{m}, \mathfrak{M} \subset \mathfrak{O}_{L^1}, \mathrm{Frob}(\mathfrak{M}|\mathfrak{m}) \in \mathrm{Fix}_1\}| + O_f(1).$$

Again using the Chebotarev density theorem, we find that

$$p/s_p = C_1(p) \cdot p + O_f(\sqrt{p}) \tag{12}$$

and thus $C_1(p) = 1/s_p + O_f(1/\sqrt{p})$. ☐

From (10) and (11) we immediately obtain $N_k(\mathbf{h}, p) = p/s_p^k + O_{k,f}(\sqrt{p})$ and the proof of Theorem 1 is concluded.

## 3. Proof of Proposition 2

We will begin by giving a proof for the case $k = 2$, and then show how the general case can be reduced to this case. By allowing worse constants in the error terms as before, we may assume that $p > \deg(f)$ and that $f(x)$ is not constant modulo $p$.

### 3.1. The case $k = 2$.

We start by showing that the field extensions $K_0, K_h$ are linearly disjoint if $h \in \mathbb{F}_p^\times$.

**Lemma 11.** *Let $\bar{f}(X) \in \mathbb{F}_p[X]$ be a non-constant polynomial. If $h \in \mathbb{F}_p^\times$ and $\deg(\bar{f}) < p$, then $\bar{f}(X) - \bar{f}(Y) + h \in \mathbb{F}_p[X,Y]$ is absolutely irreducible.*

*Proof.* Write $\bar{f}(X) = \sum_{i=0}^{d} a_i X^i$ where $d = \deg(\bar{f})$ and $a_d \in \mathbb{F}_p^\times$. The case $d = 1$ is trivial. For $d > 1$ we argue as follows: Let $Z = X - Y$. Since $\bar{f}(X) - \bar{f}(Y) + h = (X-Y)G(X,Y) + h$, where $G(X,Y) \in \mathbb{F}_p[X,Y]$, it is enough to show that $Z \cdot G(Y+Z,Y) + h$ is irreducible in $\mathbb{F}_{p^k}[Y,Z]$ for arbitrary $k$. Now, $G(Y+Z,Y) = d \cdot a_d \cdot Y^{d-1} + A(Y,Z)$ where the $Y$-degree of $A(Y,Z)$ is at most $Y^{d-2}$. Letting $W = 1/Y$, we find that

$$Z \cdot G(Y+Z,Y) + h = W^{1-d}\left(Z \cdot \left(d \cdot a_d + W \cdot \tilde{A}(W,Z)\right) + h \cdot W^{d-1}\right)$$

where $\tilde{A}$ is the reciprocal polynomial of $A$ (with respect to the first variable). Regarding

$$Z \cdot \Big(d \cdot a_d + W \cdot \tilde{A}(W, Z)\Big) + h \cdot W^{d-1}$$

as a polynomial in $W$ with coefficients in $\mathbb{F}_{p^k}[Z]$, the result follows from Eisenstein's irreducibility criterion with respect to the prime ideal $(Z)$. □

**Remark.** *The above proof, due to Peter Müller* [14]*, in fact shows that* $\bar{f}(X) - \bar{f}(Y) + h$ *is absolutely irreducible as long as* $p$ *does not divide* $\deg(\bar{f})$*.*

Proposition 2 in the case $k = 2$ now immediately follows from the following Lemma and (10).

**Lemma 12.** *There exists* $C_0 < 1$*, only depending on* $f \in \mathbf{Z}[x]$*, with the following property: for all sufficiently large* $p$ *for which* $f$ *is not a permutation polynomial modulo* $p$*,*

$$C_2(\mathbf{h}, p) \leq C_0/s_p$$

*for all* $\mathbf{h} = (h_1 \mod p)$ *such that* $h_1 \not\equiv 0 \mod p$*.*

*Proof.* For $f \in \mathbf{Z}[x]$ fixed there are only finitely many possibilities for $\mathrm{Gal}(L^2/\mathbb{F}_p(T))$, hence $C_2(\mathbf{h}, p) = |\operatorname{Fix}_{2,\mathbf{h}}|/|\operatorname{Gal}(L^2/l(T))|$ can only take finitely many values. Thus, since $C_2(\mathbf{h}, p) \leq C_1(p) = 1/s_p + O_f(p^{-1/2})$ by (12), it is enough to show that $C_2(\mathbf{h}, p) = C_1(p)$ can only happen for finitely many primes $p$ (i.e., unless $f$ is a permutation polynomial modulo $p$.)

Given $a \in \mathbb{F}_p$, let $M(a) = |\{x_0 \in \mathbb{F}_p : \bar{f}(x_0) = a\}|$. Then

$$|\{x_0, y_0 \in \mathbb{F}_p : \bar{f}(x_0) = \bar{f}(y_0) + \overline{h_1}\}| = \sum_{a \in \mathbb{F}_p} M(a)M(a - \overline{h_1})$$

On the other hand, by Lemma 11, the algebraic set defined by $\bar{f}(x_0) = \bar{f}(y_0) + \overline{h_1}$ is an absolutely irreducible curve, and hence the Riemann hypothesis for curves (e.g., see [10], Theorem 4.9) gives that

$$|\{x_0, y_0 \in \mathbb{F}_p : \bar{f}(x_0) = \bar{f}(y_0) + \overline{h_1}\}| = p + O_f(\sqrt{p})$$

We have

$$|\{a : M(a) > 0\}| = |\{a : M(a - \overline{h_1}) > 0\}| = |\mathrm{Image}(\bar{f})| = p/s_p$$

Thus, if

$$N_2(\mathbf{h}, p) = |\{a \in \mathbb{F}_p : M(a) > 0, M(a - \overline{h_1}) > 0\}| =$$
$$C_2(\mathbf{h}, p) \cdot p + O_f(\sqrt{p}) = C_1(p) \cdot p + O_f(\sqrt{p}) = \frac{1}{s_p} \cdot p + O_f(\sqrt{p})$$

then, since $|\{a : M(a-\overline{h_1}) > 0\}| = |\mathrm{Image}(\bar{f})| = p/s_p$, we have

$$|\{a \in \mathbb{F}_p : M(a) = 0, M(a-\overline{h_1}) > 0\}| = O_f(\sqrt{p})$$

Therefore

$$\begin{aligned} p + O_f(\sqrt{p}) &= \sum_{a\in\mathbb{F}_p} M(a)M(a-\overline{h_1}) \\ &\geq \sum_{a\in\mathbb{F}_p:M(a)=1} M(a-\overline{h_1}) + 2\sum_{a\in\mathbb{F}_p:M(a)>1} M(a-\overline{h_1}) \\ &= \sum_{a\in\mathbb{F}_p:M(a)>0} M(a-\overline{h_1}) + \sum_{a\in\mathbb{F}_p:M(a)>1} M(a-\overline{h_1}) \\ &= \sum_{a\in\mathbb{F}_p} M(a-\overline{h_1}) + \sum_{a\in\mathbb{F}_p:M(a)>1} M(a-\overline{h_1}) - \sum_{a\in\mathbb{F}_p:M(a)=0} M(a-\overline{h_1}) \\ &= p + \sum_{a\in\mathbb{F}_p:M(a)>1} M(a-\overline{h_1}) - O_f(\sqrt{p}) \end{aligned}$$

and thus

$$\sum_{a\in\mathbb{F}_p:M(a)>1} M(a-\overline{h_1}) = O_f(\sqrt{p})$$

Hence

$$|\{a \in \mathbb{F}_p : M(a) > 1, M(a-\overline{h_1}) > 0\}| = O_f(\sqrt{p})$$

and we similarly obtain that

$$|\{a \in \mathbb{F}_p : M(a) > 0, M(a-\overline{h_1}) > 1\}| = O_f(\sqrt{p})$$

But then

$$\begin{aligned} p + O_f(\sqrt{p}) &= \sum_{a\in\mathbb{F}_p} M(a)M(a-\overline{h_1}) \\ &= |\{a \in \mathbb{F}_p : M(a) = M(a-\overline{h_1}) = 1\}| + O_f(\sqrt{p}) \end{aligned}$$

In other words, $M(a) = 1$ for all but $O_f(\sqrt{p})$ elements, which, by Wan's result (see (6), section 1.1), can only happen if $f$ is bijection once $p$ is sufficiently large.

□

3.2. **The case $k > 2$.** Here we return to the notational conventions of Section 2, in particular $h_0 = 0, h_1, \ldots, h_{k-1}$ denote elements of $\mathbb{F}_p$. Arguing as in the proof of Lemma 7, we find that the field extensions

$$(L_{h_0}L_{h_1})/l(T), L_{h_2}/l(T), \ldots, L_{h_{k-2}}/l(T), L_{h_{k-1}}/l(T)$$

are linearly disjoint since they have disjoint ramification. Hence there is an isomorphism

$$\operatorname{Gal}\left(L_{h_0}L_{h_1}\cdot\ldots\cdot L_{h_{k-1}}/l(T)\right) \\ \simeq \operatorname{Gal}\left(L_{h_0}L_{h_1}/l(T)\right) \times \operatorname{Gal}\left(L_{h_2}/l(T)\right) \times \ldots \times \operatorname{Gal}\left(L_{h_{k-1}}/l(T)\right)$$

Putting $\mathbf{h}' = (h_1)$ and arguing as in Lemma 10, we find that

$$\frac{|\operatorname{Fix}_{k,\mathbf{h}}|}{|\operatorname{Gal}(L^k/l(T))|} = \frac{|\operatorname{Fix}_{2,\mathbf{h}'}|}{|\operatorname{Gal}(L_{h_0}L_{h_1}/l(T))|} \cdot \frac{1}{s_p^{k-2}} = C_2(\mathbf{h}', p) \cdot \frac{1}{s_p^{k-2}}.$$

By Lemma 12, $C_2(\mathbf{h}', p) \leq C_0/s_p$ and by (10) the proof is complete.

## 4. Proof of Theorem 3

In what follows we will use the convention that $h_0 = 0$. For $\mathbf{h} = (h_1, \ldots, h_{k-1}) \in \mathbf{Z}^{k-1}$ fixed, it follows immediately from the Chinese Remainder Theorem that $N_k(\mathbf{h}, q)$ is multiplicative in $q$. The following Lemma shows that we may assume that $q$ is a product of primes $p$ for which $f$ is not a permutation polynomial modulo $p$, and hence that $s_p$ is uniformly bounded away from 1 for all $p|q$.

**Lemma 13.** *Given a square free integer $q$, write $q = q_1 q_2$ where*

$$q_1 = \prod_{\substack{p|q \\ |\Omega_p|<p}} p, \quad q_2 = \prod_{\substack{p|q \\ |\Omega_p|=p}} p$$

*Then*

$$R_k(X, q) = R_k(X, q_1)$$

*Proof.* If $p|q_2$ we have $s_p = p/|\Omega_p| = 1$ and $N_k(\mathbf{h}, p) = p$ for all $\mathbf{h} \in \mathbf{Z}^{k-1}$. Thus $s_q = s_{q_1} \cdot s_{q_1} = s_{q_1}$, and since for $\mathbf{h}$ fixed, $N_k(\mathbf{h}, q)$ is multiplicative, we find that $N_k(\mathbf{h}, q) = N_k(\mathbf{h}, q_1) \cdot q_2$. Thus

$$R_k(X, q) = \frac{1}{|\Omega_q|} \sum_{\mathbf{h} \in s_q X \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q) = \frac{q_2}{|\Omega_{q_1}||\Omega_{q_2}|} \sum_{\mathbf{h} \in s_q X \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q_1)$$
$$= \frac{1}{|\Omega_{q_1}|} \sum_{\mathbf{h} \in s_{q_1} X \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q_1) = R_k(X, q_1)$$

□

We also note the following easy consequence of Theorem 1.

**Lemma 14.** *Let $l$ be the largest integer such that $R_p - h_{i_1}, R_p - h_{i_2}, \ldots, R_p - h_{i_l}$ are pairwise disjoint for some choice of indices $0 \leq i_1, i_2, \ldots, i_l \leq k-1$ (recall that $h_0 = 0$). Then*

$$N_k((h_1, h_2, \ldots, h_{k-1}), p) \leq p/s_p^l + O_{f,k}(\sqrt{p})$$

*Proof.* If $\{h'_1, h'_2, \dots h'_{l-1}\}$ is a subset of $\{h_1, h_2, \dots, h_{k-1}\}$ then trivially

$$N_k((h_1, h_2, \dots, h_{k-1}), p) \leq N_l((h'_1, h'_2, \dots, h'_{l-1}), p)$$

and the Lemma follows from Theorem 1. □

4.1. **Some remarks on affine sets.** We will partition $\mathbf{Z}^{k-1}$ according to the size of the bounds on $N_k(\mathbf{h}, q) = \prod_{p|q} N_k(\mathbf{h}, p)$ given by Theorem 1 and Proposition 2. In order to do this, we need to introduce some notation: By an *affine set* $L \subset \mathbf{Z}^{k-1}$ we mean an integer translate of a lattice $L' \subset \mathbf{Z}^{k-1}$. We then define the rank, respectively discriminant, of $L$ as the rank, respectively discriminant[f], of $L'$. Similarly, we define $\operatorname{codim}(L)$ as $k-1$ minus the rank of $L$.

Let $R$ be the set of critical values of $f$, i.e.,

$$R := \{f(\xi) : f'(\xi) = 0, \xi \in \overline{\mathbf{Q}}\}$$

and recall that $R_p = \{f(\xi) : f'(\xi) = 0, \xi \in \overline{\mathbb{F}_p}\}$ is the set of critical values of $f$ modulo $p$. Let

$$\tilde{R} := R - R = \{\alpha - \beta : \alpha, \beta \in R\},$$

put

$$\tilde{R}_\infty := \tilde{R} \cap \mathbf{Z},$$

and let

$$\tilde{R}_p := (R_p - R_p) \cap \mathbb{F}_p.$$

If $R_p + h_i \cap R_p + h_j \neq \emptyset$ then $h_i - h_j \in \tilde{R}_p$, so the affine sets to be considered will be given by equations of the form

$$h_i - h_j = r, \quad r \in \tilde{R}_\infty \tag{13}$$

or congruences of the form

$$h_i - h_j \equiv r_p \mod p, \quad r_p \in \tilde{R}_p \tag{14}$$

We note that the bounds given by Theorem 1 and Proposition 2 only depends on the congruence class of $\mathbf{h}$, but we will treat the case of equality separately since $N_k(\mathbf{h}, p)$ will be large *for all* $p|q$ if $\mathbf{h}$ satisfies an equation of the form (13).

To ensure that the equations defining the affine sets are independent, we will need the following notions: Given

$$E \subset \{(i, j) : 0 \leq i < j \leq k-1\}$$

[f]By the discriminant of $L' \subset \mathbf{Z}^{k-1}$ we mean the index of $L'$ in $\mathbf{Z}^{k-1}$.

we may associate a graph $G(E)$ on the set of vertices $\{0,1,\dots,k-1\}$ by regarding $E$ as the set of edges, i.e., two nodes $i,j$ are connected by an edge if and only if $(i,j)\in E$. Let

$$\mathcal{AG} := \{E \subset \{(i,j) : 0 \le i < j \le k-1\} : \ G(E) \text{ is acyclic.}\}$$

be the collection of edge sets whose associated graphs are acyclic.

Given $E \in \mathcal{AG}$ and a map $\alpha : E \to \tilde{R}_\infty$, define an affine set

$$L(E,\alpha) := \{\mathbf{h} \in \mathbf{Z}^{k-1} : h_i - h_j = \alpha((i,j)) \text{ for all } (i,j) \in E.\}.$$

(with the usual convention that $h_0 = 0$). Note that $G(E)$ acyclic implies that the equations defining $L(E,\alpha)$ are independent. Further, given $E \in \mathcal{AG}$, let

$$\mathcal{L}(E) := \{L(E,\alpha) \text{ where } \alpha \text{ ranges over all maps } \alpha : E \to \tilde{R}_\infty\}$$

be the collection of affine sets defined by independent relations between $h_i$ and $h_j$ for all $(i,j) \in E$. We note that $\mathcal{L}(\emptyset)$ contains exactly one element, namely the full lattice $L(\emptyset,-) = \mathbf{Z}^{k-1}$. Moreover, if $L \in \mathcal{L}(E)$, then (since we assume that $E \in \mathcal{AG}$) $\operatorname{codim}(L) = |E|$, and if $\mathbf{h} \in L$, then Proposition 2 will, for *all* $p|q$, at best give the bound

$$N_k(\mathbf{h},p) \le C_0 \frac{p}{s_p^{k-|E|}} + O_{f,k}(\sqrt{p}).$$

(The bound will not hold if the components of $\mathbf{h}$ satisfies additional equations, i.e., if $\mathbf{h} \in L'$ for some $L' \in \mathcal{L}(E')$ such that $E' \supsetneq E$.)

Given $L(E,\alpha) \in \mathcal{L}(E)$, let

$$L^\times(E,\alpha) := \{\mathbf{h} \in L(E,\alpha) : \mathbf{h} \notin L(E',\alpha') \text{ for all } E' \supsetneq E,\ \alpha' : E' \to \tilde{R}_\infty\}$$

In particular, if $\mathbf{h} \in L^\times(E,\alpha)$, the components of $\mathbf{h}$ satisfy exactly $|E|$ independent equations of the form $h_i - h_j = r_{ij}$ where $r_{ij} \in \tilde{R}_\infty$.

We also need to keep track of similar relations, modulo $p$, between the components of $\mathbf{h}$. Thus, given $E_p \in \mathcal{AG}$ and $\alpha_p : E_p \to \tilde{R}_p$, define an affine set

$$L_p(E_p,\alpha_p) := \{\mathbf{h} \in \mathbf{Z}^{k-1} : h_i - h_j \equiv \alpha_p((i,j)) \mod p \text{ for all } (i,j) \in E_p\}.$$

We note that the rank of $L_p(E_p,\alpha_p)$ is $k-1$ and that the discriminant of $L_p(E_p,\alpha_p)$ is $p^{|E_p|}$, and if $\mathbf{h} \in L_p(E_p,\alpha_p)$, then Proposition 2 will at best give the bound

$$N_k(\mathbf{h},p) \le C_0 \frac{p}{s_p^{k-|E_p|}} + O_{f,k}(\sqrt{p}).$$

Now, given $E \in \mathcal{AG}$, let

$$\mathcal{L}_p(E) := \{L_p(E_p,\alpha_p) : E_p \in \mathcal{AG}, \alpha_p : E_p \to \tilde{R}_p, E_p \cap E = \emptyset, E_p \cup E \in \mathcal{AG}\}$$

and for $L_p \in \mathcal{L}_p(E)$, let

$$L_p^\times := \{\mathbf{h} \in L_p : \mathbf{h} \notin L_p' \text{ for all } L_p' \in \mathcal{L}_p(E_p'),\ E_p' \supsetneq E_p\}$$

If $\mathbf{h} \in L^\times \cap L_p^\times$ for $L \in \mathcal{L}(E)$ and $L_p = L_p(E_p, \alpha_p) \in \mathcal{L}_p(E)$, then $\mathbf{h} = (h_1, \dots, h_{k-1})$ (also recall that $h_0 = 0$) satisfies exactly $|E|$ independent *equations* of the form $h_i - h_j = r_{ij}$ where $r_{ij} \in \tilde{R}_\infty$, and exactly $|E_p|$ independent *congruences* of the $h_i - h_j \equiv r_{ij}' \mod p$ where $r_{ij}' \in \tilde{R}_p$, and furthermore, there is no overlap between the equations and congruences. The reason for keeping track of equalities and congruences separately is that if $\mathbf{h} \in L$ for $L \in \mathcal{L}(E)$ and $|E| > 0$, then the bounds given on $N_k(\mathbf{h}, p)$ given by Proposition 2 allows $N_k(\mathbf{h}, p)$ to deviate quite a bit from its mean value *for all* $p|q$. On the other hand, if we let $c$ be the product of primes $p|q$ for which the bounds are bad because of congruence conditions, rather than equalities, then we can bound the size of $c$ (see Lemma 18). We can now partition $\mathbf{Z}^{k-1}$ according to the size of the bounds on $N_k(h, p)$ given by Theorem 1 and Proposition 2:

**Lemma 15.** *Let* $L = L(E, \alpha)$, $L_p = L_p(E_p, \alpha_p) \in \mathcal{L}_p(E)$, *and assume that* $\mathbf{h} \in L^\times \cap L_p^\times$. *If* $|E| + |E_p| = 0$, *then*

$$N_k(\mathbf{h}, p) = s_p^{-k} \cdot p + O_{k,f}(p^{1/2}),$$

*whereas if* $k > |E| + |E_p| > 0$, *then*

$$N_k(\mathbf{h}, p) \le C_0 \cdot s_p^{|E|+|E_p|-k} \cdot p + O_{k,f}(p^{1/2}).$$

*where* $C_0 < 1$ *is as in Proposition 2.*

*Proof.* The first assertion follows immediately from Theorem 1 since $R_p + h_i \cap R_p + h_j \neq \emptyset$ implies that $h_i - h_j \in \tilde{R}_p$.

For the second assertion, we argue as follows: Since $\mathbf{h} = (h_1, h_2, \dots, h_{k-1}) \in L^\times \cap L_p^\times$ there are indices $i_1, i_2, \dots, i_{k-|E|-|E_p|}$ such that $h_{i_1} \neq h_{i_2}$ and

$$(R_p - h_{i_1} \cup R_p - h_{i_2}), R_p - h_{i_3}, \dots, R_p - h_{i_{k-|E|-|E_p|}}$$

are pairwise disjoint. Putting

$$\mathbf{h}' = (h_{i_2} - h_{i_1}, h_{i_3} - h_{i_1}, \dots, h_{i_{k-|E|-|E_p|}} - h_{i_1}),$$

the result follows from the bound for $N_k(\mathbf{h}', p)$ given by Proposition 2. □

However, partitioning $\mathbf{Z}^{k-1}$ according to the size of $N_k(\mathbf{h}, p)$ for individual prime factors $p|q$ is not quite enough; we need to partition $\mathbf{Z}^{k-1}$ according to the size of $N_k(\mathbf{h}, q) = \prod_{p|q} N_k(\mathbf{h}, p)$. Thus, let

$$\mathcal{L}_c(E) := \{L \cap (\cap_{p|c} L_p) : L \in \mathcal{L}(E),\ \forall p|c\ L_p \in \mathcal{L}_p(E) \setminus L_p(\emptyset, -)\}$$

(where $L_p(\emptyset, -) \in \mathcal{L}_p(E)$ is the maximal lattice, i.e., $L_p(\emptyset, -) = \mathbf{Z}^{k-1}$) and given

$$L_c = L \cap (\cap_{p|c} L_p) \in \mathcal{L}_c(E)$$

let

$$L_c^\times := L^\times \cap (\cap_{p|c} L_p^\times) \cap (\cap_{p|\frac{q}{c}} L_p^\times(\emptyset, -))$$

We can now partition $\mathbf{Z}^{k-1}$ into subsets $L_c^\times$, where $L_c \in \mathcal{L}_c(E)$, $E \in \mathcal{AG}$, and $c|q$. Moreover, as an immediate consequence of the definitions and Lemma 15, we obtain the following:

**Lemma 16.** *Assume that* $L_c = L \cap (\cap_{p|c} L_p(E_p, \alpha_p)) \in \mathcal{L}_c(E)$ *and that* $\mathbf{h} \in L_c^\times$. *If* $p \nmid c$, *then*

$$N_k(\mathbf{h}, p) = s_p^{-k} \cdot p + O_{k,f}(p^{1/2}).$$

*If* $p \mid c$, *then*

$$N_k(\mathbf{h}, p) \leq C_0 \cdot s_p^{|E|+|E_p|-k} \cdot p + O_{k,f}(p^{1/2}).$$

*where* $C_0 < 1$ *is as in Proposition 2.*

Using the previous Lemma we can now bound sums of the form $\sum_{\mathbf{h} \in s_q X \cap L_c^\times} N_k(\mathbf{h}, q)$.

**Lemma 17.** *If*

$$L_c = L \cap (\cap_{p|c} L_p(E_p, \alpha_p)) \in \mathcal{L}_c(E),$$

*then*

$$|\{\mathbf{h} \in s_q X \cap L_c^\times\}| \leq |\{\mathbf{h} \in s_q X \cap L_c\}| \ll_{k,f,X} \frac{s_q^{k-|E|-1}}{c} + s_q^{k-|E|-2}$$

*Moreover, if* $\mathbf{h} \in L_c^\times$, *then*

$$\frac{N_k(\mathbf{h}, q)}{q/s_q} \ll \prod_{p|c} \left( \frac{s_p^{|E|+|E_p|}}{s_p^{k-1}} + O_{k,f}(p^{-1/2}) \right) \cdot \prod_{p|\frac{q}{c}} \left( C_0 \cdot \frac{s_p^{|E|}}{s_p^{k-1}} + O_{k,f}(p^{-1/2}) \right)$$

*In particular,*

$$\sum_{\mathbf{h} \in s_q X \cap L_c^\times} \frac{N_k(\mathbf{h}, q)}{q/s_q} \ll s_c^{k-1} C_0^{-\omega(c)} \left(\frac{1}{s_q} + \frac{1}{c}\right) \cdot C_0^{\omega(q)} \cdot \prod_{p|q} \left(1 + O_{k,f}(p^{-1/2})\right) \tag{15}$$

*Proof.* The first assertion follows from the Lipschitz principle[g] (e.g., see Lemma 16 in [16]) since $L_c$ is a translate of a lattice with discriminant (relative to $L$) divisible by $c$. The second assertion follows from Lemma 16. Thus

$$\sum_{\mathbf{h}\in s_q X\cap L_c^\times} \frac{N_k(\mathbf{h},q)}{q/s_q} \ll \prod_{p|c}\left(\frac{s_p^{|E_p|}}{p} + O_{k,f}(p^{-3/2})\right)\cdot\prod_{p|\frac{q}{c}}\left(C_0 + O_{k,f}(p^{-1/2})\right)$$

$$+\frac{1}{s_q}\prod_{p|c}\left(s_p^{|E_p|} + O_{k,f}(p^{-1/2})\right)\cdot\prod_{p|\frac{q}{c}}\left(C_0 + O_{k,f}(p^{-1/2})\right)$$

$$\ll C_0^{-\omega(c)}\left(\frac{s_c^{k-1}}{c} + \frac{s_c^{k-1}}{s_q}\right)\cdot C_0^{\omega(q)}\cdot\prod_{p|q}\left(1 + O_{k,f}(p^{-1/2})\right)$$

□

Since the bound in (15) is not useful for large $c$, we will also need the following:

**Lemma 18.** *Let $d$ be the degree of the field extension $\mathbf{Q}(\tilde{R})/\mathbf{Q}$. If $L_c \in \mathcal{L}_c(E)$ for some $E \in \mathcal{AG}$ and $s_q X \cap L_c^\times \neq \emptyset$ then*

$$c \ll_{X,\tilde{R}} s_q^{d\binom{k}{2}|\tilde{R}|}.$$

*Moreover, there exist a constant $D$, only depending on $k$ and $f$, such that*

$$|\mathcal{L}_c(E)| \ll_{k,f} D^{\omega(c)}.$$

*Proof.* We first assume that all elements of $\tilde{R}$ are algebraic integers. Let $B$ be the ring of integers in $\mathbf{Q}(\tilde{R})$. For each prime $p|q$ chose a prime $\mathfrak{P}_p \subset B$ lying above $p$, so that we may regard any element in $\tilde{R}_p$ as the image of an element in $\tilde{R}$ under the reduction map $B \to B/\mathfrak{P}_p$.

For $0 \leq i < j \leq k-1$, $r \in \tilde{R}$, and $\mathbf{h} \in L_c^\times$, let

$$\gamma_{i,j,r}(\mathbf{h}) = \prod_{p: h_i - h_j \equiv r \mod \mathfrak{P}_p} p$$

Then $c$ divides

$$\prod_{\substack{0\leq i<j\leq k-1\\ r\in\tilde{R}: h_i-h_j\neq r}} \gamma_{i,j,r}(\mathbf{h})$$

[g]Actually, we have to be a little careful: if we embed $L$ into $\mathbf{Z}^{k-1-|E|}$ and apply the Lipschitz principle, there is an implicit constant in the bound that will depend on $L$. However, the estimate is uniform since $L$ only can be chosen in $O_k(1)$ ways.

Since $h_i - h_j - r \equiv 0 \mod \mathfrak{P}_p$ for all $p$ dividing $\gamma_{i,j,r}$, we find that $\gamma_{i,j,r}$ divides $\mathrm{N}_{\mathbf{Q}}^{\mathbf{Q}(\tilde{R})}(h_i - h_j - r)$. Moreover, if $\mathbf{h} \in s_q X$, then $|h_i - h_j| \ll_X s_q$, thus

$$\mathrm{N}_{\mathbf{Q}}^{\mathbf{Q}(\tilde{R})}(h_i - h_j - r) \ll_{f,X} s_q^d$$

and hence

$$c \leq \prod_{\substack{0 \leq i < j \leq k-1 \\ r \in \tilde{R} : h_i - h_j \neq r}} \mathrm{N}_{\mathbf{Q}}^{\mathbf{Q}(\tilde{R})}(h_i - h_j - r) \ll_{k,f,X} s_q^{d|\tilde{R}|\binom{k}{2}}$$

(Note that $\mathrm{N}_{\mathbf{Q}}^{\mathbf{Q}(\tilde{R})}(h_i - h_j - r) \neq 0$ since $h_i - h_j - r \neq 0$).

In case $\tilde{R}$ contains elements that are not algebraic integers, we can find an integer $m$, only depending on $\tilde{R}$, such that all elements of $m \cdot \tilde{R} = \{m \cdot r : r \in \tilde{R}\}$ are algebraic integers, and apply the above argument to $m \cdot \tilde{R}$ and $m\mathbf{h}$ (for primes $p$ not dividing $m$, but since $c$ is square free this just makes the constant worse by a power of $(c, m) \leq m$, which is $O(1)$.)

The second assertion follows upon noting that there are $O_{k,f}(1)$ possible choices of $E_p$ and $\alpha_p$ for each $p|c$.

□

4.2. **Conclusion.** We can now write $\mathbf{Z}^{k-1}$ as a disjoint union of sets $L^\times$ where $L$ ranges over all elements in $\cup_{E \in \mathcal{AG}} \mathcal{L}(E)$, and hence $R_k(X, q)$ equals

$$\frac{1}{|\Omega_q|} \sum_{\mathbf{h} \in s_q X \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q) = \frac{1}{|\Omega_q|} \sum_{E \in \mathcal{AG}} \sum_{L \in \mathcal{L}(E)} \sum_{\mathbf{h} \in s_q X \cap L^\times} N_k(\mathbf{h}, q) \tag{16}$$

The term corresponding to $E = \emptyset$ in (16) will give the main contribution (note that if $E = \emptyset$, then $L = L_\infty(E, -) = \mathbf{Z}^{k-1}$.) Let

$$X' := \{\mathbf{h} \in X : h_i - h_j \notin \tilde{R}_\infty \text{ for } 0 \leq i < j \leq k-1\}$$

where we as usual use the convention that $h_0 = 0$. Then

$$s_q X \cap L^\times = s_q X' \cap \mathbf{Z}^{k-1}$$

Note that $X'$ is just $\mathbf{R}^{k-1}$ with some hyperplanes removed, so if $X$ is convex, we can write $X'$ as a finite union of convex sets. We now rewrite (16) as follows:

$$\frac{1}{|\Omega_q|} \sum_{\mathbf{h} \in s_q X \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q) = \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q) + \mathrm{Error}_1$$

where

$$\text{Error}_1 := \frac{1}{|\Omega_q|} \sum_{E\in\mathcal{AG},|E|>0} \sum_{L\in\mathcal{L}(E)} \sum_{\mathbf{h}\in s_q X\cap L^\times} N_k(\mathbf{h},q)$$

and the main term is given by

$$\sum_{\mathbf{h}\in s_q X'\cap\mathbf{Z}^{k-1}} N_k(\mathbf{h},q) \tag{17}$$

We begin by showing that $\text{Error}_1 = o(1)$ as $\omega(q)\to\infty$.

**Lemma 19.** *As $\omega(q)\to\infty$,*

$$\mathit{Error}_1 = \frac{1}{|\Omega_q|} \sum_{\substack{E\in\mathcal{AG}\\|E|>0}} \sum_{L\in\mathcal{L}(E)} \sum_{\mathbf{h}\in s_q X\cap L^\times} N_k(\mathbf{h},q) \ll C_0^{\omega(q)(1-o(1))}.$$

*Proof.* Given $E\in\mathcal{AG}$ with $|E|>0$, we find that

$$\frac{1}{|\Omega_q|} \sum_{L\in\mathcal{L}(E)} \sum_{\mathbf{h}\in s_q X\cap L^\times} N_k(\mathbf{h},q) = \frac{1}{q/s_q} \sum_{c|q} \sum_{L_c\in\mathcal{L}_c(E)} \sum_{\mathbf{h}\in s_q X\cap L_c^\times} N_k(\mathbf{h},q) \tag{18}$$

which, by Lemmas 17 and 18 is

$$\ll C_0^{\omega(q)}\cdot \prod_{p|q}\left(1+O(p^{-1/2})\right) \sum_{\substack{c|q\\ c\ll s_q^{d\binom{k}{2}|\tilde{R}|}}} D^{\omega(c)} s_c^{k-1} C_0^{-\omega(c)}\left(\frac{1}{s_q}+\frac{1}{c}\right) \tag{19}$$

Now,

$$\sum_{\substack{c|q\\ c\ll s_q^{d\binom{k}{2}|\tilde{R}|}}} D^{\omega(c)} s_c^{k-1} C_0^{-\omega(c)}\frac{1}{c} \ll \prod_{p|q}(1+O(1/p))$$

and, for any $\delta>0$,

$$\frac{1}{s_q}\sum_{\substack{c|q\\ c\ll s_q^{d\binom{k}{2}|\tilde{R}|}}} D^{\omega(c)} s_c^{k-1} C_0^{-\omega(c)} \ll \frac{1}{s_q^{1-\delta d\binom{k}{2}|\tilde{R}|}} \sum_{c|q} \frac{s_c^{k-1}C_0^{-\omega(c)}}{c^\delta}$$

$$\ll \frac{1}{s_q^{1-\delta d\binom{k}{2}|\tilde{R}|}} \prod_{p|q}\left(1+O(1/p^\delta)\right) \ll \frac{1}{s_q^{1-\delta d\binom{k}{2}|\tilde{R}|-o(1)}}$$

Thus, taking $\delta = 1/(2d\binom{k}{2}|\tilde{R}|)$, we find that (19) is

$$\ll C_0^{\omega(q)} \cdot \prod_{p|q} \left(1 + O(p^{-1/2})\right) \cdot \left( \frac{1}{s_q^{1/2-o(1)}} + \prod_{p|q} \left(1 + O(p^{-1})\right) \right)$$
$$\ll C_0^{\omega(q)} \cdot \prod_{p|q} \left(1 + O(p^{-1/2})\right) = C_0^{\omega(q)(1-o(1))}$$

Since there are $O(1)$ possible choices of $L \in \mathcal{L}(E)$ for $E$ fixed, and $E$ ranges over a finite number of subsets, we find that (18) is $C_0^{\omega(q)(1-o(1))}$. □

We proceed by rewriting the main term in terms of a divisor sum. For $p$ prime and $\mathbf{h} \in \mathbf{Z}^{k-1}$, let

$$\varepsilon_k(\mathbf{h}, p) = \frac{s_p^{k-1} \cdot N_k(\mathbf{h}, p)}{|\Omega_p|} - 1,$$

so that we may write

$$N_k(\mathbf{h}, p) = \frac{|\Omega_p|}{s_p^{k-1}} (1 + \varepsilon_k(\mathbf{h}, p))$$

(recall that $s_p = p/|\Omega_p|$.) Further, for $d > 1$ a square free integer, put

$$\varepsilon_k(\mathbf{h}, d) = \prod_{p|d} \varepsilon_k(\mathbf{h}, p)$$

and, to make $\varepsilon_k$ multiplicative in the second parameter, set $\varepsilon_k(\mathbf{h}, 1) = 1$ for all $h$. Since $N_k(\mathbf{h}, q)$ is multiplicative, we then have

$$N_k(\mathbf{h}, q) = \prod_{p|q} \frac{1}{s_p^{k-1}} |\Omega_p| \left(1 + \varepsilon_k(\mathbf{h}, p)\right) = \frac{|\Omega_q|}{s_q^{k-1}} \sum_{d|q} \varepsilon_k(\mathbf{h}, d) \tag{20}$$

The following Lemma shows that the average of $\varepsilon_k(\mathbf{h}, d)$, over a full set of residues modulo $d$, equals zero if $d > 1$.

**Lemma 20.** *If $d > 1$ then*

$$\sum_{\mathbf{h} \in (\mathbf{Z}/d\mathbf{Z})^{k-1}} \varepsilon_k(\mathbf{h}, d) = 0$$

*Proof.* Since $\varepsilon_k(\mathbf{h}, d)$ is multiplicative it is enough to show that

$$\sum_{\mathbf{h} \in (\mathbf{Z}/p\mathbf{Z})^{k-1}} \varepsilon_k(\mathbf{h}, p) = 0$$

for $p$ prime, and because

$$N_k(\mathbf{h}, p) = \frac{1}{s_p^{k-1}} |\Omega_p| (1 + \varepsilon_k(\mathbf{h}, p))$$

it is enough to show that

$$\sum_{\mathbf{h} \in (\mathbf{Z}/p\mathbf{Z})^{k-1}} N_k(\mathbf{h}, p) = \frac{1}{s_p^{k-1}} |\Omega_p| p^{k-1} = |\Omega_p|^k$$

But $\sum_{\mathbf{h} \in (\mathbf{Z}/p\mathbf{Z})^{k-1}} N_k(\mathbf{h}, p)$ equals the number of $k$-tuples of elements from $\Omega_p$, and hence $\sum_{\mathbf{h} \in (\mathbf{Z}/p\mathbf{Z})^{k-1}} N_k(\mathbf{h}, p) = |\Omega_p|^k$. □

We will also need the following bound:

**Lemma 21.** *We have*

$$\sum_{\mathbf{h} \in (\mathbf{Z}/d\mathbf{Z})^{k-1}} |\varepsilon_k(\mathbf{h}, d)| \ll d^{k-3/2+o(1)}$$

*Proof.* Since the sum is multiplicative in $d$, it is enough to show that

$$\sum_{\mathbf{h} \in (\mathbf{Z}/p\mathbf{Z})^{k-1}} |\varepsilon_k(\mathbf{h}, p)| \ll p^{k-3/2}$$

for $p$ prime. By Theorem 1, $|\varepsilon_k(\mathbf{h}, p)| \ll p^{-1/2}$ for all but $O(p^{k-2})$ residues modulo $p$, and for the remaining residues we have $|\varepsilon_k(\mathbf{h}, p)| = O_{k,f}(1)$. Thus

$$\sum_{\mathbf{h} \in (\mathbf{Z}/p\mathbf{Z})^{k-1}} |\varepsilon_k(\mathbf{h}, p)| \ll p^{k-1} p^{-1/2} + p^{k-2} \ll p^{k-3/2}$$

□

We now find that the main term (17) equals

$$\frac{1}{|\Omega_q|} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} N_k(\mathbf{h}, q) = \frac{1}{s_q^{k-1}} \sum_{d|q} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} \varepsilon_k(\mathbf{h}, d)$$

$$= \frac{1}{s_q^{k-1}} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} 1 + \text{Error}_2$$

where

$$\text{Error}_2 := \frac{1}{s_q^{k-1}} \sum_{\substack{d|q \\ d>1}} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} \varepsilon_k(\mathbf{h}, d)$$

and the modified main term is

$$\frac{1}{s_q^{k-1}} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} 1 = \frac{1}{s_q^{k-1}} \left( \mathrm{vol}(s_q X') + O(s_q^{k-2}) \right) = \mathrm{vol}(X) + O(1/s_q).$$

We conclude by showing that $\mathrm{Error}_2 = o(1)$ as $s_q \to \infty$.

**Lemma 22.** *As $s_q \to \infty$, we have*

$$\mathit{Error}_2 = \frac{1}{s_q^{k-1}} \sum_{\substack{d|q \\ d>1}} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} \varepsilon_k(\mathbf{h}, d) \ll s_q^{-1/2+o(1)} \tag{21}$$

*Proof.* In order to show that $\mathrm{Error}_2$ is small, we split the divisor sum in two parts according to the size of $d$.

*Small $d$:* We first consider $d \leq s_q^T$ where $T \in (0,1)$ is to be chosen later. A point $\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}$ is contained in a unique cube $C_{\mathbf{h},d} \subset \mathbf{R}^{k-1}$ of the form

$$C_{\mathbf{h},d} = \{(x_1, x_2, \ldots, x_{k-1}) : dt_i \leq x_i < d(t_i+1), t_i \in \mathbf{Z},\ i = 1, 2, \ldots, k-1\}$$

We say that $\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}$ is a *$d$-interior* point of $s_q X'$ if $C_{\mathbf{h},d} \subset s_q X'$, and if $C_{\mathbf{h},d}$ intersects the boundary of $s_q X'$, we say that $h$ is a *$d$-boundary point* of $s_q X'$.

By Lemma 20, the sum over the $d$-interior points is zero, and hence

$$\frac{1}{s_q^{k-1}} \sum_{\substack{d|q \\ 1<d\leq s_q^T}} \sum_{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1}} \varepsilon_k(\mathbf{h}, d) = \frac{1}{s_q^{k-1}} \sum_{\substack{d|q \\ 1<d\leq s_q^T}} \sum_{\substack{\mathbf{h} \in s_q X' \cap \mathbf{Z}^{k-1} \\ \mathbf{h} \text{ is } d\text{-boundary point}}} \varepsilon_k(\mathbf{h}, d) \tag{22}$$

Since $s_q X'$ is a union of convex sets, the number of cubes $C_{\mathbf{h},d}$ intersecting the boundary of $s_q X'$ is $\ll (s_q/d)^{k-2}$, and hence (22) is

$$\ll \frac{1}{s_q^{k-1}} \sum_{\substack{d|q \\ 1<d\leq s_q^T}} (s_q/d)^{k-2} \sum_{\mathbf{h} \in (\mathbf{Z}/d\mathbf{Z})^{k-1}} |\varepsilon_k(\mathbf{h}, d)|$$

$$= \frac{1}{s_q} \sum_{\substack{d|q \\ 1<d\leq s_q^T}} \frac{1}{d^{k-2}} \sum_{\mathbf{h} \in (\mathbf{Z}/d\mathbf{Z})^{k-1}} |\varepsilon_k(\mathbf{h}, d)| \tag{23}$$

which by Lemma 21 is, for any $\alpha > 1/2$,

$$\ll \frac{1}{s_q} \sum_{\substack{d|q \\ 1<d\leq s_q^T}} d^{1/2+o(1)} \leq s_q^{\alpha T-1} \sum_{d|q} d^{1/2-\alpha+o(1)} \ll s_q^{\alpha T-1+o(1)}$$

since

$$\sum_{d|q} d^{-\epsilon} = \prod_{p|q}(1+p^{-\epsilon}) = s_q^{o(1)}$$

if $\epsilon > 0$ (recall that $s_p$ is assumed to be uniformly bounded away from 1 and $s_q = \prod_{p|q} s_p$.)

*Large $d$:* We now consider

$$\frac{1}{s_q^{k-1}} \sum_{\substack{d|q \\ d>s_q^T}} \sum_{\mathbf{h}\in s_q X' \cap \mathbf{Z}^{k-1}} \varepsilon_k(\mathbf{h}, d) \tag{24}$$

Given $\mathbf{h}$ and $d$, let $c$ be the largest divisor of $d$ such that $\mathbf{h} \in L_c$ for some $L_c \in \mathcal{L}_c(L)$. Then

$$\varepsilon_k(\mathbf{h}, d) \ll \frac{s_c^{k-1}}{(d/c)^{1/2-o(1)}}$$

by Lemma 16. Hence, for $E \in \mathcal{AG}$ fixed,

$$\sum_{L\in\mathcal{L}(E)} \sum_{\mathbf{h}\in s_q X\cap L^\times} \varepsilon_k(\mathbf{h}, d) \ll \sum_{c|d} \sum_{L_c\in\mathcal{L}_c(E)} \sum_{\mathbf{h}\in s_q X\cap L_c^\times} |\varepsilon_k(\mathbf{h}, d)|$$

$$\ll \sum_{c|d} \frac{s_c^{k-1}}{(d/c)^{1/2-o(1)}} \sum_{L_c\in\mathcal{L}_c(E)} \sum_{\mathbf{h}\in s_q X\cap L_c^\times} 1$$

which by Lemmas 17 and 18 is

$$\ll s_q^{k-1} \cdot d^{-1/2+o(1)} \cdot \sum_{\substack{c|d \\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} s_c^{k-1} c^{1/2-o(1)} D^{\omega(c)} \left(\frac{1}{c} + \frac{1}{s_q}\right) \tag{25}$$

Now,

$$\sum_{\substack{c|d \\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} \frac{s_c^{k-1} c^{1/2-o(1)} D^{\omega(c)}}{c} \ll \sum_{\substack{c|d \\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} c^{-1/2+o(1)} \ll s_q^{o(1)}$$

and similarly

$$\frac{1}{s_q} \sum_{\substack{c|d \\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} s_c^{k-1} c^{1/2-o(1)} D^{\omega(c)} \ll \frac{1}{s_q} \sum_{\substack{c|d \\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} c^{1/2+o(1)}$$

Thus (24) is

$$(26)\qquad \ll \frac{s_q^{k-1}}{s_q^{k-1}} \sum_{\substack{d|q\\ d>s_q^T}} \left( \frac{s_q^{o(1)}}{d^{1/2-o(1)}} + \frac{1}{s_q d^{1/2-o(1)}} \sum_{\substack{c|d\\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} c^{1/2+o(1)} \right)$$

$$= s_q^{o(1)} \sum_{\substack{d|q\\ d>s_q^T}} d^{-1/2+o(1)} + \frac{1}{s_q} \sum_{\substack{d|q\\ d>s_q^T}} \frac{1}{d^{1/2-o(1)}} \sum_{\substack{c|d\\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} c^{1/2+o(1)}$$

Now, for any $\beta \in (0, 1/2)$,

$$\sum_{\substack{d|q\\ d>s_q^T}} d^{-1/2+o(1)} \ll \sum_{d|q} d^{-1/2+o(1)} \left(\frac{d}{s_q^T}\right)^{\beta}$$

$$\ll s_q^{-\beta T} \sum_{d|q} d^{\beta-1/2+o(1)} \ll s_q^{-\beta T+o(1)}.$$

Similarly, for any $\gamma > 0$,

$$\sum_{\substack{c|d\\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} c^{1/2+o(1)} \ll s_q^{\gamma d\binom{k}{2}|\tilde{R}|} \sum_{c|d} c^{1/2-\gamma+o(1)} \ll s_q^{\gamma d\binom{k}{2}|\tilde{R}|} d^{1/2-\gamma+o(1)}$$

and thus

$$\sum_{\substack{d|q\\ d>s_q^T}} \frac{1}{d^{1/2-o(1)}} \sum_{\substack{c|d\\ c\ll s^{d\binom{k}{2}|\tilde{R}|}}} c^{1/2+o(1)} \ll s_q^{\gamma d\binom{k}{2}|\tilde{R}|} \sum_{d|q} d^{-\gamma+o(1)} \ll s_q^{\gamma d\binom{k}{2}|\tilde{R}|+o(1)}$$

Hence (26) is

$$\ll s_q^{-\beta T+o(1)} + s_q^{-1+\gamma d\binom{k}{2}|\tilde{R}|+o(1)} \ll s_q^{-1/2+o(1)}$$

if we take $T = 1 - o(1)$, $\beta = 1/(2T) - o(1)$, and $\gamma = 1/(2d\binom{k}{2}|\tilde{R}|)$. Thus, with $\alpha = 1/2 + o(1)$ (to bound the contribution from small $d$), we find that

$$\text{Error}_2 = \frac{1}{s_q^{k-1}} \sum_{\substack{d|q\\ d>1}} \sum_{\mathbf{h}\in s_q X' \cap \mathbf{Z}^{k-1}} \varepsilon_k(\mathbf{h}, d) \ll s_q^{-1/2+o(1)}$$

□

## References

[1] B. J. Birch and H. P. F. Swinnerton-Dyer. Note on a problem of Chowla. *Acta Arith.*, 5:417–423 (1959), 1959.

[2] C. Cobeli, M. Vâjâitu, and A. Zaharescu. Distribution of gaps between the inverses mod $q$. *Proc. Edinb. Math. Soc. (2)*, 46(1):185–203, 2003.

[3] C. Cobeli and A. Zaharescu. On the distribution of primitive roots mod $p$. *Acta Arith.*, 83(2):143–153, 1998.

[4] S. D. Cohen. The distribution of polynomials over finite fields. *Acta Arith.*, 17:255–271, 1970.

[5] S. D. Cohen. The distribution of polynomials over finite fields. II. *Acta Arith.*, 20:53–62, 1972.

[6] H. Davenport. On the distribution of quadratic residues (mod $p$). *Jour. London Math. Soc.*, 6:49–54, 1931.

[7] H. Davenport. On character sums in finite fields. *Acta Math.*, 71:99–121, 1939.

[8] W. Feller. *An introduction to probability theory and its applications. Vol. I.* John Wiley & Sons Inc., New York, 1968.

[9] M. D. Fried. Variables separated polynomials, the genus 0 problem and moduli spaces. In *Number theory in progress, Vol. 1 (Zakopane-Kościelisko, 1997)*, pages 169–228. de Gruyter, Berlin, 1999.

[10] M. D. Fried and M. Jarden. *Field arithmetic*, volume 11 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1986.

[11] A. Granville and P. Kurlberg. Poisson statistics via the chinese remainder theorem. *Submitted*. Preprint at http://www.arxiv.org/abs/math.NT/0412135.

[12] C. Hooley. On the difference between consecutive numbers prime to $n$. II. *Publ. Math. Debrecen*, 12:39–49, 1965.

[13] C. Hooley. On the difference between consecutive numbers prime to $n$. III. *Math. Z.*, 90:355–364, 1965.

[14] P. Müller. *Personal communication.*

[15] P. Kurlberg. The distribution of spacings between quadratic residues. II. *Israel J. Math.*, 120(A):205–224, 2000.

[16] P. Kurlberg and Z. Rudnick. The distribution of spacings between quadratic residues. *Duke Math. J.*, 100(2):211–242, 1999.

[17] D. Q. Wan. A $p$-adic lifting lemma and its applications to permutation polynomials. In *Finite fields, coding theory, and advances in communications and computing (Las Vegas, NV, 1991)*, pages 209–216. Dekker, New York, 1993.

*E-mail address*: `kurlberg@math.kth.se`

Department of Mathematics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden