On k-Degenerate Distance Graphs in Finite Fields and a Class of 2-Degenerate Distance Graphs in the Integer Lattice

by

Peter Woolfitt

(Under the Direction of Neil Lyall and Ákos Magyar)

Abstract

We give conditions on dimension and density to guarantee isometric copies of k-degenerate distance graphs in the model setting of subsets of finite fields. These arguments are then translated and used in conjunction with the circle method to guarantee isometric copies of arbitrarily large dilates of a class of 2-degenerate distance graphs in subsets of the integer lattice.

Index words: [geometric Ramsey theory, arithmetic combinatorics, finite fields, distance graphs, circle method]

On k-Degenerate Distance Graphs in Finite Fields and a Class of 2-Degenerate Distance Graphs in the Integer Lattice

by

Peter Woolfitt

B.S., Georgia Institute of Technology, 2015

A Dissertation Submitted to the Graduate Faculty of the
University of Georgia in Partial Fulfillment of the Requirements for the
Degree.

Doctor of Philosophy

Athens, Georgia

2023

©2023
Peter Woolfitt
All Rights Reserved

On k-Degenerate Distance Graphs in Finite Fields and a Class of 2-Degenerate Distance Graphs in the Integer Lattice

by

Peter Woolfitt

Major Professors: Neil Lyall
Ákos Magyar

Committee: Giorgis Petridis
Paul Pollack

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
December 2023

Acknowledgments

I have so many people to thank. First and foremost come my advisors Neil Lyall and Ákos Magyar for their guidance, kindness, and unreasonable levels of patience. I will be eternally grateful for all of their help and support.

I am also extremely grateful for my other committee members: Paul Pollack and Giorgis Petridis. Both Paul and Giorgis taught me many things through many courses over many years. They are wonderful people to boot.

I would not be where I am today without the dedication of my high school teacher, the iconic and terrifying Jessica Hunt. I was unaware at the time how hard Dr. Hunt fought for us to have an accelerated high school mathematics program. The other mentor to whom I owe a special debt of thanks is Josephine Yu. I felt burned out from my last year of undergraduate, but during my subsequent gap year, Josephine included me in a small group project which reminded me how fun and rewarding the pursuit of mathematics could be.

Thank you to my undergraduate friends at Georgia Tech: Charles Wang, Gautam Goel, Dylan McKay, Alex Bettadapur, Ada Hackett, and Jay Hackett. You all gave me the support and friendship I didn't know I needed.

My seniors at UGA: Luca Schaffler, Daniel McKenzie, Eric Perkerson, Ernest Guico, Zerotti Woods, Kübra Benli, Noah Lebowitz-Lockard, and Nate Zbacnik. You were unendingly kind, generous with your time, and tolerant of my silly questions. I thank you all.

I am also very grateful for all of the amazing math graduate students who each in their own way helped me through my journey. A big thank you particularly to Alex Newman, Kirsten Morris, Erik Schreyer, Terrin Warren, Nicole Song, Zack Garza, Paco Adajar, Dustin Kasser, and Swaroop Hegde.

Additionally, I am deeply grateful for the efforts of the UGA staff members Christy McDonald and Laura Ackerley. They solved all of my administrative problems, and Christy only chided me mildly the third and fourth times I was months late turning in receipts. As I could not find a better place, I suppose here I shall also thank Mo Hendon for many engaging conversations about the sorts of small cute problems which lie close to my heart. Perhaps it is appropriate for an oddball like Mo to be thanked in an odd place like this.

I have always been someone who has needed long conversations to understand myself and the world around me. Makoto Suwama, Matthew Haggerty, David Galban, Sarah Blackwell, Komal Agrawal, Devashi Gulati, Casia Siegel, Ian Campbell, and Sayan Chattopadhyay have been there for me throughout puzzle competitions, road trips, long walks, and many hours of late-night chatting. Thank you for all of your insights and the foundation they give me.

Finally, the support from my family has been invaluable. Thank you to my parents for your constant love and support. Stephen and Hanh, thank you for coming whenever we were down and cheering us with food and laughter. Hannah, thank you for being here with me through everything.

Contents

Chapter 1

Introduction

1.1 Background

The upper Banach density of a measurable subset A of $\mathbb{R}^d$ is given by

$$\delta^*(A) := \lim_{N\to\infty} \sup_{\underline{t}\in\mathbb{R}^d} \frac{|A \cap (\underline{t} + \mathcal{C}_N)|}{|\mathcal{C}_N|},$$

where $|\cdot|$ denotes the Lebesgue measure and $\mathcal{C}_N$ is a d-dimensional cube with side length N. The way to think about a set A with positive upper Banach density is that there are arbitrarily large cubes in $\mathbb{R}^d$ on which A is appreciably dense. Positive upper Banach density is a relatively weak condition, but sets with positive upper Banach density enjoy many of the properties of a set with positive upper density. [1]

[1] An example of a set with upper Banach density 1 and upper density 0 is a union of solid balls of radius n each centered a distance n^3 away from the origin, one for each n. For any N there is a cube of side length N somewhere in $\mathbb{R}^d$ on which the density of the set is 1, but as N increases, the location of that cube varies.

In [FKW90], Furstenberg, Katznelson, and Weiss show that it is possible to realize all large distances in measurable subsets of $\mathbb{R}^2$ of positive upper Banach density. In particular, if $A \subset \mathbb{R}^2$ is measurable, then they demonstrate the existence of a $\lambda_0 = \lambda_0(A)$ such that

$$[\Lambda_0, \infty) \subset \{|\underline{x} - \underline{y}| \,:\, \underline{x}, \underline{y} \in A\}.$$

Bourgain reattains this result and a pinned variant in [Bou86] using techniques from Fourier analysis. More recently, tools from ergodic theory have been brought to bear on related problems. For example a surprising theorem of Ziegler [Zie06] states that for a given point configuration S in $\mathbb{R}^d$ with $d \geq 2$ and a set E of positive upper Banach density, there exists a threshold λ_0 such that for any $\lambda \geq \lambda_0$ and $\delta > 0$ there is an isometric copy of λS in the δ-neighborhood of E.

In [Mag08], Magyar established a version for distances in subsets of $\mathbb{Z}^d$ of positive upper Banach density when $d \geq 5$.[2] Using ergodic theory techniques, Bulinski [Bul17] established results for distance trees in $\mathbb{Z}^d$. In [LM20b], Lyall and Magyar obtained simplified[3] results for distance trees through a type of discrete spherical maximal function as first studied by Magyar, Wainger, and Stein in [MSW02]. Iosevich and Parshall investigate the analogous problem for distance graphs of bounded degree in finite fields [IP19]. In chapter 2 of this thesis, we generalize the results of Iosevich in Parshall to k-degenerate distance graphs in finite fields.[4] In chapters 3 and 4, find a generalization of the result on trees to a class of 2-degenerate graphs in $\mathbb{Z}^d$.

[2] Jacobi's four-square theorem shows that are arbitrarily large lattice spheres in dimension 4 containing only 24 lattice points. When $d \geq 5$ the lattice points on spheres obey better regularity properties e.g. the number of lattice points on the sphere of radius r is proportional to r^{d-2}.

[3] And in some sense optimal.

[4] k-degeneracy is a sort of one-sided boundedness.

1.2 k-Degenerate Distance Graphs

A *distance graph* $\Gamma = (V, E)$ is a weighted graph where the vertices are taken to be points in some ambient space, and the edge weights represent a distance between vertices in that space. When $V \subseteq \mathbb{R}^d$, the edge weights are taken to be Euclidean distance as in [LM20b] and [LM20a]. When $V \subseteq \mathbb{F}_q^d$ in analogy to the Euclidean setting we define the distance between two points $v_1, v_2 \in \mathbb{F}_p^d$ as $(v_2 - v_1) \cdot (v_2 - v_1)$. With this definition of distance, finite fields are a model setting for arguments one hopes to carry to other settings. Much work has been in this model setting see e.g. [IP19].

A *k-degenerate graph* is a graph $\Gamma = (V, E)$ for which there exists an ordering of the vertices such that each vertex is adjacent to at most k prior vertices in the ordering. This is exactly the class of graphs which can be constructed by successively adding a vertex with at most k edges, which is the key property for the inductive structure of our proofs.

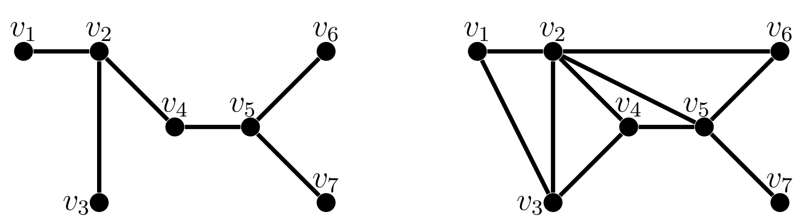

Figure 1.1: Degeneracy orderings for a 1-degenerate graph and a 2-degenerate graph.

Note that k-degenerate graphs do not have bounded maximum degree although $|E| \leq k|V|$. We will also think of the graphs as connected, though this is not part of the definition.

1.3 Overview

In Chapter 2, we demonstrate dimensional thresholds to guarantee an isometric copy of k-degenerate distance graph $\Gamma \subset \mathbb{F}_q^d$ in a subset $A \subset \mathbb{F}_q^d$ provided Γ satisfies some mild conditions[5] and provided the density of A satisfies a relatively strong condition. We also give explicit lower bounds for the number of isometric copies of Γ one is guaranteed to find in A. The strong condition on the density of A is beaten handily by the results in [IP19] when Γ has bounded degree,[6] however, any power gain is good enough to provide an avenue for an argument in $\mathbb{Z}^d$ which is our goal setting.

Chapter 3 is dedicated to establishing upper bounds for the number of lattice points which may be in the intersection of two spheres in several contexts. In Chapter 4, we obtain results analogous to Chapter 2's for a class of 2-degenerate graphs in $\mathbb{Z}^d$. This is done in essentially three parts. First an application of the classic methods of Birch in conjunction with the newer W-trick to write the number of integral solutions to a system of forms using the a so-called singular series and singular integral. The second part is a splicing together of various methods from the finite field setting along with the upper bounds from Chapter 3 to attain lower bounds on the number of isometric copies of Γ in a finite setting. The final part leverages the strong implications of positive upper Banach density to show that the result holds in much more general settings.

[5] As mentioned in that chapter, the strongest of those conditions is that Γ has non-zero edge lengths. This is not a condition which is actually necessary, but insisting on it makes the argument simpler.

[6] Iosevich and Parshall use an edge-deletion method while we use a vertex-deletion method.

1.4 Finite Fields as a Test Setting

The model case of finite fields is a useful one in which to test arguments when one's goal is to find distance graphs in $\mathbb{Z}^d$. It provides the strong advantage of not having sum convergence issues and the exponential sum estimates we use are well-known and tight. Given a distance graph Γ, we search copies of Γ in $A \subset \mathbb{F}_q^d$ which are isometric in the following sense:

Definition 1.1. $\Gamma' = (V', E')$ *is an isometric copy of distance graph* $\Gamma = (V, E)$ *in* $\mathbb{F}_q^d$ *if there is a bijection on vertices* $\varphi : V \to V'$ *and an induced bijection on edges* $E \to E'$ *in which all distances are preserved. That is, if there is an edge between* v_1 *and* v_2 *with* $(v_2 - v_1) \cdot (v_2 - v_1) = \lambda \in \mathbb{F}_q$*, then*

$$(\varphi(v_2) - \varphi(v_1)) \cdot (\varphi(v_2) - \varphi(v_1)) = \lambda.$$

Underpinning all of our arguments in $\mathbb{F}_q^d$ are estimates on the Fourier transform of the indicator function of a sphere restricted to certain affine subspaces. We here record some basic properties of the Fourier transform on $\mathbb{F}_q^d$. The

Fourier transform for a function $f : \mathbb{F}_q^d \to \mathbb{C}$ is given by

$$\widehat{f}(r) = \mathbb{E}_x f(x)\chi(x \cdot r),$$

where $\chi : \mathbb{F}_q \to \mathbb{C}$ is the canonical additive character and $\mathbb{E}_x$ denotes the expectation $q^{-d}\sum_{x\in\mathbb{F}_q^d}$. With this definition, the standard Fourier inversion, Plancherel, and convolution formulas are as follows

$$f(x) = \sum_{r\in\mathbb{F}_q^d} \hat{f}(r)\chi(-x \cdot r), \tag{1.1}$$

$$\mathbb{E}_x |f(x)|^2 = \sum_{r\in\mathbb{F}_q^d} |\hat{f}(r)|^2, \tag{1.2}$$

and

$$\widehat{f * g}(r) = \hat{f}(r)\hat{g}(r). \tag{1.3}$$

1.5 The Circle Method

In Chapter 3, we follow the machinery of the classical Hardy-Littlewood circle method to get upper bounds on the number of points on certain subsets of spheres in $\mathbb{Z}^d$. The starting point is the orthogonal identity that for $n \in \mathbb{Z}$,

$$\int_0^1 e(\alpha n)\, d\alpha = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $e(x) = e^{2\pi i x}$. This identity can be used to detect integral solutions to an equation or system of equations. In particular, by summing over $\underline{x} \in \mathbb{Z}^d$ we can count the number of solutions. The idea of the circle method is to switch the order of summation and integration, and then to carefully estimate the resulting inner summation depending on how close α is to a rational with small denominator. In our case, the set of α which lie away from rationals with small denominator, the so-called minor arcs, can be handled with relative ease by estimates dating to Weyl. The sum for the set of α lying near rationals with small denominator, the so-called major arcs, require a much more careful analysis.

We record for later use the following fact from Fourier analysis: If

$$\widehat{f}(\underline{\xi}) := \sum_{\underline{m}\in\mathbb{Z}^d} f(\underline{m})e(-\underline{m} \cdot \underline{\xi}), \tag{1.4}$$

and

$$\widetilde{f}(\underline{\xi}) := \int_{\mathbb{R}^d} f(\underline{x}) e(-\underline{x} \cdot \underline{\xi})\, d\underline{x}, \tag{1.5}$$

then Poisson summation yields

$$\widehat{f}(\underline{\xi}) = \sum_{\underline{m} \in \mathbb{Z}^d} \widetilde{f}(\underline{\xi} + \underline{m}) \tag{1.6}$$

for functions $f : \mathbb{R}^d \to [-1, 1]$ of compact support.

Remark 1.2. In chapters 3 and 4 if a variable is written with an underline, it is a vector, while quantities without an underline are one-dimensional.[7]

[7] We do not do this in chapter 2 because a key object is a vector of vectors: $X = (x_1, \ldots, x_k)$, with each $x_i \in \mathbb{F}_q^d$.

1.6 The W-Trick

A majority of the thesis works with respect to a parameter $W = 2 \prod_{p<w} p$.[8] The idea is that we can control potential bad behaviour of small primes in the singular series by only searching for solutions in a fixed residue class modulo W. This comes with the penalty of much worse $o(1)$ decay rates, but as we are only interested in main terms, it is a price we are more than happy to pay. This trick was crucial in the proof of Ben Green and Terence Tao that there are arbitrarily long arithmetic progression among the primes [GT08].

[8] The extra 2 is a technical necessity arising from the fact that we deal with only quadratic forms, which leads to unique behaviour with respect to $p = 2$.

1.7 Positive Upper Banach Density

In the setting of a set $A \subset \mathbb{Z}^d$, our starting assumption is that A has positive upper Banach density. Just this condition will have some remarkable structural implications for A which we work out below, and make use of in the final chapter of this thesis.

Definition 1.3. *The upper Banach density of a set $A \subset \mathbb{Z}^d$ is given by*

$$\delta^*(A) := \lim_{N\to\infty} \max_{\underline{t}\in\mathbb{Z}^d} \frac{|A \cap (\underline{t} + \mathcal{C}_N)|}{|\mathcal{C}_N|},$$

where $\mathcal{C}_N$ is the d-dimensional lattice cube $[1, N]^d$.[9]

[9] We use the notation $\mathcal{C}_N$ for a cube here instead of the Q_N notation used elsewhere because we set $Q_N = [-N, N]^d$ to be centered about 0. Although there is no conceptual issue with using Q_N instead, its side-length of $2N + 1$ causes unnecessary computational annoyance.

Lemma 1.4. *For any set $A \subset \mathbb{Z}^d$, the upper Banach density is well-defined.*

Proof. Set $a_N := \max_{\underline{t}\in\mathbb{Z}^d} \frac{|A\cap(\underline{t}+\mathcal{C}_N)|}{|\mathcal{C}_N|}$. We need to show $\lim_{N\to\infty} a_N$ exists.

To do this, we observe that the sequence $\{a_N\}_{N=1}^{\infty}$ satisfies the following two properties:

1. $a_{kN} \leq a_N$ for any positive integer k,
2. $a_{N+1} \leq a_N + \frac{d}{N}$.

The first property can be seen by writing $\mathcal{C}_{kN}$ as the union of k^d disjoint subcubes and noting that by the pigeon-hole principle, there is some subcube whose density is at least that of the density of A on the big cube.

The second property can be seen by calculating

$$\max_{\underline{t}\in\mathbb{Z}_d} |A \cap (\underline{t} + \mathcal{C}_{N+1})| \leq \max_{\underline{t}\in\mathbb{Z}_d} |A \cap (\underline{t} + \mathcal{C}_N)| + |\mathcal{C}_{N+1}| - |\mathcal{C}_N|.$$

Dividing through by $|\mathcal{C}_{N+1}|$ gives

$$\begin{aligned} a_{N+1} &\leq a_N \frac{|\mathcal{C}_N|}{|\mathcal{C}_{N+1}|} + \frac{|\mathcal{C}_{N+1}| - |\mathcal{C}_N|}{|\mathcal{C}_{N+1}|} \\ &\leq a_N + \frac{d(N+1)^{d-1}}{(N+1)^d} \\ &\leq a_N + \frac{d}{N}, \end{aligned}$$

where we estimated $|\mathcal{C}_{N+1}| - |\mathcal{C}_N|$ by recognizing it geometrically as a union of d cubes of dimension $(d-1)$.[10]

[10] Alternatively of course one could use the binomial theorem.

Let $1 \leq N < M$ and assume $kN < M \leq (k+1)N$. Then applying the above properties, we have

$$\begin{aligned} a_M &\leq a_{kN} + \frac{d}{kN} + \cdots + \frac{d}{M-1} \\ &\leq a_N + \frac{d}{kN}(M - kN) \\ &\leq a_N + \frac{d}{k}. \end{aligned}$$

Noting $a_N + d/k \to a_N$ as $M \to \infty$ we have $\limsup_{M\to\infty} a_M \leq a_N$ for every $N \in \mathbb{N}$. Hence

$$\limsup_{M\to\infty} a_M \leq \inf_{M\in\mathbb{Z}} a_M \leq \liminf_{M\to\infty} a_M,$$

so $\lim_{N\to\infty} a_N$ exists. □

Definition 1.5. *For $A \subset \mathbb{Z}^d$ and $m \in \mathbb{N}$ we define the relative upper Banach density with respect to spacing m as*

$$\delta_m^*(A) := \lim_{N\to\infty} \max_{\underline{t}\in\mathbb{Z}^d} \frac{|A \cap (\underline{t} + m\mathcal{C}_N)|}{|\mathcal{C}_N|}$$

Noting $|m\mathcal{C}_N| = |\mathcal{C}_N|$ and using $m\mathcal{C}_N$ instead of $\mathcal{C}_N$ in Lemma 1.4 shows that $\delta^*_m(A)$ is also well-defined for any $A \subset \mathbb{Z}$.[11]

[11] It is important to note the distinction between the solid lattice cube $\mathcal{C}_{Nm} = [1, \ldots, Nm]^d$ and the m-spaced lattice cube $m\mathcal{C}_N = [m, 2m, \ldots, Nm]^d$.

Definition 1.6. *A set $A \subset \mathbb{Z}^d$ is η-uniformly distributed with respect to spacing m if*

$$\delta^*_m(A) < \delta^*(A) + \eta,$$

and A is η-uniformly distributed up to threshold T if A is η-uniformly distributed with respect to spacing m for all $1 \le m \le T$.

Definition 1.7. *For $A \subset \mathbb{Z}^d$, $q \in \mathbb{N}$, and $\underline{s} \in \mathcal{C}_q$ define the scaled restricted set*

$$R_{q,\underline{s}}(A) := \left\{ \frac{\underline{x} - \underline{s}}{q} \ : \ \underline{x} \in A \cap (q\mathbb{Z}^d + \underline{s}) \right\}.$$

Lemma 1.8. *Fix $A \subset \mathbb{Z}^d$, $\eta > 0$, and $T \in \mathbb{N}$. There exist $q \le T^{1/\eta}$ and $\underline{s} \in \mathcal{C}_q$ such that*

1. *$\delta^*(R_{q,\underline{s}}(A)) \ge \delta^*(A)$, and*
2. *$R_{q,\underline{s}}(A)$ is η-uniformly distributed up to threshold T.*

Proof. If A is η-uniformly distributed up to threshold T, then the statement holds with $q = 1$. If not, then there exists some $1 \le m \le T$ such that $\delta^*_m(A) \ge \delta^*(A) + \eta$. Set

$$a_N := \max_{\underline{t} \in \mathbb{Z}^d} \frac{|A \cap (\underline{t} + m\mathcal{C}_N)|}{|\mathcal{C}_N|},$$

and for each N choose a representative $\underline{t}_N$ such that

$$a_N = \frac{|A \cap (\underline{t}_N + m\mathcal{C}_N)|}{|\mathcal{C}_N|}.$$

By the pigeonhole principle, an infinite subsequence $\{t_{N_i}\}_{i=1}^{\infty}$ falls into some single equivalence class modulo m with representative say $\underline{s} \in \mathcal{C}_m$. Therefore, we have that

$$\begin{aligned}
\delta^*(R_{q,\underline{s}}(A)) &= \lim_{N\to\infty} \max_{\underline{t} \in \mathbb{Z}^d} \frac{|R_{q,\underline{s}}(A) \cap (\underline{t} + \mathcal{C}_N)|}{|\mathcal{C}_N|} \\
&= \lim_{i\to\infty} \max_{\underline{t} \in \mathbb{Z}^d} \frac{|R_{q,\underline{s}}(A) \cap (\underline{t} + \mathcal{C}_{N_i})|}{|\mathcal{C}_{N_i}|} \\
&= \lim_{i\to\infty} \max_{\underline{t} \in m\mathbb{Z}^d + \underline{s}} \frac{|A \cap (\underline{t} + m\mathcal{C}_{N_i})|}{|\mathcal{C}_{N_i}|} \\
&= \lim_{i\to\infty} a_{N_i}.
\end{aligned}$$

As $\lim_{i\to\infty} a_{N_i} = \lim_{N\to\infty} a_N \geq \delta^*(A) + \eta$, we have $\delta^*(R_{q,\underline{s}}(A)) \geq \delta^*(A) + \eta$. If $R_{q,\underline{s}}(A)$ is η-uniformly distributed up to threshold T, then we are done, letting $q = m$.

If not, set $q_1 = m$, $\underline{s}_1 = \underline{s}$ and $A_1 = R_{q_1,\underline{s}_1}(A)$. This procedure can then be repeated to obtain an

$$A_2 = R_{q_2,\underline{s}_2}(A_1) = R_{q_1 q_2, q_1 \underline{s}_2 + \underline{s}_1}(A),$$

such that

$$\delta^*(A_2) \geq \delta^*(A_1) + \eta \geq \delta^*(A) + 2\eta.$$

In general, if A_n is not η-uniformly distributed up to threshold T, we can use the procedure above to obtain $A_{n+1} = R_{q_{n+1},\underline{s}_{n+1}}(A_n)$ where

$$\delta^*(A_{n+1}) \geq \delta^*(A) + (n+1)\eta.$$

As the density of a set cannot increase past 1, this process must terminate within $1/\eta$ iterations.

Therefore there is some A_n with $n \leq 1/\eta$ that is η-uniformly distributed up to threshold T. To complete the proof, we set

$$q = \prod_{i=1}^{n} q_i, \qquad \text{and} \qquad \underline{s} = \sum_{i=1}^{n} \left(\prod_{k=1}^{i-1} q_k \right) \underline{s}_i,$$

and note $A_n = R_{q,\underline{s}}(A)$ and that $q \leq T^{1/\eta}$.[12] □

[12] Following this proof, it is possible to generate a q such that $\delta^*(R_{q,s}(A)) \geq \delta^*(A)$ and $R_{q,s}(A)$ is η-uniformly distributed with respect to spacing m for all m. The downside is that there is no control on the size of q if the threshold T is not included.

If a set is η-uniformly distributed, we can now show that it contains subsets which satisfy a compact version of being η-uniformly distributed.[13]

[13] The arcane definition of (η, L)-uniformly distributed is chosen to exactly match $\|f_A\|_{U^1_{m,L}(\mathcal{C}_N)}$ which is a key object in Chapter 4.

Definition 1.9. *Fix $0 < \eta \leq 1$ and let $m, L, N \in \mathbb{N}$. A set $A \subset \mathcal{C}_N$ with density $\alpha := \frac{|A|}{|\mathcal{C}_N|}$ is (η, L)-uniformly distributed with respect to spacing m if*

$$\frac{1}{|\mathcal{C}_N|} \sum_{\underline{t} \in \mathbb{Z}^d} \left| \frac{|A \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} - \alpha \frac{|\mathcal{C}_N \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} \right|^2 < \eta^2.$$

Lemma 1.10. *Fix $0 < \eta \leq 1$ and $m \in \mathbb{N}$, let $A \subset \mathbb{Z}^d$ be $\frac{1}{16}\eta^4$-uniformly distributed with respect to spacing m and fix L such that*

$$\max_{\underline{t} \in \mathbb{Z}^d} \frac{|A \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} < \delta^*(A) + \frac{1}{16}\eta^4.$$

There exists an $N_0 = N_0(A, L, \eta, m, d)$ such that for any $N \geq N_0$, there is a $\underline{t}_N$ such that the set $(A - \underline{t}_N) \cap \mathcal{C}_N$ has density $\gamma := \frac{|(A-\underline{t}_N)\cap\mathcal{C}_N|}{|\mathcal{C}_N|} \geq \delta^*(A)$ and is (η, L)-uniformly distributed with respect to spacing m.

Proof. For convenience, set $\varepsilon := \frac{1}{4}\eta^2$. Choose N_0 such that for any $N \geq N_0$

$$|\mathcal{C}_{N+mL}| - |\mathcal{C}_{N-mL}| < \varepsilon^4 |\mathcal{C}_N|,$$

and

$$\max_{\underline{t}\in\mathbb{Z}^d} \frac{|A \cap (\underline{t} + \mathcal{C}_N)\,|}{|\mathcal{C}_N|} - \delta^*(A) < \varepsilon^2.$$

Fix $N \geq N_0$. Let $\underline{t}_N$ be a representative of the maximum density, i.e. choose $\underline{t}_N \in \mathbb{Z}^d$ such that

$$\frac{|A \cap (\underline{t}_N + \mathcal{C}_N)\,|}{|\mathcal{C}_N|} = \max_{\underline{t}\in\mathbb{Z}^d} \frac{|A \cap (\underline{t} + \mathcal{C}_N)\,|}{|\mathcal{C}_N|},$$

and set

$$A' := (A - \underline{t}_N) \cap \mathcal{C}_N.$$

As A' corresponds to the maximum density of A on grids of size Q_N and some grid of size Q_N must have density $\geq \delta^*(A)$, we have $\gamma \geq \delta^*(A)$. We will show A' is also (η, L)-uniformly distributed with respect to spacing m. To this end, let

$$T := \sum_{\underline{t}\in\mathbb{Z}^d} \left| \frac{|A' \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} - \gamma \frac{|\mathcal{C}_N \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} \right|^2$$

Observe that the summand

$$\frac{|A' \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} - \gamma \frac{|\mathcal{C}_N \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|}$$

is supported only for $\underline{t} \in \mathcal{C}_{N+mL}$, and additionally, if $t \in D$ where

$$D := \mathcal{C}_N \cap (\mathcal{C}_N + (mL, \ldots, mL))\,,$$

then $|\mathcal{C}_N \cap (\underline{t} - m\mathcal{C}_L)| = |\mathcal{C}_L|$. Noting that $|D| = |\mathcal{C}_{N-mL}|$, we have by the definition of N_0

$$T \leq \sum_{\underline{t}\in D} \left| \frac{|A' \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} - \gamma \right|^2 + \varepsilon^4 |\mathcal{C}_N|$$

By the definition of L and the observation $\gamma \geq \delta^*(A)$, for every $\underline{t} \in \mathbb{Z}^d$ we have

$$\frac{|A' \cap (\underline{t} - m\mathcal{C}_L)\,|}{|\mathcal{C}_L|} < \gamma + \varepsilon^2.$$

We set

$$B := \left\{ \underline{t} \in D \ : \ \frac{|A' \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} \leq \gamma - \varepsilon \right\},$$

so

$$T \leq |B| + \varepsilon^2|D| + \varepsilon^4|\mathcal{C}_N| \leq |B| + 2\varepsilon^2|\mathcal{C}_N|. \tag{1.7}$$

Observe

$$\sum_{\underline{t} \in \mathbb{Z}^d} |A' \cap (\underline{t} - m\mathcal{C}_L)| = |\mathcal{C}_L||A'|$$

because every $x \in A'$ is detected exactly $|m\mathcal{C}_L| = |\mathcal{C}_L|$ times. Rearranging and writing $|A'| = \gamma|\mathcal{C}_N|$, we have

$$\gamma|\mathcal{C}_N| = \sum_{\underline{t} \in \mathbb{Z}^d} \frac{|A' \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|}.$$

Noting the support of the summand and the definition of N_0, we have

$$(\gamma - \varepsilon^4)|\mathcal{C}_N| \leq \sum_{\underline{t} \in \mathcal{C}_N} \frac{|A' \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|}.$$

By the definition of L, each term is at most $\gamma + \varepsilon^2$, so we have

$$(\gamma - \varepsilon^4)|\mathcal{C}_N| \leq |B|(\gamma - \varepsilon) + (|\mathcal{C}_N| - |B|)(\gamma + \varepsilon^2).$$

Rearranging, we obtain

$$|B|(\varepsilon + \varepsilon^2) \leq |\mathcal{C}_N|(\varepsilon^2 + \varepsilon^4),$$

which gives $|B| \leq 2\varepsilon|\mathcal{C}_N|$. Plugging this back into equation (1.7) yields

$$T \leq 4\varepsilon|\mathcal{C}_n|. \tag{1.8}$$

Recalling $\varepsilon = \frac{1}{4}\eta^2$, equation (1.8) says exactly that A' is (η, L)-uniformly distributed with respect to m. □

We now define the $U^1_{m,L}(Q)$ norm and tie it to the notion of a set being (η, L)-uniformly distributed with respect to m.

Definition 1.11. *For a lattice cube $Q \subset \mathbb{Z}^d$ and a function $f : Q \to [-1, 1]$*

$$\|f\|_{U^1_{m,L}(Q)} := \left(\frac{1}{|Q|} \sum_{\underline{t} \in \mathbb{Z}^d} |f * \chi_{m\mathcal{C}_L}(\underline{t})|^2 \right)^{1/2},$$

where $\chi_{m\mathcal{C}_L}$ is the normalized characteristic function of $m\mathcal{C}_L$.

Lemma 1.12. *Let $A \subset Q$ with density $\alpha := \frac{|A|}{|Q|}$. Then A is (η, L)-uniformly distributed with respect to m iff $\|f_A\|_{U^1_{m,L}(Q)} < \eta$ where $f_A := 1_A - \alpha 1_Q$.*

Proof.

$$\begin{aligned} f_A * \chi_{mC_L}(\underline{t}) &= \frac{1}{|\mathcal{C}_L|} \sum_{\underline{x} \in \mathbb{Z}^d} f_A(\underline{t} - \underline{x}) 1_{m\mathcal{C}_L}(\underline{x}) \\ &= \frac{1}{|\mathcal{C}_L|} \sum_{\underline{x} \in \mathbb{Z}^d} [1_A(\underline{t} - \underline{x}) 1_{m\mathcal{C}_L}(\underline{x}) - \alpha 1_Q(\underline{t} - \underline{x}) 1_{m\mathcal{C}_L}(\underline{x})] \\ &= \frac{|A \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} - \alpha \frac{|Q \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|}. \end{aligned}$$

Therefore

$$\begin{aligned} \|f_A\|^2_{U^1_{m,L}(Q)} &= \frac{1}{|Q|} \sum_{\underline{t} \in \mathbb{Z}^d} |f * \chi_{m\mathcal{C}_L}(\underline{t})|^2 \\ &= \frac{1}{|Q|} \sum_{\underline{t} \in \mathbb{Z}^d} \left| \frac{|A \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} - \alpha \frac{|Q \cap (\underline{t} - m\mathcal{C}_L)|}{|\mathcal{C}_L|} \right|^2. \end{aligned}$$

□

Finally, we address a technical lemma showing that (η, L)-uniformity of A implies a type of uniformity on the q-scaled $\underline{s}$-restricted set $R_{q,\underline{s}}(A)$.

Lemma 1.13. *Fix $q \in \mathbb{N}$ and let $\underline{s} \in \mathcal{C}_q$. Let $A \subset \mathcal{C}_{qN}$ with density $\alpha := \frac{|A|}{|\mathcal{C}_{qN}|}$. Define $A_{\underline{s}} := R_{q,\underline{s}}(A)$ and $g_{\underline{s}} := 1_{A_{\underline{s}}} - \alpha 1_{\mathcal{C}_N}$. If A is $(q^{-d/2}\eta, L)$-uniformly distributed with respect to qm, then $\|g_{\underline{s}}\|_{U^1_{m,L}(\mathcal{C}_N)} < \eta$ for every $\underline{s} \in \mathcal{C}_q$.*

Proof. Writing $\underline{t} = q\underline{t}_1 + \underline{s}$, and noting $|(A - \underline{t}) \cap qm\mathcal{C}_L| = |(A_{\underline{s}} - \underline{t}_1) \cap m\mathcal{C}_L|$, we have

$$\frac{\eta^2}{q^d} > \frac{1}{|\mathcal{C}_{qN}|} \sum_{\underline{t} \in \mathbb{Z}^d} \left| \frac{|A \cap (\underline{t} - qm\mathcal{C}_L)|}{|\mathcal{C}_L|} - \alpha \frac{|\mathcal{C}_{qN} \cap (\underline{t} - qm\mathcal{C}_L)|}{|\mathcal{C}_L|} \right|^2$$

$$= \frac{1}{|\mathcal{C}_{qN}|} \sum_{\underline{t} \in \mathbb{Z}^d} \left| \frac{|(A - \underline{t}) \cap qm\mathcal{C}_L|}{|\mathcal{C}_L|} - \alpha \frac{|(\mathcal{C}_{qN} - \underline{t}) \cap qm\mathcal{C}_L|}{|\mathcal{C}_L|} \right|^2$$

$$= \frac{1}{|\mathcal{C}_{qN}|} \sum_{\underline{s} \in \mathcal{C}_q} \sum_{\underline{t}_1 \in \mathbb{Z}^d} \left| \frac{|(A_{\underline{s}} - \underline{t}_1) \cap m\mathcal{C}_L|}{|\mathcal{C}_L|} - \alpha \frac{|(\mathcal{C}_N - \underline{t}_1) \cap m\mathcal{C}_L|}{|\mathcal{C}_L|} \right|^2$$

$$= \frac{1}{q^d} \sum_{\underline{s} \in \mathcal{C}_q} \frac{1}{|\mathcal{C}_N|} \sum_{\underline{t}_1 \in \mathbb{Z}^d} \left| \frac{|(A_{\underline{s}} - \underline{t}_1) \cap m\mathcal{C}_L|}{|\mathcal{C}_L|} - \alpha \frac{|(\mathcal{C}_N - \underline{t}_1) \cap m\mathcal{C}_L|}{|\mathcal{C}_L|} \right|^2$$

$$= \frac{1}{q^d} \sum_{\underline{s} \in \mathcal{C}_q} \|g_{\underline{s}}\|^2_{U^1_{m,L}(\mathcal{C}_N)},$$

where the last line follows from the proof of Lemma 1.12.[14] Multiplying through by q^d shows $\|g_{\underline{s}}\|_{U^1_{m,L}(\mathcal{C}_N)} < \eta$ for any $\underline{s} \in \mathcal{C}_q$.

□

[14] Note that while $g_{\underline{s}}$ is not quite the balance function of $A_{\underline{s}}$, the computation in Lemma 1.12 still goes through because the only difference is the constant α.

Chapter 2

k-Degenerate Distance Graphs in Finite Fields

In this chapter we will describe two approaches to counting k-degenerate distance graphs in subsets of finite fields. The first follows a Von-Neumann type counting argument which gives nontrivial results with the requirement that the dimension d is at least linear in the graph degeneracy k, while the second is a simplified inductive approach valid for d quadratic in k, but which has the advantage of being adaptable to the integer lattice.[15]

[15] Though some elements unique to the first approach are also drawn upon for the argument in $\mathbb{Z}^d$.

The model case of finite fields is a useful one to test the viability of arguments before translating them to the technicality-ridden world of $\mathbb{Z}^d$. It does however come with its own pitfalls. One culprit in our case is that the analogue of distance we use is not a norm. For $y \in \mathbb{F}_q^d$, we use the definition $|y|^2 := y \cdot y$, but $|y|^2 = 0$ does not imply that $y = 0$, i.e. this quadratic form is isotropic. It is possible to work around this issue with the penalty of more notation and worse estimates, but as this artifact (mostly) does not appear in our goal setting of $\mathbb{Z}^d$, we will avoid it by assuming our distance graphs have no edges of length 0.

Both approaches show that under suitable assumptions on d, q, and α that every subset of $\mathbb{F}_q^d$ with density α contains an isometric copy of every k-degenerate distance graph with nonzero edge lengths.

2.1 Preliminaries

For completeness, we record the character sum estimates used to count points in the intersection of spheres in $\mathbb{F}_q^d$.

2.1.1 Spheres in Finite Fields

We start with the definitions

Definition 2.1. *If $t \in \mathbb{F}_q$ and $x \in \mathbb{F}_q^d$, then the sphere $S_t(x)$ is defined as*

$$S_t(x) := \{y \in \mathbb{F}_q^d \ : \ |y-x|^2 = t\}$$

Definition 2.2. *If $\underline{t} = (t_1, \ldots, t_k)$ and $X = (x_1, \ldots, x_k)$ are tuples with $t_i \in \mathbb{F}_q$ and $x_i \in \mathbb{F}_q^d$ for $i = 1, \ldots, k$, then the sphere $S_{\underline{t}}(X)$ is defined as*

$$S_{\underline{t}}(X) := \{y \in \mathbb{F}_q^d \ : \ |y-x_i|^2 = t_i \textit{ for } i = 1, \ldots, k\}$$

Note that if $\underline{t} = (t_1, \ldots, t_k)$ and $X = (x_1, \ldots, x_k)$, then

$$S_{\underline{t}}(X) = \bigcap_{i=1}^{k} S_{t_i}(x_i),$$

so $S_{\underline{t}}(X)$ can be thought of as the intersection of k spheres.

It will additionally be convenient to define normalized indicator functions associated to these spheres.

Definition 2.3. *If $t \in \mathbb{F}_q$ then the normalized indicator function of the sphere $S_t(x)$ is given by*

$$\sigma_t^x(y) = \begin{cases} q & \textit{if } |y-x|^2 = t, \\ 0 & \textit{otherwise.} \end{cases}$$

Definition 2.4. *If $\underline{t} = (t_1, \ldots, t_k)$ and $X = (x_1, \ldots, x_k)$ with $t_i \in \mathbb{F}_q$ and $x_i \in \mathbb{F}_q^d$ for $i = 1, \ldots, k$ then the normalized indicator function of the sphere $S_{\underline{t}}(X)$ is given by*

$$\sigma_{\underline{t}}^X(y) = \begin{cases} q^k & \textit{if } |y-x_1|^2 = t_1, \ldots, |y-x_k|^2 = t_k, \\ 0 & \textit{otherwise.} \end{cases}$$

Note that

$$|S_t(x)| = q^{-1} \sum_{y \in \mathbb{F}_q} \sigma_t^x(y),$$

and

$$|S_{\underline{t}}(X)| = q^{-k} \sum_{y \in \mathbb{F}_q} \sigma_{\underline{t}}^X(y).$$

There are however convenient ways to rewrite $\sigma_t^x(y)$ and $\sigma_{\underline{t}}^X(y)$. The way we will rewrite $\sigma_t^x(y)$ is a bit trivial, but the generalization to rewriting $\sigma_{\underline{t}}^X(y)$ will be quite computationally useful.

Lemma 2.5. *For $t \in \mathbb{F}_q$ and $x \in \mathbb{F}_q^d$ we have*

$$\sigma_t^x(y) = \sigma_t^0(x - y).$$

Proof. This is immediate as

$$\begin{aligned}\sigma_t^0(x-y) &= \begin{cases} q & \text{if } |x-y-0|^2 = t, \\ 0 & \text{otherwise.} \end{cases} \\ &= \sigma_t^x(y)\end{aligned}$$

□

Because its use is ubiquitous, we specially rename $\sigma_t^0(y)$.

Definition 2.6. *If $t \in \mathbb{F}_q$, the normalized indicator of the sphere centered at 0 is*

$$\sigma_t(y) = \begin{cases} q & \text{if } |y|^2 = t, \\ 0 & \text{otherwise.} \end{cases}.$$

Now we turn to the more interesting rewriting of $\sigma_{\underline{t}}^X(y)$. We can think of $\sigma_{\underline{t}}^X(y)$ not as an intersection of k spheres, but rather a single sphere restricted to an affine subspace.

Lemma 2.7. *If $\underline{t} = (t_1, \ldots, t_k)$ and $X = (x_1, \ldots, x_k)$ with $t_i \in \mathbb{F}_q$ and $x_i \in \mathbb{F}_q^d$ for $i = 1, \ldots, k$ and $\operatorname{char}(\mathbb{F}_q) \neq 2$, then*

$$\sigma_{\underline{t}}^X(y) = \sigma_{\underline{t}}^Z(x_1 - y),$$

where $Z = (0, x_1 - x_2, \ldots, x_1 - x_k)$. Additionally,

$$\sigma_{\underline{t}}^Z(y) = \begin{cases} q^k & \text{if } |y|^2 = t_1 \text{ and } y \cdot z_i = c_i \text{ for } i = 2, \ldots, k, \\ 0 & \text{otherwise,} \end{cases}$$

where $c_i = (|z_i|^2 - t_i + t_1)/2$ for $i = 2, \ldots, k$.

Proof. The first equality is immediate as

$$\sigma_{\underline{t}}^Z(x_1 - y) = \begin{cases} q^k & \text{if } |x_1 - y - z_1|^2 = t_1, \ldots, |x_1 - y - z_k|^2 = t_k, \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} q^k & \text{if } |x_1 - y|^2 = t_1, \ldots, |x_1 - y - x_1 + x_k|^2 = t_k, \\ 0 & \text{otherwise.} \end{cases}$$
$$= \sigma_{\underline{t}}^X(y).$$

The second equality arises by subtracting the i-th equation $|y - z_i|^2 = t_i$ from the first equation $|y|^2 = t_1$ for each $i \geq 2$:

$$|y|^2 - |y|^2 + 2z_i \cdot y - |z_i|^2 = t_1 - t_i \iff y \cdot z_i = \frac{|z_i|^2 - t_i + t_1}{2},$$

given $\mathrm{char}(\mathbb{F}_q) \neq 2$. □

2.1.2 Fourier Transform of Spheres in a Finite Field

Estimates on $|\hat{\sigma}_t^x(r)|$ and $|\hat{\sigma}_{\underline{t}}^X(r)|$ will be crucial, and for these we follow the presentation in the thesis work of Parshall [Par17], though we will make some minor changes to match the above notation and track constants.[16]

[16] The constants are relevant for the inductive arguments used later.

Let χ be the canonical additive character and let η be the multiplicative character or order 2, i.e. $\eta(s) = 1$ if s is a square and $\eta(s) = -1$ if s is a nonsquare for $s \in \mathbb{F}_q^*$. Additionally let $G(\eta, \chi)$ be the Gaussian sum

$$G(\eta, \chi) = \sum_{s \in \mathbb{F}_q^*} \eta(s)\chi(s).$$

We have the following standard computations from the book of Lidl and Niederreiter [LN96] when $\mathrm{char}(\mathbb{F}_q) \neq 2$:

$$|G(\eta, \chi)| = \sqrt{q}, \tag{2.1}$$

and

$$\sum_{x \in \mathbb{F}_q} \chi(ax^2 + bx) = G(\eta, \chi)\eta(a)\chi(-b^2(4a)^{-1}), \tag{2.2}$$

for $a \in \mathbb{F}_q^*$ and $b \in \mathbb{F}_q$.

We will make use of the deep character sum estimates of Sailé [Sal32] and Weil [Wei48] relating to the Riemann hypothesis over finite fields that

$$\left| \sum_{s \in \mathbb{F}_q^*} \eta^d(s)\chi\left(as + bs^{-1}\right) \right| \leq 2\sqrt{q}, \tag{2.3}$$

where $a, b \in \mathbb{F}_q$ and at least one of a and b is nonzero.

We first carry through the estimate of $|\hat{\sigma}_t(r)|$ in the case where $t \neq 0$ as in [IP19] or [Par17].[17]

Lemma 2.8. *If* $t \in \mathbb{F}_q^*$ *and* $x \in \mathbb{F}_q^d$ *with* $\mathrm{char}(\mathbb{F}_q) \neq 2$, *then*

$$|\hat{\sigma}_t^x(r)| = 1_{\{0\}}(r) + \mathcal{E}(r),$$

where $|\mathcal{E}(r)| \leq 2q^{(1-d)/2}$.[18]

Proof. First from Lemma 2.5 we have

$$\sigma_t^x(y) = \sigma_t(x-y),$$

so

$$\begin{aligned} \hat{\sigma}_t^x(r) &= \mathbb{E}_y \sigma_t^x(y)\chi(y \cdot r) \\ &= \mathbb{E}_y \sigma_t(x-y)\chi(y \cdot r) \\ &= \mathbb{E}_u \sigma_t(u)\chi((x-u)\cdot r) \\ &= \chi(x \cdot r)\hat{\sigma}_t(-r), \end{aligned}$$

and therefore $|\hat{\sigma}_t^x(r)| = |\hat{\sigma}_t(-r)|$.

Via orthogonality we have

$$\sigma_t(y) = \sum_{s \in \mathbb{F}_q} \chi(s(|y|^2 - t)).$$

Now we compute

$$\begin{aligned} \hat{\sigma}_t(r) &= \mathbb{E}_y \sigma_t(y)\chi(r \cdot y) \\ &= \mathbb{E}_y \sum_{s \in \mathbb{F}_q} \chi(s(|y|^2 - t))\chi(r \cdot y) \\ &= \sum_{s \in \mathbb{F}_q} \chi(-st)\mathbb{E}_y \chi(s|y|^2 + r \cdot y) \\ &= 1_{\{0\}}(r) + \sum_{s \in \mathbb{F}_q^*} \chi(-st)\mathbb{E}_y \chi(s|y|^2 + r \cdot y) \end{aligned}$$

Setting

$$\mathcal{E}(r) = \sum_{s \in \mathbb{F}_q^*} \chi(-st)\mathbb{E}_y \chi(s|y|^2 + r \cdot y),$$

it remains to show $|\mathcal{E}(r)| < 2q^{(1-d)/2}$.

For this, we first observe that the inner expectation splits as a sum and then apply equation (2.2):

[17] It is also possible to work out these estimates with $t = 0$ and obtain slightly larger error bounds as Parshall does in [Par17].

[18] By $1_{\{0\}}(r)$ we mean the indicator function of the condition $r = 0$, i.e. $1_{\{0\}}(r) = \begin{cases} 1 & \text{if } r = 0, \\ 0 & \text{otherwise.} \end{cases}$

$$\begin{aligned}\mathbb{E}_y \chi(s|y|^2 + r\cdot y) &= q^{-d}\sum_{y\in\mathbb{F}_q^d}\chi(s|y|^2 + r\cdot y)\\ &= q^{-d}\prod_{i=1}^d\sum_{y_i\in\mathbb{F}_q}\chi(sy_i^2 + r_iy_i)\\ &= q^{-d}\prod_{i=1}^d G(\eta,\chi)\eta(s)\chi(-r_i^2(4s)^{-1})\\ &= q^{-d}G(\eta,\chi)^d\eta^d(s)\chi(-|r|^2(4s)^{-1}).\end{aligned}$$

Hence

$$\mathcal{E}(r) = q^{-d}G(\eta,\chi)^d\sum_{s\in\mathbb{F}_q^*}\eta^d(s)\chi(-st - |r|^2(4s)^{-1}). \tag{2.4}$$

Now as we assumed $t \neq 0$ we may apply equation (2.3) and finally equation (2.1) to obtain the desired

$$\begin{aligned}|\mathcal{E}(r)| &\leq 2q^{-d}|G(\eta,\chi)|^d\sqrt{q}\\ &\leq 2q^{(1-d)/2}.\end{aligned}$$

Noting $1_{\{0\}}(-r) = 1_{\{0\}}(r)$, we get

$$|\hat{\sigma}_t^x(r)| = |\hat{\sigma}_t(-r)| = 1_{\{0\}}(r) + \mathcal{E}(-r),$$

where we may as well redefine $\mathcal{E}(r)$ as $\mathcal{E}(-r)$ since the bound is uniform in r.

□

Corollary 2.9. *If* $t \in \mathbb{F}_q^*$ *and* $x \in \mathbb{F}_q^d$ *where* $\mathrm{char}(\mathbb{F}_q) \neq 2$, *then*

$$|S_t(x)| = q^{d-1} + \mathcal{E},$$

where $|\mathcal{E}| \leq 2q^{(d-1)/2}$.

Proof. We compute

$$|S_t(x)| = q^{-1}\sum_{y\in\mathbb{F}_q}\sigma_t^x(y) = q^{d-1}\hat{\sigma}_t^x(0),$$

where via Lemma 2.8

$$|\hat{\sigma}_t^x(0)| = 1 + \mathcal{E},$$

with $\mathcal{E} \leq 2q^{(1-d)/2}$. □

We now turn to the more general case of estimating $\hat{\sigma}_{\underline{t}}^X(r)$ where we assume $t_i \neq 0$.[19] First we need a definition to ensure that $S_{\underline{t}}(X)$ is nonempty.

[19] Again, it is possible to carry out estimates when some collection of the t_i are 0 as is done in [Par17], but we shall not need those as our goal case is in $\mathbb{Z}^d$ where $t_i = 0$ would correspond to a length zero edge.

Definition 2.10. *A tuple $X = (x_1, \ldots, x_k)$ is said to be in general position if setting $z_i = x_1 - x_i$ for $i = 2, \ldots, k$, the set $z_2, \ldots, z_k$ is linearly independent.*

Lemma 2.11. *If $\underline{t} = (t_1, \ldots, t_k)$ and $X = (x_1, \ldots, x_k)$ with $t_i \in \mathbb{F}_q^*$ and $x_i \in \mathbb{F}_q^d$ for $i = 1, \ldots, k$ where X is in general position and $\operatorname{char}(\mathbb{F}_q) \neq 2$, then*

$$|\hat{\sigma}_{\underline{t}}^X(r)| = 1_{\langle z_2,\ldots,z_k\rangle}(r) + \mathcal{E}(r),$$

where $z_i = x_1 - x_i$ for $i = 2, \ldots, k$ and $|\mathcal{E}(r)| \leq 2q^{(2k-1-d)/2}$.

Proof. The proof follows the same lines as the proof of Lemma 2.8. First from Lemma 2.7 we have

$$\sigma_{\underline{t}}^X(y) = \sigma_{\underline{t}}^Z(x_1 - y),$$

where $Z = (0, x_1 - x_2, \ldots, x_1 - x_k)$, so

$$\begin{aligned}
\hat{\sigma}_{\underline{t}}^X(r) &= \mathbb{E}_y \sigma_{\underline{t}}^X(y)\chi(y \cdot r) \\
&= \mathbb{E}_y \sigma_{\underline{t}}^Z(x_1 - y)\chi(y \cdot r) \\
&= \mathbb{E}_u \sigma_{\underline{t}}^Z(u)\chi((x_1 - u) \cdot r) \\
&= \chi(x_1 \cdot r)\hat{\sigma}_{\underline{t}}^Z(-r),
\end{aligned}$$

and therefore

$$|\hat{\sigma}_{\underline{t}}^X(r)| = |\hat{\sigma}_{\underline{t}}^Z(-r)|.$$

For notational convenience, fix $v_Z \in \mathbb{F}_q^d$ such that $v_Z \cdot z_i = c_i$ for $i = 2, \ldots, k$.[20] First note via orthogonality that

[20] Such a v_Z exists because $z_2, \ldots, z_k$ are linearly independent.

$$\begin{aligned}
\sigma_{\underline{t}}^Z(y) &= \sum_{s \in \mathbb{F}_q} \chi(s(|y|^2 - t_1)) \prod_{i=2}^{k} \sum_{s_i \in \mathbb{F}_q} \chi(s_i(y \cdot z_i - c_i)) \\
&= \sum_{s \in \mathbb{F}_q} \chi(s(|y|^2 - t_1)) \prod_{i=2}^{k} \sum_{s_i \in \mathbb{F}_q} \chi((y - v_Z) \cdot s_i z_i) \\
&= \sum_{s \in \mathbb{F}_q} \chi(s(|y|^2 - t_1)) \sum_{s_2,\ldots,s_k \in \mathbb{F}_q} \chi\left((y - v_Z) \cdot \sum_{i=2}^{k} s_i z_i\right) \\
&= \sum_{s \in \mathbb{F}_q} \chi(s(|y|^2 - t_1)) \sum_{v \in \langle z_2,\ldots,z_k\rangle} \chi\left((y - v_Z) \cdot v\right).
\end{aligned}$$

Now

$$
\begin{aligned}
\hat{\sigma}_{\underline{t}}^{Z}(r) &= \mathbb{E}_y \sigma_{\underline{t}}^{Z}(y)\chi(r\cdot y) \\
&= \mathbb{E}_y \sum_{s\in\mathbb{F}_q} \chi(s(|y|^2-t_1)) \sum_{v\in\langle z_2,\dots,z_k\rangle} \chi\left((y-v_Z)\cdot v\right)\chi(r\cdot y) \\
&= \sum_{s\in\mathbb{F}_q} \chi(-st_1) \sum_{v\in\langle z_2,\dots,z_k\rangle} \chi(-v_Z\cdot v)\mathbb{E}_y\chi(s|y|^2+y\cdot(v+r)).
\end{aligned}
$$

If $s=0$, the inner expectation is zero unless $v=-r$ which only happens when $-r\in\langle z_2,\dots,z_k\rangle$, so

$$
\hat{\sigma}_{\underline{t}}^{Z}(r) = \chi(v_Z\cdot r)1_{\langle z_2,\dots,z_k\rangle}(-r) + \mathcal{E}(r),
$$

where

$$
\mathcal{E}(r) = \sum_{s\in\mathbb{F}_q^*} \chi(-st_1) \sum_{v\in\langle z_2,\dots,z_k\rangle} \chi(-v_Z\cdot v)\mathbb{E}_y\chi(s|y|^2+y\cdot(v+r)).
$$

It remains to show $|\mathcal{E}(r)|\le 2q^{(2k-1-d)/2}$. Noting that the inner expectation splits and applying equation (2.2) again gives

$$
\begin{aligned}
\mathbb{E}_y\chi(s|y|^2+y\cdot(v+r)) &= q^{-d}\sum_{y\in\mathbb{F}_q^d}\chi(s|y|^2+y\cdot(v+r)) \\
&= q^{-d}\prod_{i=1}^{d}\sum_{y_i\in\mathbb{F}_q}\chi(sy_i^2+y_i(v_i+r_i)) \\
&= q^{-d}\prod_{i=1}^{d} G(\eta,\chi)\eta(s)\chi(-(v_i+r_i)^2(4s)^{-1}) \\
&= q^{-d}G(\eta,\chi)^d\eta^d(s)\chi(-|v+r|^2(4s)^{-1}).
\end{aligned}
$$

Using the Gaussian sum computation equation (2.1) gives

$$
|\mathcal{E}(r)| \le q^{-d/2}\sum_{v\in\langle z_2,\dots,z_k\rangle}\left|\sum_{s\in\mathbb{F}_q^*}\eta^d(s)\chi(-st_1-|v+r|^2(4s)^{-1})\right|. \quad (2.5)
$$

The inner sum is amenable to the estimate in equation (2.3) as t_1 was assumed to be nonzero. This yields the desired

$$
\begin{aligned}
|\mathcal{E}(r)| &\leq q^{-d/2} \sum_{v \in \langle z_2, \ldots, z_k \rangle} 2\sqrt{q} \\
&\leq 2q^{(2k-1-d)/2}.
\end{aligned}
$$

Once again redefining $\mathcal{E}(r)$ as $\mathcal{E}(-r)$ we get

$$|\hat{\sigma}_{\underline{t}}^X(r)| = |\hat{\sigma}_{\underline{t}}^Z(-r)| = 1_{\langle z_2, \ldots, z_k \rangle}(r) + \mathcal{E}(r).$$

□

Corollary 2.12. *If* $\underline{t} = (t_1, \ldots, t_k)$ *and* $X = (x_1, \ldots, x_k)$ *with* $t_i \in \mathbb{F}_q^*$ *and* $x_i \in \mathbb{F}_q^d$ *for* $i = 1, \ldots, k$ *where* X *is in general position and* $\mathrm{char}(\mathbb{F}_q) \neq 2$*, then*

$$|S_{\underline{t}}(X)| = q^{d-k} + \mathcal{E},$$

where $|\mathcal{E}| \leq 2q^{(d-1)/2}$.

Proof.

$$|S_{\underline{t}}(X)| = q^{-k} \sum_{y \in \mathbb{F}_q} \sigma_{\underline{t}}^X(y) = q^{d-k} \hat{\sigma}_{\underline{t}}^X(0).$$

Via Lemma 2.11 we have

$$|\hat{\sigma}_{\underline{t}}^X(0)| = 1 + \mathcal{E}_1,$$

where $|\mathcal{E}_1| \leq 2q^{(2k-1-d)/2}$. Plugging this in, we get

$$|S_{\underline{t}}(X)| \leq q^{d-k} + q^{d-k}\mathcal{E}_1,$$

with $q^{d-k}|\mathcal{E}_1| \leq 2q^{(d-1)/2}$. □

2.2 Direct Counting Function Approach

In this first approach, we will prove that a subset $A \subset \mathbb{F}_q^d$ contains any k-degenerate graph with n vertices under suitable assumptions on the dimension d and the density of A. The main point of this approach is that a nontrivial estimate exists when $d > 2k - 1$. We will only obtain a quadratic dependence of d on k in the second approach, though that approach is more suited to the case of the integer lattice.

The general strategy we will follow to estimate an averaged counting function Λ_Γ which detects appearances of copies of the distance graph Γ in a set $A \subset \mathbb{F}_q^d$. This is an adaptation of the method used by Lyall and Magyar in [LM20b] from setting of $\mathbb{Z}^d$ to the case of finite fields. It also contains a generalization to k-degenerate graphs, which is much easier in finite fields than $\mathbb{Z}^d$.

For a beginning example, let Γ be the distance graph with two vertices and one edge of length $t \neq 0$. The averaged counting function here will boil down to

$$\Lambda_\Gamma = \mathbb{E}_{x,y} 1_A(x) 1_A(x+y) \sigma_t(y),$$

where as before $\sigma_t(y)$ is the L^1-normalized indicator function of the sphere $S_t(0) = \{y \in \mathbb{F}_q^d \; : \; |y|^2 = t\}$. Showing Λ_Γ is positive is equivalent to showing A contains at least one isometric copy of Γ.

We now want to define the averaged counting function. We shall need notation to track the edge requirements of distance graph Γ. First we make a notation for the lengths of the backwards edges associated to vertex v_i in Γ.

Definition 2.13. *Let v_i be the i-th vertex of distance graph Γ. If v_i has m backwards edges in Γ to vertices $v_{i_1}, \ldots, v_{i_m}$ with $i_1 < \cdots < i_m < i$, define*

$$\underline{t}(i) = \underline{t}(i, \Gamma) = \left(|v_i - v_{i_1}|^2, \ldots, |v_i - v_{i_m}|^2\right).$$

Now if the tuple $(x_1, \ldots, x_{i-1})$ represents chosen vertices of an isometric copy of Γ we need notation to select those vertices which should be the backwards neighbors of x_i.

Definition 2.14. *Let $(x_1, \ldots, x_{i-1})$ be a tuple of points in $\mathbb{F}_q^d$. Let v_i be the i-th vertex of distance graph Γ. If v_i has m backwards edges in Γ to vertices $v_{i_1}, \ldots, v_{i_m}$ with $i_1 < \cdots < i_m < i$, define*

$$X(i) = X(i, x_1, x_2, \ldots, x_{i-1}, \Gamma) = (x_{i_1}, \ldots, x_{i_m}).$$

We will want to ensure each $X(i)$ is in general position for each i so that Lemma 2.11 is applicable. To that end we define a restricted average.

Definition 2.15. *Let k be a fixed integer and let $x_1, \ldots, x_n \in \mathbb{F}_q^d$.*

$$\mathbb{E}_x^* f(x) := \mathbb{E}_x^{k, x_1, \ldots, x_n} f(x) = q^{-d} \sum_{\substack{x \in \mathbb{F}_q^d, \\ x \cup X \text{ in general position for} \\ X \subset \{x_1, \ldots, x_n\} \text{ with } |X| = k-1}} f(x).$$

In other words $\mathbb{E}_x^*$ only considers those x's which can be included with $x_1, \ldots, x_n$ with the condition that any subset of size k of them are linearly independent. We make this definition to ensure each $X(i)$ is linearly independent when Γ is a k-degenerate distance graph. It is wasteful as really we only need to make sure a total of $n-1$ subsets of $x_1, \ldots, x_n$ of size at most k are linearly independent, but we will not worry about tracking a more precise estimate as it will be included in an error term in any case.

Lemma 2.16. *Let k be a fixed integer and let $x_1, \ldots, x_n \in \mathbb{F}_q^d$. Additionally, let $f : \mathbb{F}_q^d \to \mathbb{C}$, then*

$$|\mathbb{E}_x f(x) - \mathbb{E}_x^* f(x)| = O(q^{k-1-d}) \max_{x \in \mathbb{F}_q^d} |f(x)|.$$

Proof.

$$\begin{aligned}
|\mathbb{E}_x f(x) - \mathbb{E}_x^* f(x)| &\leq q^{-d} \sum_{\substack{x \in \langle x_{i_1}, \ldots, x_{i_{k-1}} \rangle, \\ \{x_{i_1}, \ldots, x_{i_{k-1}}\} \subset \{x_1, \ldots, x_n\}}} |f(x)| \\
&\leq q^{-d} \binom{n}{k-1} q^{k-1} \max_{x \in \mathbb{F}_q^d} |f(x)| \\
&= O(q^{k-1-d}) \max_{x \in \mathbb{F}_q^d} |f(x)|
\end{aligned}$$

□

We are finally in a position to define the averaged counting function Λ_Γ.

Definition 2.17. *Let Γ be a fixed distance graph. If $f_i : \mathbb{F}_q^d \to [-1, 1]$ for $i = 1, \ldots, n$, then we define*

$$\Lambda_\Gamma(f_1, \ldots, f_n) := \mathbb{E}_{x_1, \ldots, x_n}^* \prod_{i=1}^{n} f_i(x_i) \sigma_{\underline{t}(i)}^{X(i)}(x_i).$$

Note that $\underline{t}(1)$ and $X(1)$ are empty, so for notational convenience we let $\sigma_{\underline{t}(1)}^{X(1)}(y) := 1$.

For a set $A \subseteq \mathbb{F}_q^d$ if $\Lambda_\Gamma(1_A, \ldots, 1_A) > 0$, then A contains at least one isometric copy of distance graph Γ. We apply this counting function first to counting copies of Γ in the whole space $\mathbb{F}_q^d$.

Lemma 2.18. *For a fixed k-degenerate distance graph Γ in $\mathbb{F}_q^d$ where $d > 2k-1$ and $\operatorname{char}(\mathbb{F}_q) \neq 2$,*

$$\Lambda(1, \ldots, 1) = 1 + O(q^{(2k-1-d)/2}).$$

Proof. From Lemma 2.11 we have

$$\mathbb{E}_{x_i}\sigma_{\underline{t}(i)}^{X(i)}(x_i) = |\hat{\sigma}_{\underline{t}(i)}^{X(i)}(0)| = 1 + \mathcal{E}_i,$$

with $|\mathcal{E}_i| \leq 2q^{(2k-1-d)/2}$.

Additionally applying Lemma 2.16 gives

$$\mathbb{E}^*_{x_i}\sigma_{\underline{t}(i)}^{X(i)}(x_i) = 1 + \mathcal{E}_i + O(q^{2k-1-d}) = 1 + O(q^{(2k-1-d)/2}).$$

$$\begin{aligned}
\Lambda_\Gamma(1,\ldots,1) &= \mathbb{E}^*_{x_1,\ldots,x_n}\prod_{i=1}^{n}\sigma_{\underline{t}(i)}^{X(i)}(x_i) \\
&= \mathbb{E}^*_{x_1}\mathbb{E}^*_{x_2}\sigma_{\underline{t}(2)}^{X(2)}(x_2)\ldots\mathbb{E}^*_{x_n}\sigma_{\underline{t}(n)}^{X(n)}(x_n) \\
&= \prod_{i=2}^{n}\left(1 + O(q^{(2k-1-d)/2})\right) \\
&= 1 + O(q^{(2k-1-d)/2})
\end{aligned}$$

□

As before, if $A \subset \mathbb{F}_q^d$ with density $\alpha := q^{-d}|A|$, we define the *balance function* $f_A(x) := 1_A(x) - \alpha$.

Lemma 2.19. *Let* Γ *be a* k*-degenerate distance graph in* $\mathbb{F}_q^d$ *where* $d > 2k-1$ *and* $\operatorname{char}(\mathbb{F}_q) \neq 2$*, and let* $A \subset \mathbb{F}_q^d$ *with density* $\alpha \geq q^{\frac{3}{2}(2k-1-d)}$*. Then*

$$\Lambda_\Gamma(\underbrace{1_A,\ldots,1_A}_{j-1 \text{ copies}}, f_A, 1,\ldots,1) = \alpha^{\frac{1}{2}}O(q^{(2k-1-d)/4}).$$

Proof. First note that the computation in Lemma 2.18 gives us

$$\mathbb{E}^*_{x_{j+1},\ldots,x_n}\prod_{i=j+1}^{n}\sigma_{\underline{t}(i)}^{X(i)}(x_i) = 1 + O(q^{(2k-1-d)/2}),$$

where the $\mathbb{E}^*$ above is also in reference to some fixed $x_1,\ldots,x_j$. We can pull out this estimate to get

$$\Lambda_\Gamma(\underbrace{1_A,\ldots,1_A}_{j-1 \text{ copies}}, f_A, 1,\ldots,1) = (1 + O(q^{(2k-1-d)/2}))\Lambda_\Gamma(\underbrace{1_A,\ldots,1_A}_{j-1 \text{ copies}}, f_A).$$

In the following, let b be the index of the first of the backwards neighbors of v_j. Abbreviating $\Lambda_\Gamma(\underbrace{1_A, \ldots, 1_A}_{j-1 \text{ copies}}, f_A)$ as Λ, we have

$$\begin{aligned}
\Lambda &= \mathbb{E}^*_{x_1,\ldots,x_{j-1}} \left(\prod_{i=1}^{j-1} 1_A(x_i) \sigma^{X(i)}_{\underline{t}(i)}(x_i) \right) \mathbb{E}^*_{x_j} f_A(x_j) \sigma^{X(j)}_{\underline{t}(j)}(x_j) \\
&= \mathbb{E}^*_{x_1,\ldots,x_{j-1}} \left(\prod_{i=1}^{j-1} 1_A(x_i) \sigma^{X(i)}_{\underline{t}(i)}(x_i) \right) \mathbb{E}^*_{x_j} f_A(x_j) \sigma^{Z(j)}_{\underline{t}(j)}(x_b - x_j) \\
&= \mathbb{E}^*_{x_1,\ldots,x_{j-1}} \left(\prod_{i=1}^{j-1} 1_A(x_i) \sigma^{X(i)}_{\underline{t}(i)}(x_i) \right) \left[f_A * \sigma^{Z(j)}_{\underline{t}(j)}(x_b) + O(q^{2k-1-d}) \right].
\end{aligned}$$

Using Lemma 2.18 again for the error term:

$$|\Lambda| \leq O(q^{2k-1-d}) + \mathbb{E}^*_{x_1,\ldots,x_{j-1}} \left(\prod_{i=1}^{j-1} \sigma^{X(i)}_{\underline{t}(i)}(x_i) \right) \left| f_A * \sigma^{Z(j)}_{\underline{t}(j)}(x_b) \right|.$$

Let

$$S = \mathbb{E}^*_{x_1,\ldots,x_{j-1}} \left(\prod_{i=1}^{j-1} \sigma^{X(i)}_{\underline{t}(i)}(x_i) \right) \left| f_A * \sigma^{Z(j)}_{\underline{t}(j)}(x_b) \right|.$$

We now rewrite the expectation with respect to the basepoint x_1, i.e. write $x_i = x_1 + y_i$ for $i = 1, \ldots, j-1$. Observe that under this rewriting

$$\sigma^{X(i)}_{\underline{t}(i)}(x_i) = \sigma^{X(i)}_{\underline{t}(i)}(x_1 + y_i) = \sigma^{Z(i)}_{\underline{t}(i)}(y_a - y_i),$$

where a is the index of the first backwards neighbor of v_i in Γ, and $Z(i) = (0, y_a - y_{i_2}, \ldots, y_a - y_{i_m})$. Noting $\sigma^{X(1)}_{\underline{t}(1)}(x_1) = 1$, each of the $\sigma^{X(1)}_{\underline{t}(1)}(x_1) = 1$ is independent of x_1. Under the rewriting, the $*$ condition in the expectation transforms from the general position of any subset of size k of $\{x_1, \ldots, x_{j-1}\}$ to the condition that any subset of size k of $\{y_2, \ldots, y_{j-1}\}$ is linearly independent. This is also independent of x_1, so we can pull the expectation in x_1 inside the sum to attain

$$\begin{aligned}
S &= \mathbb{E}^*_{y_2\ldots,y_{j-1}} \left(\prod_{i=2}^{j-1} \sigma^{Z(i)}_{\underline{t}(i)}(y_{i_1} - y_i) \right) \mathbb{E}_{x_1} \left| f_A * \sigma^{Z(j)}_{\underline{t}(j)}(x_1 + y_b) \right| \\
&= \mathbb{E}^*_{y_2\ldots,y_{j-1}} \left(\prod_{i=2}^{j-1} \sigma^{Z(i)}_{\underline{t}(i)}(y_{i_1} - y_i) \right) \mathbb{E}_{x_1} \left| f_A * \sigma^{Z(j)}_{\underline{t}(j)}(x_1) \right|.
\end{aligned}$$

To condense notation, let

$$\mathcal{P} = \prod_{i=2}^{j-1} \sigma_{\underline{t}(i)}^{Z(i)}(y_{i_1} - y_i).$$

Now by Cauchy-Schwarz on the functions

$$\mathcal{P}^{\frac{1}{2}} \quad \text{and} \quad \mathcal{P}^{\frac{1}{2}} \mathbb{E}_{x_1} \left| f_A * \sigma_{\underline{t}(j)}^{Z(j)}(x_1) \right|,$$

and applying Lemma 2.18 to get

$$\mathbb{E}^*_{y_2 \ldots, y_{j-1}} \mathcal{P} = 1 + O(q^{(2k-1-d)/2}), \tag{2.6}$$

we have

$$S^2 \leq \left(1 + O(q^{(2k-1-d)/2})\right) T,$$

where

$$T = \mathbb{E}^*_{y_2 \ldots, y_{j-1}} \mathcal{P} \mathbb{E}_{x_1} \left| f_A * \sigma_{\underline{t}(j)}^{Z(j)}(x_1) \right|^2.$$

We next rewrite T using Plancherel's identity:

$$T = \mathbb{E}^*_{y_2 \ldots, y_{j-1}} \mathcal{P} \sum_{r \in \mathbb{F}_q^d} |\hat{f}_A(r)|^2 |\hat{\sigma}_{\underline{t}(j)}^{Z(j)}(r)|^2.$$

We can now apply the computation in Lemma 2.11, that

$$|\hat{\sigma}_{\underline{t}(j)}^{Z(j)}(r)| = 1_{\langle z_2, \ldots, z_m \rangle}(r) + O(q^{(2k-1-d)/2}),$$

where $z_i = y_{j_1} - y_{j_i}$ for $i = 2, \ldots, m$.

We therefore split the sum in the above expression on the condition that $r \in \langle z_2, \ldots, z_m \rangle$, so we may write

$$T = T_1 + T_2,$$

where

$$T_1 = \mathbb{E}^*_{y_2 \ldots, y_{j-1}} \mathcal{P} \sum_{r \in \langle z_2, \ldots, z_m \rangle} |\hat{f}_A(r)|^2 \left(1 + O(q^{(2k-1-d)/2})\right),$$

and

$$T_2 = O(q^{2k-1-d}) \mathbb{E}^*_{y_2 \ldots, y_{j-1}} \mathcal{P} \sum_{r \notin \langle z_2, \ldots, z_m \rangle} |\hat{f}_A(r)|^2.$$

Note

$$\sum_{r\in\mathbb{F}_q^d}|\hat{f}_A(r)|^2 = E_x|f_A(x)|^2 = \alpha(1-\alpha)\le\alpha, \tag{2.7}$$

and applying this estimate and equation (2.6), we can dispatch with T_2:

$$T_2 = \alpha O(q^{2k-1-d}).$$

It remains to estimate T_1. First we can again use equation (2.6) for the error and write

$$T_1 = \alpha O(q^{(2k-1-d)/2}) + \mathbb{E}^*_{y_2\dots,y_{j-1}}\mathcal{P}\sum_{r\in\langle z_2,\dots,z_m\rangle}|\hat{f}_A(r)|^2.$$

Let us consider just the last part of this expression

$$E := \mathbb{E}^*_{y_2\dots,y_{j-1}}\mathcal{P}\sum_{r\in\langle z_2,\dots,z_m\rangle}|\hat{f}_A(r)|^2.$$

Observe that if $r=0$, then $E=0$ as $\hat{f}_A(0)=\mathbb{E}_x f_A(x)=0$. If $r\neq 0$, then r only appears in some of the $\langle z_2,\dots,z_m\rangle$, and the actual value of r turns out to be immaterial. Examining $\mathcal{P}$ again, and recognizing that it is an indicator function of isometric copies of Γ restricted to the vertices $v_1,\dots,v_{j-1}$, we can recognize that it is invariant under any distance preserving transformation $u\in SO\left(\mathbb{F}_q^d\right)$. Thus

$$E = \mathbb{E}^*_{y_2\dots,y_{j-1}}\mathcal{P}\sum_{u^{-1}(r)\in\langle z_2,\dots,z_m\rangle}|\hat{f}_A(u^{-1}(r))|^2.$$

We can now average over $SO\left(\mathbb{F}_q^d\right)$ as is demonstrated in the proof of Proposition 1 in [LM20a]. The average can be pulled into the expression to obtain

$$\begin{aligned}
E &= \mathbb{E}^*_{y_2\dots,y_{j-1}}\mathcal{P}\mathbb{E}_{u\in SO(\mathbb{F}_q^d)}\sum_{u(r)\in\langle z_2,\dots,z_m\rangle}|\hat{f}_A(u(r))|^2\\
&\le \mathbb{E}^*_{y_2\dots,y_{j-1}}\mathcal{P}q^{m-1-d}\sum_{r\in\mathbb{F}_q^d}|\hat{f}_A(r)|^2\\
&\le \alpha O(q^{m-1-d})\\
&\le \alpha O(q^{k-1-d}),
\end{aligned}$$

where in the penultimate step we applied equation (2.6) and equation (2.7), and in the last step we recalled that m was the backwards degree of vertex v_j which is at most k as Γ is k-degenerate. Hence

$$T_1 = \alpha O(q^{(2k-1-d)/2}) + \alpha O(q^{k-1-d}) = \alpha O(q^{(2k-1-d)/2}).$$

Putting everything back together, we have

$$T = T_1 + T_2 = \alpha O(q^{(2k-1-d)/2}) + \alpha O(q^{2k-1-d}) = \alpha O(q^{(2k-1-d)/2}),$$

and

$$S = \alpha^{\frac{1}{2}} O(q^{(2k-1-d)/4}),$$

so

$$\Lambda = O(q^{2k-1-d}) + \alpha^{\frac{1}{2}} O(q^{(2k-1-d)/4}).$$

When $\alpha \geq q^{\frac{3}{2}(2k-1-d)}$, the second term dominates, so

$$\Lambda = \alpha^{\frac{1}{2}} O(q^{(2k-1-d)/4}),$$

and

$$\Lambda_\Gamma(\underbrace{1_A, \ldots, 1_A}_{j-1 \text{ copies}}, f_A, 1, \ldots, 1) = \alpha^{\frac{1}{2}} O(q^{(2k-1-d)/4})$$

as desired. □

The proof of the main theorem follows immediately from this lemma and the linearity of Λ_Γ.

Theorem 2.20. *Let Γ be a k-degenerate distance graph in $\mathbb{F}_q^d$ with n vertices and nonzero edge lengths where $\mathrm{char}(\mathbb{F}_q^d) \neq 2$ and $d > 2k - 1$. Then there exists a constant $C = C(n, k)$ such that if $A \subseteq F_q^d$ with density $\alpha \geq Cq^{\frac{2k-1-d}{4n-2}}$, then A contains an isometric copy of Γ.*

Proof. By the linearity of Λ and an application of Lemmas 2.18 and 2.19,

$$\begin{aligned}\Lambda(1_A, \ldots, 1_A) &= \alpha^n \Lambda(1, \ldots, 1) + \sum_{j=1}^{n} \alpha^{n-j} \Lambda(\underbrace{1_A, \ldots, 1_A}_{j-1 \text{ copies}}, f_A, 1, \ldots, 1) \\ &= \alpha^n(1 + O(q^{(2k-1-d)/2})) + \frac{1-\alpha^n}{1-\alpha} \alpha^{\frac{1}{2}} O(q^{(2k-1-d)/4}) \\ &= \alpha^n + \alpha^{\frac{1}{2}} O(q^{(2k-1-d)/4}).\end{aligned}$$

If $\alpha \geq Cq^{\frac{2k-1-d}{4n-2}}$ where C is a constant depending only on n and k, then $\Lambda(1_A, \ldots, 1_A) > 0$ and A necessarily contains an isometric copy of Γ. Note that C could be explicitly computed from Lemmas 2.18 and 2.19, taking the minimal choices for the other parameters, i.e. $q = 3$ and $d = 2k$. □

2.3 Exceptional Set Approach

In this approach, we use the idea that most of the time the number of ways it is possible to add an extra point $x \in A$ connected by k edges to a distance graph $\Gamma \subset \mathbb{F}_q^d$ is about as many as expected just considering the density of A. Specifically, if k points $x_1, x_2, \ldots, x_k$ are fixed in $\mathbb{F}_q^d$, and k nonzero edge lengths $t_1, t_2, \ldots, t_k$ are specified, then there are $\approx q^{d-k}$ points x in $\mathbb{F}_q^d$ satisfying $|x-x_1|^2 = t_1, \ldots, |x-x_k|^2 = t_k$. If the density of $A \subset \mathbb{F}_q^d$ is α, we expect that αq^{d-k} points satisfying the restrictions also lie in A. This idea is made precise in the following theorem.

Theorem 2.21. *Fix $\varepsilon > 0$ and $A \subset \mathbb{F}_q^d$ with density α where $\operatorname{char}(\mathbb{F}_q) \neq 2$. Let $\underline{t} \in \left(\mathbb{F}_q^*\right)^k$ be a tuple of k distances, and $X \in \left(\mathbb{F}_q^d\right)^k$ be a tuple of k points in $\mathbb{F}_q^d$ in general position with $d > 2k-1$. Then if X is not in an exceptional set B_ε then*

$$|A \cap S_{\underline{t}}(X)| \geq (\alpha - \varepsilon) q^{d-k} - 2\alpha q^{\frac{d-1}{2}},$$

where $|B_\varepsilon| < 6\alpha\varepsilon^{-2} q^{kd+2k-1-d}$.

Proof. In order to write $|A \cap S_{\underline{t}}(X)|$ as a convolution, it is useful to write the points as relative to the chosen basepoint x_1. Now we can compute

$$|A \cap S_{\underline{t}}(X)| = \sum_{y \in \mathbb{F}_q^d} 1_A(x_1 + y) \prod_{i=1}^{k} q^{-1} \sigma_{t_i}(x_i - x_1 - y).$$

This is the desired count as $1_A(x_1 + y)$ checks if $x_1 + y$ is in A while each $q^{-1}\sigma_{t_i}(x_i - x_1 - y)$ checks if x_i is the appropriate distance from $x_1 + y$.

To see this expression as a convolution, let $Z = (z_1, \ldots, z_k)$ with $z_i := x_i - x_1$ for $i = 1, \ldots, k$, and define

$$\sigma_{\underline{t}}^Z(y) := \prod_{i=1}^{k} \sigma_{t_i}(z_i + y).$$

Then

$$\begin{aligned}
|A \cap S_{\underline{t}}(X)| &= q^{-k} \sum_{y \in \mathbb{F}_q^d} 1_A(x_1 + y) \sigma_{\underline{t}}^Z(-y) \\
&= q^{d-k} 1_A * \sigma_{\underline{t}}^Z(x_1) \\
&= q^{d-k} (\alpha + f_A) * \sigma_{\underline{t}}^Z(x_1) \\
&= \alpha |S_{\underline{t}}(X)| + q^{d-k} f_A * \sigma_{\underline{t}}^Z(x_1),
\end{aligned}$$

where as usual f_A is the balance function, defined as $f_A = 1_A - \alpha$.

By Corollary 2.12 we know $|S_{\underline{t}}(X)| = q^{d-k} + \mathcal{E}$ where $|\mathcal{E}| \le 2q^{\frac{d-1}{2}}$, so we have

$$|A \cap S_{\underline{t}}(X)| = \alpha q^{d-k} + q^{d-k} f_A * \sigma_{\underline{t}}^Z(x_1) + \alpha\mathcal{E}.$$

Define the exceptional set

$$B_\varepsilon := \{X \in \left(\mathbb{F}_q^d\right)^k \ : \ |f_A * \sigma_{\underline{t}}^Z(x_1)| > \varepsilon, \ X \text{ in general position}\}.$$

We can make use of Plancherel's identity and estimates of the Fourier transform of the sphere to get a handle on the size of B_ε. For the following, we shall write $\sum_Z^*$ to mean the sum over $Z = (z_2, \dots, z_k) \in (\mathbb{F}_q^d)^{k-1}$ where the elements of Z are linearly independent. This makes sense as $z_1 = x_1 - x_1 = 0$, but every other $\{z_i\}_{i=2}^k$ varies freely over $\mathbb{F}_q^d$ as x_i varies, with the restriction that $z_2, \dots, z_k$ are linearly independent since X is in general position. Now

$$\begin{aligned}
|B_\varepsilon|\varepsilon^2 &\le \sum_{\substack{X \in (\mathbb{F}_q^d)^k, \\ X \text{ in general position}}} |f_A * \sigma_{\underline{t}}^Z(x_1)|^2 \\
&\le \sum_Z^* \sum_{x_1 \in \mathbb{F}_q^d} |f_A * \sigma_{\underline{t}}^Z(x_1)|^2 \\
&\le \sum_Z^* q^d \mathbb{E}_{x_1} |f_A * \sigma_{\underline{t}}^Z(x_1)|^2 \\
&\le q^d \sum_Z^* \sum_{r \in \mathbb{F}_q^d} |\hat{f}_A(r)|^2 |\hat{\sigma}_{\underline{t}}^Z(r)|^2 \\
&\le q^d I_1 + q^d I_2,
\end{aligned}$$

where

$$I_1 = \sum_Z^* \sum_{r \notin \langle z_2, \dots, z_k \rangle} |\hat{f}_A(r)|^2 |\hat{\sigma}_{\underline{t}}^Z(r)|^2, \text{ and}$$

$$I_2 = \sum_Z^* \sum_{r \in \langle z_2, \dots, z_k \rangle} |\hat{f}_A(r)|^2 |\hat{\sigma}_{\underline{t}}^Z(r)|^2.$$

As $z_2, \dots, z_k$ are linearly independent, Lemma 2.11 applies:

$$|\hat{\sigma}_{\underline{t}}^Z(r)| = 1_{r \in \langle z_2, \dots, z_k \rangle}(r) + \mathcal{E}(r),$$

with $|\mathcal{E}(r)| \le 2q^{(2k-1-d)/2}$, and for I_1, we get the estimate

$$I_1 = \sum_Z^* \sum_{r \notin \langle z_2, \dots, z_k \rangle} |\hat{f}_A(r)|^2 |\hat{\sigma}_{\underline{t}}^Z(r)|^2$$

$$\leq \sum_Z{}^* \sum_{r \notin \langle z_2, \ldots, z_k \rangle} |\hat{f}_A(r)|^2 4q^{2k-1-d}$$
$$\leq 4q^{2k-1-d} \sum_Z{}^* \sum_{r \in \mathbb{F}_q^d} |\hat{f}_A(r)|^2$$
$$\leq 4q^{2k-1-d} \sum_Z{}^* \mathbb{E}_x |f_A(x)|^2$$
$$\leq 4q^{2k-1-d} q^{d(k-1)} \alpha(1-\alpha)$$
$$\leq 4\alpha q^{dk+2k-1-2d}.$$

On the other hand,

$$I_2 = \sum_Z{}^* \sum_{r \in \langle z_2, \ldots, z_k \rangle} |\hat{f}_A(r)|^2 |\hat{\sigma}_{\underline{t}}^Z(r)|^2$$
$$\leq \left(1 + 4q^{(2k-1-d)/2} + 4q^{2k-1-d}\right) \sum_Z{}^* \sum_{r \in \langle z_2, \ldots, z_k \rangle} |\hat{f}_A(r)|^2.$$

This last line can be rewritten by observing each nonzero r appears an equal number of times in the sum.[21] Indeed if $r_1, r_2 \neq 0$, then let $A \in \mathrm{GL}(d, q)$ be some matrix sending r_1 to r_2 (there are many such matrices, though to make it explicit, one could construct a change of basis matrix). Now if $Z = (z_2, \ldots, z_k)$ is fixed and $r_1 \in \langle z_2, \ldots, z_k \rangle$, then

[21] A very similar observation and calculation occurred in the calculation of E in the proof of Lemma 2.19.

$$r_1 = \sum_{m=2}^{k} c_m z_m \iff r_2 = \sum_{m=2}^{k} c_m A z_m,$$

i.e. $r_1 \in \langle z_2, \ldots, z_k \rangle \iff r_2 \in \langle Az_2, \ldots, Az_k \rangle$. Therefore, the number of Z for which $r \in \langle z_2, \ldots, z_k \rangle$ is invariant for nonzero r. We can compute explicitly for $r \neq 0$

$$\#\{Z : r \in \langle z_2, \ldots, z_k \rangle\} = \frac{(q^{k-1} - 1) \prod_{m=0}^{k-2} (q^d - q^m)}{q^d - 1} \leq q^{k-1+d(k-2)}.$$

The one exception, $r = 0$, which is in every $\langle z_2, \ldots, z_k \rangle$ does not actually contribute to the sum for I_2 as $\hat{f}_A(0) = \mathbb{E}_x f_A(x) = 0$. Thus

$$I_2 \leq \left(1 + 4q^{(2k-1-d)/2} + 4q^{2k-1-d}\right) \sum_Z{}^* \sum_{r \in \langle z_2, \ldots, z_k \rangle} |\hat{f}_A(r)|^2$$
$$\leq \left(1 + 4q^{(2k-1-d)/2} + 4q^{2k-1-d}\right) q^{dk+k-1-2d} \sum_{r \in \mathbb{F}_q} |\hat{f}_A(r)|^2$$

$$\leq \alpha q^{dk+k-1-2d} + 4\alpha q^{dk+2k-3/2-5d/2} + 4\alpha q^{dk+3k-2-3d}.$$

Comparing the exponents, we see the I_1 term is the main term, but specifically, we get

$$\begin{aligned}
|B_\varepsilon|\varepsilon^2 &\leq q^d I_1 + q^d I_2 \\
&\leq \alpha q^{dk-2k-1-d}\left(4 + q^{-k} + 4q^{(-1-d)/2} + 4q^{k-1-d}\right) \\
&\leq 6\alpha q^{dk-2k-1-d},
\end{aligned}$$

where in the last line we use $d > 2k-1$, $k \geq 1$, and $q \geq 3$. □

The above is the crucial working theorem, though to put all the pieces together, we also make use of an easy upper bound computation for the number of isometric copies of a k-degenerate distance graph in the whole space $\mathbb{F}_q^d$.

If $\Gamma \subset F_q^d$ is a k-degenerate distance graph in k-general position, let

$$\mathcal{N}(\Gamma) := \#\{\Gamma' \subset \mathbb{F}_q^d \,:\, \Gamma' \sim \Gamma\}$$

where by $\Gamma' \sim \Gamma$ we mean both that Γ' is isometric to Γ and that Γ' is in k-general position.

Lemma 2.22. *Let $\Gamma \subset \mathbb{F}_q^d$ be a k-degenerate distance graph with $\mathrm{char}(\mathbb{F}_q) \neq 2$ and $d > 2k-1$. If Γ is in k-general position with n vertices and e edges, then*

$$\mathcal{N}(\Gamma) \leq q^{nd-e}\prod_{i=1}^{n}(1 + 2q^{(2m_i-1-d)/2}) \leq q^{nd-e}(1+2q^{-1/2})^n,$$

where m_i is the backwards degree of the i-th vertex.

Proof. The proof is a direct inductive argument. If Γ is the one vertex graph, then the number of isometric copies of Γ is q^d. If Γ has vertices $v_1, \ldots, v_{n+1}$ with $n \geq 1$, we can count the number of isometric copies of Γ by considering how many ways there are to add the vertex v_{n+1} to an isometric copy of the induced subgraph Γ_0 on vertices $v_1, \ldots, v_n$. Let e be the number of edges in Γ, let e_0 be the number of edges in Γ_0, and let m be the degree of v_{n+1} in Γ. Note that $e_0 + m = e$, and $m \leq k$ as Γ is k-degenerate. For induction, we now assume

$$\mathcal{N}(\Gamma_0) \leq q^{nd-e_0}\prod_{i=1}^{n}(1 + 2q^{(2m_i-1-d)/2}).$$

Fix some particular $\Gamma_0' \sim \Gamma_0$ and let $v_{i_1}, \ldots, v_{i_m}$ be the vertices connected to v_{n+1} in Γ, and define

$$X := (v_{i_1}(\Gamma_0'), \ldots, v_{i_m}(\Gamma_0')) .$$

Let $\underline{t} = (t_1, \ldots, t_m)$ be the lengths of the edges connected to v_{n+1} in Γ, i.e. $t_j = |v_{n+1}(\Gamma) - v_{i_j}(\Gamma)|^2$ for $j = 1, \ldots, m$.

Again by Corollary 2.12 we know $|S_{\underline{t}}(X)| \leq q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right)$, and so for each Γ_0', we can add v_{n+1} in at most $q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right)$ ways.

Therefore

$$\begin{aligned}\mathcal{N}(\Gamma) &\leq \mathcal{N}(\Gamma_0) q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right) \\ &\leq q^{nd-e_0} q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right) \prod_{i=1}^{n}(1 + 2q^{(2m_i-1-d)/2}) \\ &\leq q^{(n+1)d-e} \prod_{i=1}^{n+1}(1 + 2q^{(2m_i-1-d)/2}).\end{aligned}$$

Noting $d > 2k - 1 \geq 2m - 1$ gives

$$\prod_{i=1}^{n+1}(1 + 2q^{(2m_i-1-d)/2}) \leq \left(1 + 2q^{-1/2}\right)^n .$$

□

We will also make use of the count of the number of isometric copies of Γ in which certain vertices are fixed.

Definition 2.23. *Let $I = (i_1, \ldots, i_m)$ with $i_j \in \{1, \ldots, n\}$ be an index tuple and let $X = (x_1, \ldots, x_m)$ with $x_j \in \mathbb{F}_q^d$ be a tuple of points. Then set*

$$\mathcal{N}_I(\Gamma, X) := \#\{\Gamma' \subset \mathbb{F}_q^d \,:\, \Gamma' \sim \Gamma,\ v_{i_1}(\Gamma') = x_1, \ldots, v_{i_m}(\Gamma') = x_m\}.$$

The following lemma follows the same proof idea as Lemma 2.22, though needs to pay the minor price of assuming Γ is in $2k$-general position.

Lemma 2.24. *Let $\Gamma \subset F_q^d$ be a k-degenerate distance graph with $\operatorname{char}(\mathbb{F}_q) \neq 2$ and $d > 4k - 1$. Additionally assume Γ is in $2k$-general position with n vertices and e edges. Let $I = (i_1, \ldots, i_s)$ be an index tuple with $s \leq k$, let e_I be the number of edges in the induced subgraph on vertices $v_{i_1}, \ldots, v_{i_s}$, and let $X = (x_1, \ldots, x_s)$ be a tuple of points $x_j \in \mathbb{F}_q^d$, then*

$$\mathcal{N}_I(\Gamma, X) \leq q^{(n-s)d-e+e_I}\left(1 + 2q^{-1/2}\right)^{n-s} .$$

Proof. We immediately consider a reindexed version of Γ where all the edges remain the same, but we move the s vertices $(v_{i_1}, \ldots, v_{i_s})$ to the beginning of the ordering, i.e. let $\widetilde{\Gamma} = (v_{i_1}, \ldots, v_{i_s}, v_1, \ldots, v_n)$ where with an abuse of notation $v_1, \ldots, v_n$ is meant to skip over those indices in I. Because s vertices have been brought forward in the degenerecy ordering, to each vertex in Γ we have added at most s new backwards edges. Hence $\widetilde{\Gamma}$ is a $(k+s)$-degenerate graph. In particular as $s \leq k$, it is a $2k$-degenerate graph. We observe $\mathcal{N}_I(\Gamma, X) = \mathcal{N}_{[s]}(\widetilde{\Gamma}, X)$, where by $[s]$ we mean $\{1, \ldots, s\}$.

We start the induction with $n = s$. If Γ has s vertices, then $e = e_I$ and

$$\mathcal{N}_{[s]}(\widetilde{\Gamma}, X) \leq 1.$$

Note that if the tuple of points in X fail to satisfy the edge restrictions of $\widetilde{\Gamma}$, then $\mathcal{N}_{[s]}(\widetilde{\Gamma}, X) = 0$, but otherwise there is exactly one isometric graph because we have fixed all vertices.

The induction proceeds exactly as in the previous lemma, though unfortunately the notation is cumbersome.

If Γ has $n+1$ vertices with $n \geq s$, then $\widetilde{\Gamma}$ has $n+1$ vertices, and we count how many ways there were to add an $(n+1)$-th vertex to $\widetilde{\Gamma}_0$, the $2k$-degenerate distance graph induced by removing the $(n+1)$-th vertex from $\widetilde{\Gamma}$. Again, let e be the number of edges in $\widetilde{\Gamma}$, let e_0 be the number of edges in $\widetilde{\Gamma}_0$, and let m be the degree of v_{n+1} in $\widetilde{\Gamma}$. By the inductive assumption,

$$\mathcal{N}_{[s]}(\widetilde{\Gamma}_0, X) \leq q^{(n-s)d - e_0 + e_I} \left(1 + 2q^{-1/2}\right)^{n-s}.$$

For each copy of $\widetilde{\Gamma}_0$, we can add v_{n+1} in at most $q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right)$ ways and therefore

$$\begin{aligned}
\mathcal{N}_I(\Gamma, X) &= \mathcal{N}_{[s]}(\widetilde{\Gamma}, X) \\
&\leq \mathcal{N}_{[s]}(\widetilde{\Gamma}_0, X) q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right) \\
&\leq q^{(n-s)d - e_0 + e_I} q^{d-m}\left(1 + 2q^{(2m-1-d)/2}\right)\left(1 + 2q^{-1/2}\right)^{n-s} \\
&\leq q^{(n+1-s)d - e + e_I}\left(1 + 2q^{-1/2}\right)^{n+1-s},
\end{aligned}$$

where the last inequality arises as $m \leq 2k$ and $d > 4k - 1$ by assumption. □

We are now in a position to prove the main theorem for finite fields, though we need the final notation

$$\mathcal{N}^A(\Gamma) := \#\{\Gamma' \subset A \,:\, \Gamma' \sim \Gamma\}.$$

Theorem 2.25. *Let* Γ *be a* k*-degenerate distance graph in* $2k$*-general position on* n *vertices with* e *edges and nonzero edge lengths in* $\mathbb{F}_q^d$ *with* $\mathrm{char}(\mathbb{F}_q) \neq 2$. *Fix* $0 < \delta < \frac{1}{3}$ *and let* $A \subset \mathbb{F}_q^d$ *with density* α. *If* $q > \max\{2^{1/\delta}, 2^{2/(1-2\delta)}\}$, $d > k(k-1)/2 + 2k - 1$, *and*

$$\alpha > 6c_{\delta,q}^{-1}\left(1 + 2q^{-1/2}\right) q^{\frac{k(k-1)/2+2k-1-d+3\delta}{n+1}},$$

then

$$\mathcal{N}^A(\Gamma) \geq c_{\delta,q}^n \alpha^n q^{nd-e},$$

where

$$c_{\delta,q} = \left(1 - 2q^{-\delta}\right)\left(1 - q^{-\delta}\right).$$

Proof. We again use induction. If Γ is the single-vertex distance graph, then A contains αq^d isometric copies of Γ. Now let Γ be a k-degenerate distance graph on $n+1$ vertices with nonzero edgelengths in $2k$-general position with $n \geq 1$. Assume that $\deg(v_{n+1}) = m \leq k$, and as before, let Γ_0 be the induced subgraph of Γ on the first n vertices, and let e_0 denote the number of edges in Γ_0. Note that if e is the number of edges in Γ, then $e_0 + m = e$. The inductive assumption says

$$\mathcal{N}^A(\Gamma_0) \geq c_{\delta,q}^n \alpha^n q^{nd-e_0}.$$

We proceed by counting the way to add an $(n+1)$-th vertex to each isometric copy of Γ_0 provided the would-be neighbors of the $(n+1)$-th vertex do not fall in a bad set. Let $\underline{t} \in \left(\mathbb{F}_q^*\right)^m$ be the tuple of distances from the $(n+1)$-th vertex in Γ. Fixing $0 < \delta < \frac{1}{3}$ and setting $\varepsilon = \alpha q^{-\delta}$ in Theorem 2.21 gives

$$|A \cap S_{\underline{t}}(X)| \geq \alpha q^{d-m}\left(1 - q^{-\delta} - 2q^{(2m-1-d)/2}\right) \qquad \text{for } X \notin B,$$

where $|B| < 6\alpha^{-1}q^{md+2m-1-d+2\delta}$. Noting $k \geq m$ and $d > 2k-1$ we can replace $q^{(2m-1-d)/2}$ with $q^{-1/2}$. Then noting we carefully chosen q such that $q > 2^{2/(1-2\delta)}$ which implies $2q^{-1/2} < q^{-\delta}$, we can write

$$|A \cap S_{\underline{t}}(X)| \geq \alpha q^{d-m}\left(1 - 2q^{-\delta}\right) \qquad \text{for } X \notin B.$$

The other requirement $q > 2^{1/\delta}$ ensures the term $\left(1 - 2q^{-\delta}\right)$ is positive.

Setting I to be the indices of the neighbors of the $(n+1)$-th vertex in Γ, we are left with at least

$$\mathcal{N}^A(\Gamma_0) - \sum_{X \in B} \mathcal{N}_I(\Gamma_0, X) \geq \mathcal{N}^A(\Gamma_0) - |B| \max_{X \in B} \mathcal{N}_I(\Gamma_0, X)$$

isometric copies of Γ_0 for which we have about αq^{d-m} options for adding a $(n+1)$-th vertex. From Lemma 2.24, we have

$$\mathcal{N}_I(\Gamma_0, X) \leq q^{(n-m)d-e_0+e_I}\left(1+2q^{-1/2}\right)^{n-m},$$

and so

$$\begin{aligned}|B| \max_{X\in B} \mathcal{N}_I(\Gamma_0, X) &\leq \alpha^{-1} 6\left(1+2q^{-1/2}\right)^{n-m} q^{nd-e_0+e_I+2m-1-d+2\delta} \\ &\leq \alpha^n q^{nd-e_0} \frac{6\left(1+2q^{-1/2}\right)^{n-m} q^{e_I+2m-1-d+2\delta}}{\alpha^{n+1}} \\ &\leq c_{\delta,q}^n \alpha^n q^{nd-e_0} q^{-\delta},\end{aligned}$$

where in the last step we have applied the lower bound of α.

Putting everything together yields

$$\begin{aligned}\mathcal{N}^A(\Gamma) &\geq \alpha q^{d-m}\left(1-2q^{-\delta}\right)\left(\mathcal{N}^A(\Gamma_0) - \sum_{X\in B} \mathcal{N}_I(\Gamma_0, X)\right) \\ &\geq \alpha q^{d-m}\left(1-2q^{-\delta}\right)\left(c_{\delta,q}^n \alpha^n q^{nd-e_0} - c_{\delta,q}^n \alpha^n q^{nd-e_0} q^{-\delta}\right) \\ &\geq c_{\delta,q}^n \alpha^{n+1} q^{(n+1)d-e}\left(1-2q^{-\delta}\right)\left(1-q^{-\delta}\right) \\ &\geq c_{\delta,q}^{n+1} \alpha^{n+1} q^{(n+1)d-e}\end{aligned}$$

□

The above theorem is notation dense, so we record a particular case to emphasize the nature of the result.

Corollary 2.26. *Let Γ be a k-degenerate distance graph in $2k$-general position on n vertices with e edges and nonzero edge lengths in $\mathbb{F}_q^d$ with $\operatorname{char}(\mathbb{F}_q) \neq 2$. Let $A \subset \mathbb{F}_q^d$ with density α. If $q > 1000$, $d > k(k-1)/2 + 2k - 1$, and*

$$\alpha > 26 q^{\frac{k(k-1)/2+2k-1/2-d}{n+1}},$$

then

$$\mathcal{N}^A(\Gamma) \geq \left(\frac{1}{4}\right)^n \alpha^n q^{nd-e}.$$

Proof. The corollary is the above theorem after setting $\delta = \frac{1}{6}$ and choosing 1000 as a threshold for q. Note that it would be possible to get better lower bounds for larger q as $c_{\delta,q} \to 1$ as $q \to \infty$. □

Chapter 3

Upper Bounds

3.1 Introduction

In this chapter, we attain upper bounds on the number of lattice points on the intersections of spheres using the machinery of the circle method. We first search for solutions in all of $\mathbb{Z}^d$ and then restrict to looking for solutions in a fixed residue class. In particular, we estimate $|S_{\underline{x}}(\lambda,\mu)|$ and $|S_{\underline{x},W,\underline{v}}(\lambda,\mu)|$ as defined below.

Definition 3.1. *Given* $\lambda^2, \mu \in \mathbb{Z}$ *and* $\underline{x} \in \mathbb{Z}^d$*,*

$$S_{\underline{x}}(\lambda,\mu) := \{\underline{y} \in \mathbb{Z}^d \ : \ |\underline{y}|^2 = \lambda^2,\ \underline{y}\cdot\underline{x} = \mu\}.$$

Definition 3.2. *Given* $\lambda^2, \mu \in \mathbb{Z}$*,* $\underline{x} \in \mathbb{Z}^d$*, and residue class* $\underline{v} \pmod{W}$*,*

$$S_{\underline{x},W,\underline{v}}(\lambda,\mu) := \{\underline{y} \in \mathbb{Z}^d \ : |\underline{y}|^2 = \lambda^2,\ \underline{y}\cdot\underline{x} = \mu,\ \underline{y} \equiv \underline{v} \pmod{W}\}.$$

We will also only consider those $\underline{x}$ of the following simple form.[22]

Definition 3.3. $\underline{x} = (x_1,\dots,x_d)$ *is primitive if* $\gcd(x_1,\dots,x_d) = 1$.

[22] This is because some of the crucial estimates depend on $\beta\underline{x}$ not being too close to lattice points for $\beta \in [0,1]$.

3.2 Upper Bounds in $\mathbb{Z}^d$

3.2.1 Preliminaries

In order to apply the circle method, we write

$$|S_{\underline{x}}(\lambda,\mu)| = \sum_{\underline{y}\in\mathbb{Z}^d} \int_0^1 e(\alpha(|\underline{y}|^2 - \lambda^2))\,d\alpha \int_0^1 e(\beta(\underline{y}\cdot\underline{x} - \mu))\,d\beta.$$

We will break the integration in α into two pieces: the so-called major and minor arcs. We make the following definitions:

Definition 3.4. *Let $0 < \theta \leq 1$. The major arc $\mathfrak{M}_{a/q}(\theta)$ for $q > 1$ is given by*

$$\mathfrak{M}_{a/q}(\theta) := \{\alpha \in [0,1] \ : \ \left|\alpha - \frac{a}{q}\right| \leq q^{-1}\lambda^{-2+\theta}\}.$$

For $q = 1$, make the special definition

$$\mathfrak{M}_{0/1}(\theta) := \left[0, \lambda^{-2+\theta}\right] \cup \left[1 - \lambda^{-2+\theta}, 1\right].$$

Additionally we define the collection of major arcs $\mathfrak{M}(\theta)$ as

$$\mathfrak{M}(\theta) := \bigcup_{1 \leq q \leq \lambda^\theta} \bigcup_{\substack{0 \leq a < q, \\ (a,q)=1}} \mathfrak{M}_{a/q}(\theta).$$

The minor arcs $\mathfrak{m}(\theta)$ are the complement, i.e. $\mathfrak{m}(\theta) = [0,1] \setminus \mathfrak{M}(\theta)$.

Now write

$$|S_{\underline{x}}(\lambda, \mu)| = I_{\mathfrak{m}(\theta)} + I_{\mathfrak{M}(\theta)}, \tag{3.1}$$

with

$$I_{\mathfrak{m}(\theta)} = \sum_{\underline{y} \in \mathbb{Z}^d} \int_{\mathfrak{m}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta,$$

$$I_{\mathfrak{M}(\theta)} = \sum_{\underline{y} \in \mathbb{Z}^d} \int_{\mathfrak{M}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta.$$

We note that $I_{\mathfrak{m}(\theta)}$ and $I_{\mathfrak{M}(\theta)}$ depend on $\underline{x}$ though we do not notate the explicit dependency. For the minor arcs all estimates are in fact independent of $\underline{x}$, although analysis of $\underline{x}$ features heavily for estimates of the major arcs.

3.2.2 Minor Arcs Estimate

To estimate $I_{\mathfrak{m}(\theta)}$, note that the sum over $\mathbb{Z}^d$ may as well be replaced with a sum over the lattice cube Q_N where $N := \lfloor \lambda \rfloor$, as the only nonzero terms in the sum occur when $|\underline{y}| = \lambda$. Hence

$$I_{\mathfrak{m}(\theta)} = \sum_{\underline{y} \in Q_N} \int_{\mathfrak{m}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta$$

$$= \int_{\mathfrak{m}(\theta)} \int_0^1 e(-\alpha\lambda^2 - \beta\mu) \sum_{\underline{y} \in Q_N} e(\alpha|\underline{y}|^2 + \beta \underline{y} \cdot \underline{x})\, d\beta\, d\alpha \qquad (3.2)$$

Definition 3.5. *For the minor arcs, we define the Gaussian sum*

$$g_{\underline{x},\lambda}(\alpha, \beta) = \sum_{\underline{y} \in Q_N} e(\alpha|\underline{y}|^2 + \beta \underline{y} \cdot \underline{x}).$$

Lemma 3.6 (Weyl estimate)**.** *If* $\alpha \in \mathfrak{m}(\theta)$*, then for any* $\varepsilon > 0$*,*

$$|g_{\underline{x},\lambda}(\alpha, \beta)| \lesssim_\varepsilon \lambda^{d - \frac{d\theta}{2} + \varepsilon}.$$

Proof. Note that $g_{\underline{x},\lambda}(\alpha, \beta)$ splits as

$$g_{\underline{x},\lambda}(\alpha, \beta) = \prod_{i=1}^{d} \sum_{y_i=-N}^{N} e(\alpha y_i^2 + \beta y_i x_i).$$

A bound for $\sum_{y=-N}^{N} e(\alpha y^2 + \beta x y)$ comes from the Weyl inequality. A version dating to Vinogradov in [Vin27] says if

$$S = \sum_{y=M+1}^{M+P} e(f(y)),$$

with $f(y) = \alpha y^n + c_{n-1} y^{n-1} + \cdots + c_0$, and

$$\left| \alpha - \frac{a}{q} \right| < \frac{1}{q^2}, \qquad (a, q) = 1, \qquad 0 < q < P^n,$$

then

$$S = O\left(P^{1+\varepsilon}(q^{-1} + P^{-1} + P^{-n+1} + qP^{-n})^\sigma\right), \qquad \sigma = 2^{-n+1}.$$

In our case, letting $P = 2N + 1$ and $n = 2$ gives

$$S = O\left(N^{1+\varepsilon}(q^{-1} + N^{-1} + qN^{-2})^{1/2}\right). \qquad (3.3)$$

We now use a standard application of the pigeonhole principle. Note that for any $Q \in \mathbb{Z}^+$, some two of $0, \alpha \bmod 1, 2\alpha \bmod 1, \ldots, Q\alpha \bmod 1$ must lie within $\frac{1}{Q}$ of each other, so there exist integers $0 \leq n_1 < n_2 \leq Q$, and $a \geq 0$ such that $|n_2\alpha - n_1\alpha - a| < \frac{1}{Q}$. Setting $q := n_2 - n_1$ we have $1 \leq q \leq Q$

and rearranging yields

$$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{qQ} \leq \frac{1}{q^2}, \tag{3.4}$$

so for any Q we get a q with which the Weyl inequality applies. Now looking at the left inequality of the above and setting $Q = \lceil \lambda^{2-\theta} \rceil$ gives a q with the property that

$$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{q\lambda^{2-\theta}}.$$

However as α is in the minor arcs $\mathfrak{m}(\theta)$, this forces this particular q to be bigger than λ^{θ}. Since $\lambda^{\theta} < q \leq \lceil \lambda^{2-\theta} \rceil$, the Weyl estimate with this q boils down to

$$|S| \lesssim_{\varepsilon} \lambda^{1-\theta/2+\varepsilon}.$$

Applying this estimate to each term of the d-fold product form of $g_{\underline{x},\lambda}(\alpha, \beta)$, yields the desired

$$|g_{\underline{x},\lambda}(\alpha, \beta)| \lesssim_{\varepsilon} \lambda^{d-\frac{d\theta}{2}+\varepsilon},$$

for any $\alpha \in \mathfrak{m}(\theta)$. □

Corollary 3.7. *For any* $\varepsilon > 0$

$$|I_{\mathfrak{m}(\theta)}| \lesssim_{\varepsilon} \lambda^{d-\frac{d\theta}{2}+\varepsilon}.$$

Proof. Applying the Weyl estimate in equation (3.2), we find for any $\varepsilon > 0$

$$\begin{aligned} |I_{\mathfrak{m}(\theta)}| &\leq \int_{\mathfrak{m}(\theta)} \int_0^1 |g_{\underline{x},\lambda}(\alpha, \beta)| \, d\beta \, d\alpha \\ &\lesssim_{\varepsilon} \lambda^{d-\frac{d\theta}{2}+\varepsilon} \end{aligned}$$

□

3.2.3 Major Arcs Estimate

To ensure convergence in the minor arcs, we restricted the sum to a lattice cube. To ensure convergence in the major arcs, we instead introduce a smooth cutoff function φ which will allow us to make sense of $\sum_{\underline{y}\in\mathbb{Z}^d} f(\underline{y})\varphi\left(\frac{\underline{y}}{\lambda}\right)$ as all but finitely many terms are 0.

Definition 3.8. *Let* $\psi : \mathbb{R} \to \mathbb{R}$ *be a smooth function bounded by* 1 *and satisfying*

$$\psi(y) = \begin{cases} 1 & \text{if } |y| \leq 1, \\ 0 & \text{if } |y| \geq 2. \end{cases}$$

For $\underline{y} = (y_1, \ldots, y_d)$*, then we set*

$$\varphi(\underline{y}) := \psi(y_1) \ldots \psi(y_d).$$

Now we may write

$$\begin{aligned} I_{\mathfrak{M}(\theta)} &= \sum_{\underline{y} \in \mathbb{Z}^d} \int_{\mathfrak{M}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta \\ &= \sum_{\underline{y} \in \mathbb{Z}^d} \varphi\left(\frac{\underline{y}}{\lambda}\right) \int_{\mathfrak{M}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta \\ &= \int_{\mathfrak{M}(\theta)} \int_0^1 e(-\alpha\lambda^2 - \beta\mu) \sum_{\underline{y} \in \mathbb{Z}^d} e(\alpha|\underline{y}|^2 + \beta\underline{y} \cdot \underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right) d\beta\, d\alpha. \end{aligned}$$

Definition 3.9. *For the major arcs, we define the Gaussian sum*

$$G_{\underline{x},\lambda}(\alpha, \beta) = \sum_{\underline{y} \in \mathbb{Z}^d} e(\alpha|\underline{y}|^2 + \beta\underline{y} \cdot \underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right).$$

Pulling in the absolute value, we have

$$|I_{\mathfrak{M}(\theta)}| \leq \int_{\mathfrak{M}(\theta)} \int_0^1 |G_{\underline{x},\lambda}(\alpha, \beta)|\, d\beta\, d\alpha \tag{3.5}$$

To simplify notation, we combine the cut-off function with $e(\alpha|\underline{y}|^2)$.

Definition 3.10. *We define*

$$\varphi_{\alpha,\lambda}(\underline{y}) := e(\alpha|\underline{y}|^2)\varphi\left(\frac{\underline{y}}{\lambda}\right)$$

With this notation we have

$$G_{\underline{x},\lambda}(\alpha, \beta) = \sum_{\underline{y} \in \mathbb{Z}^d} e(\beta\underline{y} \cdot \underline{x})\varphi_{\alpha,\lambda}(\underline{y}).$$

Additionally, recall the notation set up in equation (1.5):

$$\widetilde{f}(\underline{\xi}) := \int_{\mathbb{R}^d} f(\underline{y}) e(-\underline{y} \cdot \underline{\xi})\, d\underline{y}$$

Lemma 3.11. *If* $\alpha = \frac{a}{q} + \tau$ *with* $0 \leq a < q$ *and* $(a, q) = 1$*, then*

$$G_{\underline{x},\lambda}(\alpha, \beta) = q^{-d} \sum_{\underline{m} \in \mathbb{Z}^d} \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\underline{m} - q\beta\underline{x}}{q} \right) \sum_{\underline{s} \pmod{q}} e\left(\frac{a}{q} |\underline{s}|^2 + \frac{\underline{s} \cdot \underline{m}}{q} \right).$$

Proof. The proof is cleaner in the language of lattices, of which we will make a note of at the end, but instead we shall go through a more pedestrian path of elementary Fourier analysis manipulations. The key player will be the Poisson summation formula equation (1.6).

First we rewrite $G_{\underline{x},\lambda}(\alpha, \beta)$ as $G_{\underline{x},\lambda}\left(\frac{a}{q} + \tau, \beta\right)$:

$$\begin{aligned} G_{\underline{x},\lambda}(\alpha, \beta) &= \sum_{\underline{y} \in \mathbb{Z}^d} e(\beta \underline{y} \cdot \underline{x}) \varphi_{\frac{a}{q}+\tau,\lambda}(\underline{y}) \\ &= \sum_{\underline{y} \in \mathbb{Z}^d} e(\beta \underline{y} \cdot \underline{x}) e\left(\frac{a}{q} |\underline{y}|^2 \right) \varphi_{\tau,\lambda}(\underline{y}). \end{aligned}$$

Writing $\underline{y}$ as $q\underline{y}' + \underline{s}$ where $\underline{s}$ is a residue mod q gives

$$\begin{aligned} e\left(\frac{a}{q} |\underline{y}|^2 \right) &= e\left(\frac{a}{q} |q\underline{y}' + \underline{s}|^2 \right) \\ &= e\left(aq|\underline{y}'|^2 + 2a\underline{y}' \cdot \underline{s} + \frac{a}{q} |\underline{s}|^2 \right) \\ &= e\left(\frac{a}{q} |\underline{s}|^2 \right), \end{aligned}$$

and so

$$\begin{aligned} G_{\underline{x},\lambda}(\alpha, \beta) &= \sum_{\underline{y} \in \mathbb{Z}^d} e(\beta \underline{y} \cdot \underline{x}) e\left(\frac{a}{q} |\underline{y}|^2 \right) \varphi_{\tau,\lambda}(\underline{y}) \\ &= \sum_{\underline{s}(q)} \sum_{\underline{y} \in \mathbb{Z}^d} e(\beta (q\underline{y} + \underline{s}) \cdot \underline{x}) e\left(\frac{a}{q} |\underline{s}|^2 \right) \varphi_{\tau,\lambda}(q\underline{y} + \underline{s}) \\ &= \sum_{\underline{s}(q)} e\left(\frac{a}{q} |\underline{s}|^2 + \beta \underline{s} \cdot \underline{x} \right) \sum_{\underline{y} \in \mathbb{Z}^d} e(q\beta \underline{y} \cdot \underline{x}) \varphi_{\tau,\lambda}(q\underline{y} + \underline{s}) \end{aligned}$$

$$= \sum_{\underline{s}(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{s}\cdot\underline{x}\right) \sum_{\underline{y}\in\mathbb{Z}^d} e(q\beta\underline{y}\cdot\underline{x})\varphi_{\tau,\lambda,q,\underline{s}}(\underline{y}), \quad (3.6)$$

where

$$\varphi_{\tau,\lambda,q,\underline{s}}(\underline{y}) := \varphi_{\tau,\lambda}(q\underline{y}+\underline{s}).$$

We can recognize the inner sum in equation (3.6) as $\widehat{\varphi}_{\tau,\lambda,q,\underline{s}}(-q\beta\underline{x})$, i.e.

$$\begin{aligned} G_{\underline{x},\lambda}(\alpha,\beta) &= \sum_{\underline{s}(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{s}\cdot\underline{x}\right) \sum_{\underline{y}\in\mathbb{Z}^d} e(q\beta\underline{y}\cdot\underline{x})\varphi_{\tau,\lambda,q,\underline{s}}(\underline{y}) \\ &= \sum_{\underline{s}(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{s}\cdot\underline{x}\right) \widehat{\varphi}_{\tau,\lambda,q,\underline{s}}(-q\beta\underline{x}). \end{aligned} \quad (3.7)$$

Using the Poisson summation as stated in equation (1.6), we have

$$\widehat{\varphi}_{\tau,\lambda,q,\underline{s}}(-q\beta\underline{x}) = \sum_{\underline{m}\in\mathbb{Z}^d} \widetilde{\varphi}_{\tau,\lambda,q,\underline{s}}(\underline{m}-q\beta\underline{x}). \quad (3.8)$$

Note

$$\begin{aligned} \widetilde{\varphi}_{\tau,\lambda,q,\underline{s}}(\underline{\xi}) &= \int_{\mathbb{R}^d} \varphi_{\tau,\lambda,q,\underline{s}}(\underline{y}) e(-\underline{y}\cdot\underline{\xi})\,d\underline{y} \\ &= \int_{\mathbb{R}^d} \varphi_{\tau,\lambda}(q\underline{y}+\underline{s}) e(-\underline{y}\cdot\underline{\xi})\,d\underline{y} \\ &= q^{-d}\int_{\mathbb{R}^d} \varphi_{\tau,\lambda}(\underline{y}+\underline{s}) e\left(-\frac{\underline{y}\cdot\underline{\xi}}{q}\right)\,d\underline{y} \\ &= q^{-d}\int_{\mathbb{R}^d} \varphi_{\tau,\lambda}(\underline{y}) e\left(-\frac{\underline{y}\cdot\underline{\xi}}{q} + \frac{\underline{s}\cdot\underline{\xi}}{q}\right)\,d\underline{y} \\ &= q^{-d} e\left(\frac{\underline{s}\cdot\underline{\xi}}{q}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{\xi}}{q}\right), \end{aligned} \quad (3.9)$$

and thus via equation (3.8)

$$\begin{aligned} \widehat{\varphi}_{\tau,\lambda,q,\underline{s}}(-q\beta\underline{x}) &= \sum_{\underline{m}\in\mathbb{Z}^d} \widetilde{\varphi}_{\tau,\lambda,q,\underline{s}}(\underline{m}-q\beta\underline{x}) \\ &= \sum_{\underline{m}\in\mathbb{Z}^d} q^{-d} e\left(\frac{\underline{s}\cdot(\underline{m}-q\beta\underline{x})}{q}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m}-q\beta\underline{x}}{q}\right) \\ &= \sum_{\underline{m}\in\mathbb{Z}^d} q^{-d} e\left(\frac{\underline{s}\cdot\underline{m}}{q} - \beta\underline{s}\cdot\underline{x}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m}-q\beta\underline{x}}{q}\right) \end{aligned}$$

$$= q^{-d} e(-\beta \underline{s} \cdot \underline{x}) \sum_{\underline{m} \in \mathbb{Z}^d} e\left(\frac{\underline{s} \cdot \underline{m}}{q}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - q\beta \underline{x}}{q}\right).$$

Plugging this into equation (3.7) and reordering the summation, we obtain

$$\begin{aligned} G_{\underline{x},\lambda}(\alpha, \beta) &= \sum_{\underline{s}(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta \underline{s} \cdot \underline{x}\right) \widehat{\varphi}_{\tau,\lambda,q,\underline{s}}(-q\beta \underline{x}) \\ &= q^{-d} \sum_{\underline{s}(q)} e\left(\frac{a}{q}|\underline{s}|^2\right) \sum_{\underline{m} \in \mathbb{Z}^d} e\left(\frac{\underline{s} \cdot \underline{m}}{q}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - q\beta \underline{x}}{q}\right) \\ &= q^{-d} \sum_{\underline{m} \in \mathbb{Z}^d} \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - q\beta \underline{x}}{q}\right) \sum_{\underline{s}(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \frac{\underline{s} \cdot \underline{m}}{q}\right). \end{aligned}$$

This completes the proof, though we make the additional note that if we let $L = q\mathbb{Z}^d$ and $f(\underline{y}) = e(\beta \underline{y} \cdot \underline{x}) \varphi_{\tau,\lambda}(\underline{y})$, then the identity we have shown can be quickly computed from a Poisson summation formula for lattices:

$$\sum_{\underline{y} \in L} f(\underline{y}) = \frac{1}{\det(L)} \sum_{\underline{\xi} \in L^*} \widetilde{f}(\underline{\xi}),$$

where L^* is the dual lattice.

□

Lemma 3.12. *If* $\alpha = \frac{a}{q} + \tau$ *with* $0 \leq a < q$ *and* $(a, q) = 1$*, then*

$$|G_{\underline{x},\lambda}(\alpha, \beta)| \leq 2^{d/2} q^{-d/2} \sum_{\underline{m} \in \mathbb{Z}^d} \left| \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - q\beta \underline{x}}{q}\right) \right|.$$

Proof. This follows directly from by pulling in absolute values in the expression for $G_{\underline{x},\lambda}(\alpha, \beta)$ given in Lemma 3.11, and recalling the standard Gaussian estimate

$$\left| \sum_{s=0}^{q-1} e\left(\frac{a}{q}s^2 + \frac{sm}{q}\right) \right| \leq \sqrt{2q}. \tag{3.10}$$

Note that this is also true in the case $a = m = 0$ because we have insisted $(a, q) = 1$ which means that $q = 1$ and the sum only has a single term. □

In light of Lemma 3.12, we turn to estimating $\left|\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi})\right|$ when α is in a major arc. It will turn out that terms where $|\underline{\xi}|$ is small will far outweigh the other terms. We introduce notation to capture when this happens in our case where $\underline{\xi} = (\underline{m} - q\beta \underline{x})/q$.

Definition 3.13. *For a point $\underline{y} \in \mathbb{R}^d$, let $[\underline{y}]$ denote the closest point in $\mathbb{Z}^d$ to $\underline{y}$, i.e.*

$$[\underline{y}] := \underline{m} \in \mathbb{Z}^d \text{ such that } \underline{y} - \underline{m} \in (-1/2, 1/2]^d.$$

One might be concerned in the above definition that one may have multiple reasonable choices of closest lattice points (though note with the way we have set it up, the choice is unique). However, it will turn out with the way we will apply $[\underline{y}]$ that those points which have multiple closest lattice points will contribute least to the estimates, i.e. will we only really care when $\underline{y} - [\underline{y}]$ is small. We make an additional definition in analogy to the fractional part of an integer notation, though below we allow the fractional part to have negative coordinates.

Definition 3.14. *For a point $\underline{y} \in \mathbb{R}^d$, let $\{\underline{y}\}$ denote the fractional part of the point, i.e.*

$$\{\underline{y}\} := \underline{y} - [\underline{y}].$$

With this notation, we have

$$\sum_{\underline{m} \in \mathbb{Z}^d} \left| \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\underline{m} - q\beta\underline{x}}{q} \right) \right| = M + \mathcal{E}, \tag{3.11}$$

where the main term M is

$$M := \left| \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\{-q\beta\underline{x}\}}{q} \right) \right|$$

and the remainder

$$\mathcal{E} := \sum_{\underline{m} \neq 0} \left| \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\{-q\beta\underline{x}\} + \underline{m}}{q} \right) \right|$$

will be an error term.

We will use oscillatory integral estimates to get bounds on M and $\mathcal{E}$.

Lemma 3.15. *Let $\underline{\xi} \in \mathbb{R}^d$ such that $\max_{1 \leq i \leq d} |\xi_i| \geq \frac{1}{2q}$ where $q < \lambda^\theta$. If $\lambda^{1-\theta} \geq 16$ and $0 < |\tau| \leq q^{-1}\lambda^{-2+\theta}$, then for any $K \geq 0$ there exists a constant C_K such that*

$$|\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi})| \lesssim_d C_K \lambda^d \min \left(1, |\tau|^{-1}\lambda^{-2}(\lambda \max_{1 \leq i \leq d} |\xi_i|)^{-K} \right).$$

Proof. We shall adapt the principle of nonstationary phase as outlined in [Ste93] or [Tao20]. Recall that $\varphi(\underline{y}) := \psi(y_1) \ldots \psi(y_d)$ and expand

$$\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi}) = \int_{\mathbb{R}^d} \varphi_{\tau,\lambda}(\underline{y}) e(-\underline{y} \cdot \underline{\xi}) \, d\underline{y}$$

$$= \int_{\mathbb{R}^d} e(\tau|\underline{y}|^2 - \underline{y} \cdot \underline{\xi})\varphi\left(\frac{\underline{y}}{\lambda}\right)\, d\underline{y}$$

$$= \prod_{i=1}^{d} \int_{\mathbb{R}} e(\tau y_i^2 - \xi_i y_i)\psi\left(\frac{y_i}{\lambda}\right)\, dy_i, \tag{3.12}$$

Let

$$I(\lambda,\tau,\xi) := \int_{\mathbb{R}} e(\tau y^2 - \xi y)\psi\left(\frac{y}{\lambda}\right)\, dy.$$

Since $\psi(y)$ is 1-bounded and supported on $[-2,2]$, we have the trivial estimate

$$|I(\lambda,\tau,\xi)| \leq 4\lambda.$$

As $\tau \neq 0$, we may substitute $y = (\xi/\tau)x$ to get

$$I(\lambda,\tau,\xi) = \frac{\xi}{\tau} \int_{\mathbb{R}} e\left(\frac{\xi^2}{\tau}(x^2 - x)\right)\psi\left(\frac{\xi x}{\lambda\tau}\right)\, dx.$$

Writing $\phi(x) := 2\pi(x^2 - x)$ and

$$J(\eta,\gamma) := \int_{\mathbb{R}} e^{i\eta\phi(x)}\psi(\gamma x)\, dx,$$

we have

$$I(\lambda,\tau,\xi) = \frac{\xi}{\tau} J\left(\frac{\xi^2}{\tau}, \frac{\xi}{\lambda\tau}\right). \tag{3.13}$$

Because ψ is a smooth cut-off function, we will be able to use integration by parts to obtain estimates on $J(\eta,\gamma)$ with constants depending on derivatives of ψ. We will also need $|\phi'(x)| \geq c > 0$ on the support of $\psi(\gamma x)$. As

$$\mathrm{supp}(\psi(\gamma x)) = \{x \,:\, |x| < 2|\gamma|^{-1}\},$$

on $\mathrm{supp}(\psi(\gamma x))$ we have

$$|\phi'(x)| = 2\pi|2x - 1| \geq 2\pi(1 - 2|x|) \geq 2\pi\left(1 - 4|\gamma|^{-1}\right),$$

so if we assume say $|\gamma| \geq 8$ we have $|\phi'(x)| \geq \pi$ on the support of $\psi(\gamma x)$.

We perform the first integration by parts explicitly to give the idea:

$$J(\eta,\gamma) = \int_{\mathbb{R}} e^{i\eta\phi(x)}\psi(\gamma x)\, dx$$

$$
\begin{aligned}
&= \int_{\mathbb{R}} e^{i\eta\phi(x)}\phi'(x)\frac{\psi(\gamma x)}{\phi'(x)}\,dx \\
&= -\frac{1}{i\eta}\int_{\mathbb{R}} e^{i\eta\phi(x)}\frac{d}{dx}\left[\frac{\psi(\gamma x)}{\phi'(x)}\right]\,dx \\
&= -\frac{1}{i\eta}\int_{\mathbb{R}} e^{i\eta\phi(x)}\left[\gamma\frac{\psi'(\gamma x)}{\phi'(x)} - \frac{\phi''(x)\psi(\gamma x)}{(\phi'(x))^2}\right]\,dx.
\end{aligned}
$$

As $\phi''(x) = 4\pi$, we rewrite

$$
J(\eta,\gamma) = -\frac{\gamma}{i\eta}\int_{\mathbb{R}} e^{i\eta\phi(x)}\frac{\psi'(\gamma x)}{\phi'(x)}\,dx + \frac{4\pi}{i\eta}\int_{\mathbb{R}} e^{i\eta\phi(x)}\frac{\psi(\gamma x)}{(\phi'(x))^2}\,dx.
$$

In general, we note

$$
\frac{d}{dx}\left[\frac{f(\gamma x)}{(\phi'(x))^n}\right] = \gamma\frac{f'(\gamma x)}{(\phi'(x))^n} - 4\pi n\frac{f(\gamma x)}{(\phi'(x))^{n+1}},
$$

so by repeating the integration by parts trick K times, we obtain

$$
|J(\eta,\gamma)| \lesssim C_K|\eta|^{-K}\left(1+|\gamma|+\cdots+|\gamma|^K\right),
$$

where C_K depends on the bounds of up to the K-th derivative of ψ. As we assumed $|\gamma| \geq 8$ to ensure a lower bound of π for $|\phi'(x)|$, the above estimate simplifies to

$$
|J(\eta,\gamma)| \lesssim C_K|\eta|^{-K}|\gamma|^K. \tag{3.14}
$$

In our case, as per equation (3.13), we apply this estimate setting

$$
\eta = \frac{\xi^2}{\tau}, \quad \text{and} \quad \gamma = \frac{\xi}{\lambda\tau},
$$

with the assumption $|\xi| \geq \frac{1}{2q}$.

Recalling $|\tau| \leq q^{-1}\lambda^{-2+\theta}$, we note

$$
|\gamma| = \left|\frac{\xi}{\lambda\tau}\right| \geq \frac{1}{2}\lambda^{1-\theta} \geq 8,
$$

as was assumed for equation (3.14), and was why we had the mild condition $\lambda^{1-\theta} \geq 16$ in the initial assumptions.

Applying equation (3.13) and equation (3.14) we find when $|\xi| \geq \frac{1}{2q}$

$$\begin{aligned}|I(\lambda,\tau,\xi)| &= \left|\frac{\xi}{\tau} J\left(\frac{\xi^2}{\tau}, \frac{\xi}{\lambda\tau}\right)\right| \\ &\lesssim C_K \frac{|\xi|}{|\tau|}\frac{|\tau|^K}{|\xi^2|^K}\frac{|\xi|^K}{|\lambda\tau|^K} \\ &\lesssim C_K |\tau|^{-1}\lambda^{-K}|\xi|^{-K+1}.\end{aligned}$$

Combining this with the trivial estimate, we have

$$\begin{aligned}|I(\lambda,\tau,\xi)| &\lesssim C_K \min(4\lambda, |\tau|^{-1}\lambda^{-K}|\xi|^{-K+1}) \\ &\lesssim C_K\lambda \min(1, |\tau|^{-1}\lambda^{-K-1}|\xi|^{-K+1}).\end{aligned}$$

Reindexing in K, we get for any $K \geq 0$

$$|I(\lambda,\tau,\xi)| \lesssim C_K\lambda \min\left(1, |\tau|^{-1}\lambda^{-2}(\lambda|\xi|)^{-K}\right), \tag{3.15}$$

under the assumption $|\xi| \geq \frac{1}{2q}$.[23]

[23] Now of course C_K depends on up to the $(K+1)$-th derivative of $\psi(y)$.

Finally we can return to equation (3.12). The above estimate holds for any coordinate of $\underline{\xi}$ satisfying $|\xi_i| \geq \frac{1}{2q}$, but in particular it holds for the coordinate of largest magnitude by assumption. For all of the other coordinates, we use the trivial estimate $|I(\lambda,\tau,\xi_i)| \leq 4\lambda$ to obtain the desired

$$\begin{aligned}|\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi})| &\leq \prod_{i=1}^{d}\left|\int_{\mathbb{R}} e(\tau y_i^2 - \xi_i y_i)\psi\left(\frac{y_i}{\lambda}\right)\,dy_i\right| \\ &\lesssim_d C_K\lambda^d \min\left(1, |\tau|^{-1}\lambda^{-2}(\lambda \max_{1\leq i\leq d}|\xi_i|)^{-K}\right).\end{aligned}$$

□

Lemma 3.16. *If* $\lambda^{1-\theta} \geq 16$ *then for any integer* $K > d$ *there exists a* C_K *such that*

$$\int_{\mathfrak{M}_{a/q}(\theta)}\int_0^1 \mathcal{E}\,d\beta\,d\alpha \lesssim_d C_K q^K \lambda^{d-K-2}\log(\lambda).$$

Proof. Set

$$\underline{\xi} = \frac{\{-q\beta\underline{x}\} + \underline{m}}{q}$$

where $\underline{m} \neq 0$. As $\{-q\beta\underline{x}\} \in Q_{\frac{1}{2}}$ we have $\max\limits_{1\leq i\leq d} |\xi_i| \geq \frac{1}{2q}$ so an application of Lemma 3.15 yields

$$\begin{aligned}\mathcal{E} &= \sum_{\underline{m}\neq 0} \left| \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\{-q\beta\underline{x}\} + \underline{m}}{q}\right)\right| \\ &\lesssim_d C_K\lambda^d \sum_{\underline{m}\neq 0} \min\left(1, |\tau|^{-1}\lambda^{-2}\left(\frac{\lambda}{q}\max_{1\leq i\leq d} |\{-q\beta x_i\} + m_i|\right)^{-K}\right).\end{aligned}$$

Noting that $|\{-q\beta x_i\} + m_i| \geq \frac{|m_i|}{2}$, we can give an estimate independent of β while changing C_K by a factor of 2^K:

$$\mathcal{E} \lesssim_d C_K\lambda^d \sum_{\underline{m}\neq 0} \min\left(1, |\tau|^{-1}\lambda^{-2}\left(\frac{\lambda}{q}\max_{1\leq i\leq d} |m_i|\right)^{-K}\right).$$

We observe

$$1 \leq |\tau|^{-1}\lambda^{-2}\left(\frac{\lambda}{q}\max_{1\leq i\leq d} |m_i|\right)^{-K} \iff \max_{1\leq i\leq d} |m_i| \leq q|\tau|^{-1/K}\lambda^{-1-2/K}.$$

In particular as $\underline{m} \neq 0$ we know $\max\limits_{1\leq i\leq d} |m_i| \geq 1$ and so if

$$|\tau| > q^K\lambda^{-K-2},$$

then we do not need to use the trivial estimate at all and we have

$$\begin{aligned}\mathcal{E} &\lesssim_d C_K q^K \lambda^{d-K-2}|\tau|^{-1} \sum_{\underline{m}\neq 0}\left(\max_{1\leq i\leq d} |m_i|\right)^{-K} \\ &\lesssim_d C_K q^K \lambda^{d-K-2}|\tau|^{-1}, \qquad (3.16)\end{aligned}$$

where the last line follows because the sum is a dimensional constant since it was assumed $K > d$.

We now examine what happens to $\mathcal{E}$ for τ in the remaining range

$$0 < |\tau| \leq q^K\lambda^{-K-2}.$$

For notational convenience, let N be the cut-off, i.e.

$$N := \left\lfloor q|\tau|^{-1/K}\lambda^{-1-2/K}\right\rfloor.$$

We have

$$\#\{\underline{m} \in \mathbb{Z}^d \ : \ \max_{1 \le i \le d} |m_i| \le N\} \lesssim_d N^d.$$

Therefore for $|\tau|$ in this range, we have

$$\mathcal{E} \lesssim_d C_K \lambda^d N^d + C_K q^K \lambda^{d-K-2} |\tau|^{-1} \sum_{\underline{m} \notin Q_N} \left(\max_{1 \le i \le d} |m_i| \right)^{-K}.$$

Here we make the note[24]

[24] We use the observation that a hollow cube in d dimensions is the union of $2d$ cubes of dimension $d-1$.

$$\begin{aligned}
\sum_{\underline{m} \notin Q_N} \left(\max_{1 \le i \le d} |m_i| \right)^{-K} &= \sum_{n=N+1}^{\infty} n^{-K} \#\{\underline{m} \in \mathbb{Z}^d \ : \ \max_{1 \le i \le d} |m_i| = n\} \\
&\le \sum_{n=N+1}^{\infty} n^{-K} 2d(2n+1)^{d-1} \\
&\lesssim_d \sum_{n=N+1}^{\infty} n^{d-1-K} \\
&\lesssim_d \int_N^{\infty} x^{d-1-K} \, dx \\
&\lesssim_d N^{d-K}.
\end{aligned}$$

Hence for $0 < |\tau| \le q^K \lambda^{-K-2}$, we have

$$\begin{aligned}
\mathcal{E} &\lesssim_d C_K \lambda^d N^d + C_K q^K \lambda^{d-K-2} |\tau|^{-1} N^{d-K} \\
&\lesssim_d C_K q^d \lambda^{-2d/K} |\tau|^{-d/K} + C_K q^d \lambda^{-2d/K} |\tau|^{-d/K} \\
&\lesssim_d C_K q^d \lambda^{-2d/K} |\tau|^{-d/K}, \qquad (3.17)
\end{aligned}$$

where plugging in the definition of N shows the terms balance.

Finally we put equation (3.16) and equation (3.17) into the integral, first noting that as we have removed the dependence on β, we can drop the integration in β.

$$\int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 \mathcal{E} \, d\beta \, d\alpha \lesssim_d \mathcal{E}_1 + \mathcal{E}_2,$$

where

$$\mathcal{E}_1 = \int\limits_0^{q^K\lambda^{-K-2}} C_K q^d \lambda^{-2d/K} \tau^{-d/K}\, d\tau$$
$$\lesssim_d C_K q^d \lambda^{-2d/K} \left(q^K \lambda^{-K-2}\right)^{1-d/K}$$
$$\lesssim_d C_K q^K \lambda^{d-K-2},$$

and

$$\mathcal{E}_2 = \int\limits_{q^K\lambda^{-K-2}}^{q^{-1}\lambda^{-2+\theta}} C_K q^K \lambda^{d-K-2} \tau^{-1}\, d\tau$$
$$= C_K q^K \lambda^{d-K-2} \log\left(\frac{q^{-1}\lambda^{-2+\theta}}{q^K\lambda^{-K-2}}\right)$$
$$\leq C_K q^K \lambda^{d-K-2} \log(\lambda).$$

Putting these together yields the desired

$$\int\limits_{\mathfrak{M}_{a/q}(\theta)} \int\limits_0^1 \mathcal{E}\, d\beta\, d\alpha \lesssim_d C_K q^K \lambda^{d-K-2} \log(\lambda).$$

$\square$

We now turn to the oscillatory integral estimate to be used with the main term.

Lemma 3.17. *For any* $\tau \in \mathbb{R}$

$$|\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi})| \lesssim \lambda^d \left(1 + \lambda^2|\tau| + \lambda|\underline{\xi}|\right)^{-d/2}.$$

Proof. Expanding achieves

$$\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi}) = \int\limits_{\mathbb{R}^d} \varphi_{\tau,\lambda}(\underline{y}) e(-\underline{y}\cdot\underline{\xi})\, d\underline{y}$$
$$= \int\limits_{\mathbb{R}^d} e(\tau|\underline{y}|^2 - \underline{y}\cdot\underline{\xi})\varphi\left(\frac{\underline{y}}{\lambda}\right)\, d\underline{y}$$
$$= \lambda^d \int\limits_{\mathbb{R}^d} e(\lambda^2\tau|\underline{y}|^2 - \lambda\underline{y}\cdot\underline{\xi})\varphi(\underline{y})\, d\underline{y},$$

and we know

$$\int_{\mathbb{R}^d} e(\lambda^2\tau|\underline{y}|^2 - \lambda\underline{y}\cdot\underline{\xi})\varphi(\underline{y})\,d\underline{y} = O((\lambda^2|\tau| + \lambda|\underline{\xi}|)^{-d/2})$$

which is estimate 5.13(a) in chapter VIII of Stein's book Harmonic Analysis [Ste93]. Accounting for the trivial estimate, we have

$$|\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi})| \lesssim \lambda^d \left(1 + \lambda^2|\tau| + \lambda|\underline{\xi}|\right)^{-d/2},$$

as desired. □

We now apply Lemma 3.17 to estimate M in equation (3.11).

Corollary 3.18. *If M is the main term appearing in equation* (3.11) *then*

$$M \lesssim \lambda^d \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\,|\{q\beta\underline{x}\}|\right)^{-d/2}.$$

Proof. We have

$$M = \left|\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\{-q\beta\underline{x}\}}{q}\right)\right|,$$

and an application of Lemma 3.17 gives

$$M \lesssim \lambda^d \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\,|\{-q\beta\underline{x}\}|\right)^{-d/2},$$

and we observe $|\{-q\beta\underline{x}\}| = |\{q\beta\underline{x}\}|$. □

In view of Corollary 3.18, we now want to study $|\{q\beta\underline{x}\}|$, so we will give it a name.

Definition 3.19. *If $\underline{y} \in \mathbb{R}$, let its distance to the nearest lattice point be denoted*

$$\|\underline{y}\| := |\{\underline{y}\}|\,.$$

Rewriting Corollary 3.18 with this notation, we have

$$M \lesssim \lambda^d \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\,\|q\beta\underline{x}\|\right)^{-d/2} \tag{3.18}$$

Lemma 3.20. *Fix $\underline{x} \in \mathbb{Z}^d$. If $\underline{m} \in \mathbb{Z}^d$ and* $\mathrm{Dist}\,(\underline{m}, \langle\underline{x}\rangle) < |\underline{x}|^{-1}$, *then $\underline{m} \in \langle\underline{x}\rangle$.*

Proof. Consider the parallelogram P generated by $\underline{m}$ and $\underline{x}$, that is the parallelogram which has vertices 0, $\underline{m}$, $\underline{x}$, and $\underline{m}+\underline{x}$. Then

$$\mathrm{Area}(P) = \mathrm{Dist}\left(\underline{m}, \langle\underline{x}\rangle\right)|\underline{x}| < 1.$$

This is only possible if in fact $\mathrm{Area}(P) = 0$, as any nondegenerate parallelogram with integral points has area at least 1.[25] □

[25] Alternatively one could note $\mathrm{Dist}\left(\underline{m}, \langle\underline{x}\rangle\right) = \min_{t\in\mathbb{R}} |\underline{m} - t\underline{x}|$ and then use calculus and the Cauchy-Schwarz inequality.

Lemma 3.21. *Fix $\underline{x} \in \mathbb{Z}^d$. If $\underline{m}_1, \underline{m}_2 \in \mathbb{Z}^d$ and*

$$\mathrm{Dist}\left(\underline{m}_i, \langle\underline{x}\rangle\right) < \frac{1}{\sqrt{6\pi|\underline{x}|}} \qquad \textit{for} \qquad i = 1, 2,$$

then $\underline{m}_1$, $\underline{m}_2$, and $\underline{x}$ are linearly dependent.

Proof. Let $\langle\underline{x}\rangle^\perp$ denote the subspace of $\mathbb{R}^d$ orthogonal to $\underline{x}$. In this subspace we define a disk around 0 which is also in $\langle\underline{m}_1, \underline{m}_2\rangle$

$$D := \left\{ \underline{y} \in \langle\underline{x}\rangle^\perp \cap \langle\underline{m}_1, \underline{m}_2\rangle \ : \ |\underline{y}| \leq \frac{1}{\sqrt{6\pi|\underline{x}|}} \right\}.$$

Finally we define the cylinder $C := D \times \underline{x}$.

Without loss of generality we may assume $\underline{m}_1, \underline{m}_2 \in C$, for if they are not we may add a multiples of $\underline{x}$ without changing $\mathrm{Dist}\left(\underline{m}_i, \langle\underline{x}\rangle\right)$ and of course this does not affect the linear independence of $\underline{m}_1$, $\underline{m}_2$, and $\underline{x}$.

Now consider the tetrahedron T generated by $\underline{m}_1$, $\underline{m}_2$, and $\underline{x}$, that is the tetrahedron which has vertices $0, \underline{m}_1, \underline{m}_2$, and $\underline{x}$. This tetrahedron is completely contained in C because of convexity, so we may compute

$$\mathrm{Vol}(T) < \mathrm{Vol}(C) = \pi\left(\frac{1}{\sqrt{6\pi|\underline{x}|}}\right)^2 |\underline{x}| = \frac{1}{6}.$$

This is only possible if in fact $\mathrm{Vol}(T) = 0$, as any nondegenerate tetrahedron with integral points has volume at least $1/6$.[26] □

[26] The constants in this argument could easily be improved, but the important thing is the power of $|\underline{x}|$.

Lemma 3.22. *If $\underline{x} \in \mathbb{Z}^d$ is primitive and $|\underline{x}| \leq \lambda$ then*

$$\mathrm{meas}\left\{0 \leq \beta \leq 1 \ : \ \|q\beta\underline{x}\| < \lambda^{-1}\right\} \leq 2\lambda^{-1}|\underline{x}|^{-1}.$$

Proof. Let $\underline{m} := [q\beta\underline{x}]$ be the closest lattice point to $q\beta\underline{x}$. If $\|q\beta\underline{x}\| < \lambda^{-1}$ then Lemma 3.20 guarantees $\underline{m} \in \langle\underline{x}\rangle$. As $\underline{x}$ is primitive, this means $\underline{m} = a\underline{x}$

for some $a \in \mathbb{Z}$ with $0 \leq a \leq q$. Then we have

$$|a\underline{x} - q\beta\underline{x}| \leq \lambda^{-1} \iff \left|\frac{a}{q} - \beta\right| \leq q^{-1}\lambda^{-1}|\underline{x}|^{-1}$$

for some integer $0 \leq a \leq q$.

Summing over the a's and noting that the $a = 0$ and $a = q$ terms both correspond to half-intervals[27] gives

$$\text{meas}\left\{0 \leq \beta \leq 1 \ : \ \|q\beta\underline{x}\| \leq \lambda^{-1}\right\} \leq 2\lambda^{-1}|\underline{x}|^{-1}.$$

[27] This computation could alternatively be phrased using the language of measure preserving transformations.

□

Lemma 3.23. *If $\underline{x} \in \mathbb{Z}^d$ is primitive with $|\underline{x}| \leq \lambda$ and n is an integer with $1 \leq n < \frac{1}{\sqrt{6\pi}}\lambda^{1/2}$, then*

$$\text{meas}\left\{0 \leq \beta \leq 1 \ : \ \frac{n-1}{\lambda} \leq \|q\beta\underline{x}\| < \frac{n}{\lambda}\right\} \lesssim \frac{n}{\lambda|\underline{x}|}.$$

Proof. The case $n = 1$ is handled by Lemma 3.22.

Set $\eta = \frac{1}{\sqrt{6\pi\lambda}}$ and note $\eta \leq \frac{1}{\sqrt{6\pi|\underline{x}|}}$. Lemma 3.21 tells us that there are not three linearly independent integral vectors in the η neighborhood of $\langle\underline{x}\rangle$, i.e.

$$\mathbb{Z}^d \cap \left\{\underline{y} \in \mathbb{R}^d \ : \ \text{Dist}(\underline{y}, \langle\underline{x}\rangle) < \eta\right\} \subset \langle\underline{x}, \underline{x}^*\rangle,$$

where $\underline{x}^*$ is some fixed integer vector which we may take to be primitive and within the η cylindrical neighborhood of $\underline{x}$.[28]

[28] One can make $\underline{x}^*$ explicit: Given $\underline{m} \notin \langle\underline{x}\rangle$ in the η neighborhood of $\langle\underline{x}\rangle$, we can shift $\underline{m}$ by a multiple of $\underline{x}$ to get it into the cylindrical neighborhood of $\underline{x}$, and then we can divide this $\underline{m}$ by its coordinate-wise greatest common divisor.

It is possible that $\underline{x}^* \in \langle\underline{x}\rangle$, but if this is the case, the desired estimates come from a very slight modification of the proof of Lemma 3.22: The only lattice points under consideration are multiples of $\underline{x}$, so we are looking for the measure of the βs such that $q\beta|\underline{x}|$ lies within a certain intervals of length λ^{-1} and there are $2q$ such intervals. In particular, each interval is $[a - n\lambda^{-1}, a - (n-1)\lambda^{-1}]$ or $[a + (n-1)\lambda^{-1}, a + n\lambda^{-1}]$ where $a = 0, \ldots, q$.

Henceforth we assume $\underline{x}^* \notin \langle\underline{x}\rangle$ and additionally assume $\underline{x}^*$ is chosen so $|\underline{x}^*|$ is minimized. Now we set up some notation to match our notation in Lemma 3.21. Again, let $\langle\underline{x}\rangle^\perp$ denote the hyperplane in $\mathbb{R}^d$ orthogonal to $\underline{x}$, and let D denote the interval

$$D := \left\{\underline{y} \in \langle\underline{x}\rangle^\perp \cap \langle\underline{x}^*\rangle \ : \ |\underline{y}| \leq \eta\right\}.$$

Finally we define the strip $S := D \times q\underline{x}$. The point is that S contains all the lattice points which could be the closest lattice point to $q\beta\underline{x}$ for $0 \leq \beta \leq 1$ as long as $\|q\beta\underline{x}\| \leq \eta$.

By Lemma 3.20 we know

$$\mathrm{Dist}(\underline{x}^*, \langle \underline{x} \rangle) \geq |\underline{x}|^{-1} \geq \frac{1}{\lambda}.$$

Let $\underline{m} \in S \cap \mathbb{Z}^d$ and write $\underline{m} = \ell\underline{x} + k\underline{x}^*$ where $\ell, k \in \mathbb{Z}$ with $0 \leq \ell < q$ and $0 \leq |k| < \lambda\eta$. Because $\mathrm{Dist}(\underline{m}, \langle \underline{x} \rangle) = |k| \,\mathrm{Dist}(\underline{x}^*, \langle \underline{x} \rangle)$ and $\mathrm{Dist}(\underline{x}^*, \langle \underline{x} \rangle) \geq \frac{1}{\lambda}$, there is at most one choice for $|k|$ such that

$$\frac{n-1}{\lambda} \leq \mathrm{Dist}(\underline{m}, \langle \underline{x} \rangle) < \frac{n}{\lambda},$$

where $1 \leq n \leq \lambda\eta$. Note that a choice of $|k| > 0$ corresponds to a total of $2q$ lattice points in S. Let

$$m_{n,q,1/\lambda} := \mathrm{meas}\left\{ q\beta\underline{x} \;:\; \beta \in [0,1],\, \frac{n-1}{\lambda} \leq \|q\beta\underline{x}\| < \frac{n}{\lambda} \right\}.$$

For an upper bound, we assume $\mathrm{Dist}(\underline{x}^*, \langle \underline{x} \rangle) = \frac{1}{\lambda}$ and compute in this case

$$\begin{aligned}
m_{n,q,1/\lambda} &= 2q \sum_{r=0}^{n-1} 2 \left[\sqrt{\left(\frac{n}{\lambda}\right)^2 - \left(\frac{r}{\lambda}\right)^2} - \sqrt{\left(\frac{n-1}{\lambda}\right)^2 - \left(\frac{r}{\lambda}\right)^2} \right] \\
&= \frac{4q}{\lambda} \sum_{r=0}^{n-1} \left[\sqrt{n^2 - r^2} - \sqrt{(n-1)^2 - r^2} \right] \\
&= \frac{4q(2n-1)}{\lambda} \sum_{r=0}^{n-1} \frac{1}{\sqrt{n^2 - r^2} + \sqrt{(n-1)^2 - r^2}} \\
&\lesssim \frac{qn}{\lambda} \left[\frac{1}{\sqrt{n}} + \sum_{r=0}^{n-2} \frac{1}{\sqrt{(n-1)^2 - r^2}} \right] \\
&\lesssim \frac{qn}{\lambda} \left[\frac{1}{\sqrt{n}} + \sum_{r=0}^{n-2} \frac{1}{n-1} \frac{1}{\sqrt{1 - \left(\frac{r}{n-1}\right)^2}} \right].
\end{aligned}$$

Noting that the above sum is the left-handed Riemann sum for the function $(1-x^2)^{-1}$, we have that it is in fact bounded by $\arcsin(1)$, i.e. $\pi/2$. Hence

$$m_{n,q,1/\lambda} \lesssim \frac{qn}{\lambda}, \tag{3.19}$$

and so

$$\text{meas}\left\{0 \le \beta \le 1 \;:\; \frac{n-1}{\lambda} \le \|q\beta\underline{x}\| < \frac{n}{\lambda}\right\} \lesssim \frac{n}{\lambda|\underline{x}|},$$

as desired. □

The goal of Lemmas 3.20, 3.21, 3.22, and 3.23 is of course all to get a handle on the integration of the main term M in 3.11 with respect to β. A direct application of Corollary 3.18 yields

$$\int_0^1 M \, d\beta \lesssim \lambda^d \int_0^1 \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\|q\beta\underline{x}\|\right)^{-d/2} d\beta. \tag{3.20}$$

Now we split the integration on the condition $\|q\beta\underline{x}\| < \frac{1}{\lambda}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor$, a cutoff chosen so that when the integration is divided over regions

$$\frac{n-1}{\lambda} \le \|q\beta\underline{x}\| < \frac{n}{\lambda},$$

we can safely let n be an integer.

We write

$$\int_0^1 M \, d\beta \lesssim \lambda^d I_1 + \lambda^d I_2, \tag{3.21}$$

where

$$I_1 := \int_{\|q\beta\underline{x}\| < \frac{1}{\lambda}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\|q\beta\underline{x}\|\right)^{-d/2} d\beta,$$

and

$$I_2 := \int_{\|q\beta\underline{x}\| \ge \frac{1}{\lambda}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\|q\beta\underline{x}\|\right)^{-d/2} d\beta.$$

First we estimate I_1.

Lemma 3.24. *If $\underline{x} \in \mathbb{Z}^d$ is primitive with $|\underline{x}| \le \lambda$ and $d \ge 5$, then*

$$I_1 \lesssim q^2\lambda^{-1}|\underline{x}|^{-1}\left(1 + \lambda^2|\tau|\right)^{(4-d)/2}.$$

Proof. Splitting up the integral and applying Lemma 3.23 we get

$$
\begin{aligned}
I_1 &= \int\limits_{\|q\beta\underline{x}\|<\frac{1}{\lambda}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} \left(1+\lambda^2|\tau|+\frac{\lambda}{q}\|q\beta\underline{x}\|\right)^{-d/2} d\beta \\
&= \sum_{n=1}^{\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} \int\limits_{\frac{n-1}{\lambda}\le\|q\beta\underline{x}\|<\frac{n}{\lambda}} \left(1+\lambda^2|\tau|+\frac{\lambda}{q}\|q\beta\underline{x}\|\right)^{-d/2} d\beta \\
&\le \sum_{n=1}^{\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} \int\limits_{\frac{n-1}{\lambda}\le\|q\beta\underline{x}\|<\frac{n}{\lambda}} \left(1+\lambda^2|\tau|+\frac{n-1}{q}\right)^{-d/2} d\beta \\
&\lesssim \sum_{n=1}^{\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} n\lambda^{-1}|\underline{x}|^{-1}\left(1+\lambda^2|\tau|+\frac{n-1}{q}\right)^{-d/2} \\
&\lesssim \lambda^{-1}|\underline{x}|^{-1}\sum_{n=0}^{\infty} n\left(1+\lambda^2|\tau|+\frac{n}{q}\right)^{-d/2}.
\end{aligned}
$$

Set

$$S(c,q,d) := \sum_{n=0}^{\infty} n\left(c+\frac{n}{q}\right)^{-d/2}.$$

To estimate $S(c,q,d)$, we note that a quick calculus exercise shows that

$$f(x) := x\left(c+\frac{x}{q}\right)^{-d/2}$$

is decreasing when $x > \frac{2cq}{d-2}$. In particular $f(x)$ is decreasing when $x > cq$ for any $d \ge 5$. Hence for $d \ge 5$ and $c > 0$

$$S(c,q,d) \le \sum_{n=0}^{\lceil cq\rceil} n\left(c+\frac{n}{q}\right)^{-d/2} + \int\limits_{cq}^{\infty} f(x)\,dx.$$

For the sum, note that there are $\lceil cq\rceil$ nonzero terms and $\left(c+\frac{n}{q}\right)^{-d/2} \le c^{-d/2}$. For the integral, we perform the substitution $x = cqu$. Now

$$
\begin{aligned}
S(c,q,d) &\le c^{-d/2}\lceil cq\rceil^2 + c^{2-d/2}q^2\int\limits_{1}^{\infty} u(1+u)^{-d/2}\,du \\
&\lesssim c^{(4-d)/2}q^2, \qquad (3.22)
\end{aligned}
$$

where the last integral is less than the constant that occurs setting $d = 5$. Setting $c = 1 + \lambda^2|\tau|$ and plugging $S(c, q, d)$ back into the bound for I_1 completes the proof. $\square$

The I_2 term in equation (3.21) does not require such careful handling.

Lemma 3.25. *If $\underline{x} \in \mathbb{Z}^d$ is primitive with $|\underline{x}| \leq \lambda$, then*

$$I_2 \lesssim q^{d/2}\lambda^{-d/4}.$$

Proof. We insert the bound on $\|q\beta\underline{x}\|$ and extend the integral to be over all $\beta \in [0, 1]$:

$$\begin{aligned}
I_2 &= \int\limits_{\|q\beta\underline{x}\| \geq \frac{1}{\lambda}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor} \left(1 + \lambda^2|\tau| + \frac{\lambda}{q}\|q\beta\underline{x}\|\right)^{-d/2} d\beta \\
&\leq \int\limits_0^1 \left(1 + \lambda^2|\tau| + \frac{1}{q}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor\right)^{-d/2} d\beta \\
&\leq \left(\frac{1}{q}\left\lfloor\sqrt{\frac{\lambda}{6\pi}}\right\rfloor\right)^{-d/2} \\
&\lesssim q^{d/2}\lambda^{-d/4}.
\end{aligned}$$

$\square$

We also need to integrate I_1 and I_2 over the major arcs.

Lemma 3.26. *If $\underline{x} \in \mathbb{Z}^d$ is primitive with $|\underline{x}| \leq \lambda$ and $d \geq 7$, then*

$$\int\limits_{\mathfrak{M}_{a/q}(\theta)} I_1 \, d\alpha \lesssim q^2\lambda^{-3}|\underline{x}|^{-1}.$$

Proof. By Lemma 3.24

$$\begin{aligned}
\int\limits_{\mathfrak{M}_{a/q}(\theta)} I_1 \, d\alpha &\lesssim q^2\lambda^{-1}|\underline{x}|^{-1} \int\limits_{|\tau| \leq q^{-1}\lambda^{-2+\theta}} \left(1 + \lambda^2|\tau|\right)^{(4-d)/2} d\tau \\
&\lesssim q^2\lambda^{-1}|\underline{x}|^{-1} \int_0^\infty \left(1 + \lambda^2\tau\right)^{(4-d)/2} d\tau \\
&\lesssim q^2\lambda^{-3}|\underline{x}|^{-1} \int_0^\infty (1 + x)^{(4-d)/2} dx \\
&\lesssim q^2\lambda^{-3}|\underline{x}|^{-1},
\end{aligned}$$

where the last inequality follows as we have assumed $d \geq 7$. $\square$

Lemma 3.27. *If $\underline{x} \in \mathbb{Z}^d$ is primitive with $|\underline{x}| \leq \lambda$, then*

$$\int\limits_{\mathfrak{M}_{a/q}(\theta)} I_2 \, d\alpha \lesssim q^{d/2-1}\lambda^{-2+\theta-d/4}.$$

Proof. By Lemma 3.25

$$\begin{aligned}\int\limits_{\mathfrak{M}_{a/q}(\theta)} I_2 \, d\alpha &\lesssim \int\limits_{\mathfrak{M}_{a/q}(\theta)} q^{d/2}\lambda^{-d/4} \, d\alpha \\ &\lesssim q^{d/2-1}\lambda^{-2+\theta-d/4},\end{aligned}$$

as the length of the major arc $\mathfrak{M}_{a/q}(\theta)$ is $2q^{-1}\lambda^{-2+\theta}$. $\square$

We finally have all the pieces to estimate the total contribution of the major arcs.

Theorem 3.28. *Let $\underline{x} \in \mathbb{Z}^d$ be primitive with $|\underline{x}| \leq \lambda$ where $\lambda^{1-\theta} \geq 16$. If $d \geq \max(9, 8+8\theta)$, then*

$$|I_{\mathfrak{M}(\theta)}| \lesssim_d \lambda^{d-3}|\underline{x}|^{-1} + \lambda^{d-\frac{d\theta}{2}}.$$

Proof. By equation (3.5), Lemma 3.12, and equation (3.11)

$$\begin{aligned}|I_{\mathfrak{M}(\theta)}| &\leq \int_{\mathfrak{M}(\theta)} \int_0^1 |G_{\underline{x},\lambda}(\alpha,\beta)| \, d\beta \, d\alpha \\ &\leq \sum_{1 \leq q < \lambda^\theta} \sum_{\substack{0 \leq a < q, \\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 |G_{\underline{x},\lambda}(\alpha,\beta)| \, d\beta \, d\alpha \\ &\lesssim_d \sum_{1 \leq q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \leq a < q, \\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 [M + \mathcal{E}] \, d\beta \, d\alpha \\ &\lesssim_d S_1 + S_2 + S_3,\end{aligned}$$

where using equation (3.21) we have

$$S_1 := \sum_{1 \leq q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \leq a < q, \\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \lambda^d I_1 \, d\alpha,$$

$$S_2 := \sum_{1 \le q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \le a < q, \\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \lambda^d I_2 \, d\alpha,$$

and

$$S_3 := \sum_{1 \le q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \le a < q, \\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 \mathcal{E} \, d\beta \, d\alpha.$$

From the estimate in Lemma 3.26:

$$\begin{aligned} S_1 &\lesssim \lambda^{d-3} |\underline{x}|^{-1} \sum_{1 \le q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \le a < q, \\ (a,q)=1}} q^2 \\ &\lesssim \lambda^{d-3} |\underline{x}|^{-1} \sum_{1 \le q < \lambda^\theta} q^{(6-d)/2} \\ &\lesssim \lambda^{d-3} |\underline{x}|^{-1} \sum_{q=1}^{\infty} q^{(6-d)/2} \\ &\lesssim \lambda^{d-3} |\underline{x}|^{-1}, \end{aligned}$$

where the last inequality follows as $d \ge 9$.

From the estimate in Lemma 3.27:

$$\begin{aligned} S_2 &\lesssim \lambda^{d-2+\theta-d/4} \sum_{1 \le q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \le a < q, \\ (a,q)=1}} q^{d/2-1} \\ &\lesssim \lambda^{d-2+\theta-d/4} \sum_{1 \le q < \lambda^\theta} 1 \\ &\lesssim \lambda^{d-2+2\theta-d/4}. \end{aligned}$$

As we have assumed $d \ge 8 + 8\theta$,

$$S_2 \lesssim \lambda^{d-4}.$$

And finally, from the estimate in Lemma 3.16:

$$\begin{aligned} S_3 &\lesssim_d C_K \lambda^{d-K-2} \log(\lambda) \sum_{1 \le q < \lambda^\theta} q^{-d/2} \sum_{\substack{0 \le a < q, \\ (a,q)=1}} q^K \\ &\lesssim_d C_K \lambda^{d-K-1} \sum_{1 \le q < \lambda^\theta} q^{K+1-d/2} \\ &\lesssim_d C_K \lambda^{d-K-1+\theta(K+1-d/2)} \\ &\lesssim_d C_K \lambda^{d-(K+1)(1-\theta)-\frac{d\theta}{2}} \end{aligned}$$

$$\lesssim_d C_K \lambda^{d-\frac{d\theta}{2}}.$$

Choosing say $K = d + 1$,[29] we can remove the dependence on K to get

$$S_3 \lesssim_d \lambda^{d-\frac{d\theta}{2}}.$$

Noting that $\lambda^{-1} \leq |\underline{x}|^{-1}$ and adding the estimates for S_1, S_2, and S_3 gives the desired result. $\square$

[29] For a fixed $\theta < 1$ the penultimate inequality shows we can easily beat the critical exponent $d - 4$ with a choice of K, but this way we don't need a fixed θ and get an estimate subsumed by that for the minor arcs.

3.2.4 Theorem on Upper Bounds in $\mathbb{Z}^d$

Theorem 3.29. *Let $\underline{x} \in \mathbb{Z}^d$ be primitive with $|\underline{x}| \leq \lambda$. If $d \geq 13$, then*

$$|S_{\underline{x}}(\lambda, \mu)| \lesssim_d \lambda^{d-3} |\underline{x}|^{-1}.$$

Proof. Putting the minor arc estimate from Corollary 3.7 and major arc estimate from Theorem 3.28 into equation (3.1) gives

$$\begin{aligned} |S_{\underline{x}}(\lambda, \mu)| &\leq |I_{\mathfrak{m}(\theta)}| + |I_{\mathfrak{M}(\theta)}| \\ &\lesssim_{\varepsilon,d} \lambda^{d-\frac{d\theta}{2}+\varepsilon} + \lambda^{d-3}|\underline{x}|^{-1}, \end{aligned}$$

where for the estimate on $|I_{\mathfrak{M}(\theta)}|$ to hold we need to ensure

$$\lambda^{1-\theta} \geq 16, \qquad d \geq 9, \qquad \text{and} \qquad d \geq 8 + 8\theta.$$

To get the claimed bound, we additionally need to choose $\varepsilon > 0$ such that

$$-\frac{d\theta}{2} + \varepsilon \leq -4.$$

We find the optimal choice is $\theta = \frac{\sqrt{5}-1}{2} \approx 0.618$. Choosing $\varepsilon = 0.01$ is good enough with this choice of θ to allow $d = 13$. Additionally with this choice of θ, we get $\lambda \geq 1500$ is good enough to ensure $\lambda^{1-\theta} \geq 16$, and we can increase the implied dimensional constant so the estimate is also valid for all small spheres. $\square$

3.3 Upper Bounds in a Fixed Residue Class

We now develop the techniques of the previous section to get a more refined estimate for solutions in a fixed residue class modulo W.

Let $\underline{x} \in \mathbb{Z}^d$ with $\underline{x}$ primitive and $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$ for some fixed constants c_1 and c_2. Recall Definition 3.2: Given λ, μ, $\underline{x}$ and residue class $\underline{v}$ modulo W,

$$S_{\underline{x},W,\underline{v}}(\lambda, \mu) := \{\underline{y} \in \mathbb{Z}^d \ : \underline{y} \equiv \underline{v} \pmod{W},\ |\underline{y}|^2 = \lambda^2,\ \underline{y} \cdot \underline{x} = \mu\}.$$

We will assume $|\underline{v}|^2 \equiv \lambda^2 \pmod{W}$ and $\underline{x} \cdot \underline{v} \equiv \mu \pmod{W}$, as otherwise $|S_{\underline{x},W,\underline{v}}(\lambda, \mu)| = 0$. We capture the congruence condition for $\underline{y}$ by writing $\underline{y} = W\underline{y}_1 + \underline{v}$ and summing over $\underline{y}_1$:

$$|S_{\underline{x},W,\underline{v}}(\lambda, \mu)| = \sum_{\underline{y}_1 \in \mathbb{Z}^d} \int_0^1 e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta.$$

We again split the integration into major and minor arcs as given in Definition 3.4. We set

$$|S_{\underline{x},W,\underline{v}}(\lambda, \mu)| = I_{\mathfrak{m}^W(\theta)} + I_{\mathfrak{M}^W(\theta)}, \tag{3.23}$$

with

$$I_{\mathfrak{m}^W(\theta)} = \sum_{\underline{y}_1 \in \mathbb{Z}^d} \int_{\mathfrak{m}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta,$$

$$I_{\mathfrak{M}^W(\theta)} = \sum_{\underline{y}_1 \in \mathbb{Z}^d} \int_{\mathfrak{M}(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta,$$

where again we are using $\underline{y} = W\underline{y}_1 + \underline{v}$.

3.3.1 Minor Arcs Estimate in a Fixed Residue Class

As before, we note that the only nonzero terms of the sum are those such that $|\underline{y}| = \lambda$, so the sum over $\underline{y}_1 \in \mathbb{Z}^d$ could be replaced with a sum over those $\underline{y}_1$ such that $W\underline{y}_1 + \underline{v} \in Q_\lambda$. Those $\underline{y}_1$ lie in a shifted cube of sidelength $\lesssim \lambda W^{-1}$. Let $\mathcal{C}$ denote this cube. We have

$$\begin{aligned}
I_{\mathfrak{m}^W(\theta)} &= \sum_{\underline{y}_1 \in \mathbb{Z}^d} \int_{\mathfrak{m}^W(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta \\
&= \sum_{\underline{y}_1 \in \mathcal{C}} \int_{\mathfrak{m}^W(\theta)} e(\alpha(|\underline{y}|^2 - \lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y} \cdot \underline{x} - \mu))\, d\beta \\
&= \int_{\mathfrak{m}^W(\theta)} \int_0^1 e(-\alpha\lambda^2 - \beta\mu) \sum_{\underline{y}_1 \in \mathcal{C}} e(\alpha|\underline{y}|^2 + \beta\underline{y} \cdot \underline{x})\, d\beta\, d\alpha.
\end{aligned} \tag{3.24}$$

Definition 3.30. *The Gaussian sum* $g_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta)$ *is given by*

$$\begin{aligned} g_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta) &:= \sum_{\underline{y}_1\in\mathcal{C}} e(\alpha|\underline{y}|^2+\beta\underline{y}\cdot\underline{x}) \\ &= \sum_{\underline{y}_1\in\mathcal{C}} e\left(\alpha|W\underline{y}_1+\underline{v}|^2+\beta\underline{x}\cdot(W\underline{y}_1+\underline{v})\right) \\ &= \sum_{\underline{y}_1\in\mathcal{C}} e(W^2\alpha|\underline{y}_1|^2+(2W\alpha\underline{v}+\beta\underline{x})\cdot\underline{y}_1+(\alpha\underline{v}+\beta\underline{x})\cdot\underline{v}), \end{aligned}$$

where $\mathcal{C}$ *is the cube* $\{\underline{y}_1\in\mathbb{Z}^d\ :\ W\underline{y}_1+\underline{v}\in Q_\lambda\}$*.*

Lemma 3.31. *Let* $0<\delta<\theta$*. If* $\alpha\in\mathfrak{m}(\theta)$ *and* $W^2\leq\lambda^\delta$*, then*

$$\{W^2\alpha\}\in\mathfrak{m}(\theta-\delta).$$

Proof. Suppose $q\leq\lambda^{\theta-\delta}$. Let a be some integer with $0\leq a<q$. Then

$$\begin{aligned} \left|\{W^2\alpha\}-\frac{a}{q}\right| &= \left|W^2\alpha-\lfloor W^2\alpha\rfloor-\frac{a}{q}\right| \\ &= W^2\left|\alpha-\frac{a'}{W^2q}\right|, \end{aligned}$$

for some integer a'.

Note that $W^2q\leq\lambda^\theta$ as $W^2\leq\lambda^\delta$ by assumption. Since $\alpha\in\mathfrak{m}(\theta)$ we have

$$\left|\alpha-\frac{a'}{q'}\right|>\frac{1}{q'}\lambda^{-2+\theta}$$

for every $q'\leq\lambda^\theta$. Plugging in W^2q for q' shows

$$\left|\{W^2\alpha\}-\frac{a}{q}\right|>W^2\frac{1}{W^2q}\lambda^{-2+\theta}$$

for every integer a and $q\leq\lambda^{\theta-\delta}$. Noting $\lambda^{-2+\theta}>\lambda^{-2+\theta-\delta}$, we see

$$\{W^2\alpha\}\in\mathfrak{m}(\theta-\delta).$$

□

Lemma 3.32 (Weyl estimate)**.** *Let* $0<\delta<\theta$*. If* $\alpha\in\mathfrak{m}(\theta)$ *and* $W^2\leq\lambda^\delta$*, then for any* $\varepsilon>0$*,*

$$|g_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta)|\lesssim_\varepsilon\lambda^{d-\frac{d(\theta-\delta)}{2}+\varepsilon}.$$

Proof. By Lemma 3.31, we have $\{W^2\alpha\} \in \mathfrak{m}(\theta-\delta)$. Replacing θ with $\theta-\delta$ in Lemma 3.6 immediately yields the result as the coefficients of the lower order terms do not affect the Weyl estimate. □

Corollary 3.33. *Let* $0<\delta<\theta$. *If* $W^2 \le \lambda^\delta$, *then for any* $\varepsilon>0$,

$$|I_{\mathfrak{m}^W(\theta)}| \lesssim_\varepsilon \lambda^{d-\frac{d(\theta-\delta)}{2}+\varepsilon}.$$

Proof. Bringing in the absolute values in equation (3.24) gives

$$|I_{\mathfrak{m}^W(\theta)}| \le \int_{\mathfrak{m}(\theta)} \int_0^1 |g_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta)|\, d\beta\, d\alpha \lesssim_\varepsilon \lambda^{d-\frac{d(\theta-\delta)}{2}+\varepsilon}$$

□

3.3.2 Major Arcs Estimate in a Fixed Residue Class

The inclusion of W in the major arcs estimate shall add some technical difficulty and necessitate a description of an exceptional set in which our desired estimates do not quite hold.[30]

[30] This set, however, is not too large.

Let $\varphi(\underline{y})$ be as defined in Definition 3.8. Now

$$\begin{aligned} I_{\mathfrak{M}^W(\theta)} &= \sum_{\underline{y}_1\in\mathbb{Z}^d} \int_{\mathfrak{M}(\theta)} e(\alpha(|\underline{y}|^2-\lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y}\cdot\underline{x}-\mu))\, d\beta \\ &= \sum_{\underline{y}_1\in\mathbb{Z}^d} \varphi\left(\frac{\underline{y}}{\lambda}\right) \int_{\mathfrak{M}(\theta)} e(\alpha(|\underline{y}|^2-\lambda^2))\, d\alpha \int_0^1 e(\beta(\underline{y}\cdot\underline{x}-\mu))\, d\beta \\ &= \int_{\mathfrak{M}(\theta)} \int_0^1 e(-\alpha\lambda^2-\beta\mu) \sum_{\underline{y}_1\in\mathbb{Z}^d} e(\alpha|\underline{y}|^2+\beta\underline{y}\cdot\underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right) d\beta\, d\alpha \end{aligned}$$

Definition 3.34. *We define the Gaussian sum for major arcs as*

$$G_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta) = \sum_{\underline{y}_1\in\mathbb{Z}^d} e(\alpha|\underline{y}|^2+\beta\underline{y}\cdot\underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right),$$

where $\underline{y} = W\underline{y}_1 + \underline{v}$.

Pulling in the absolute value, we have

$$|I_{\mathfrak{M}^W(\theta)}| \le \int_{\mathfrak{M}(\theta)} \int_0^1 |G_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta)|\, d\beta\, d\alpha. \tag{3.25}$$

Recall the notation set up in Definition 3.10.

Lemma 3.35. *If* $\alpha = \frac{a}{q} + \tau$ *with* $0 \leq a < q$ *and* $(a, q) = 1$*, then*

$$G_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta) = W^{-d}q^{-d} \sum_{\underline{s}_1 \pmod q} e\left(\frac{a}{q}|\underline{s}|^2\right) \times \sum_{\underline{m}\in\mathbb{Z}^d} e\left(\frac{\underline{s}\cdot\underline{m}}{Wq}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - Wq\beta\underline{x}}{Wq}\right),$$

where $\underline{s} = W\underline{s}_1 + \underline{v}$.

Proof. The proof follows the ideas of Lemma 3.11. We write $\underline{y}_1 = q\underline{y}_2 + \underline{s}_1$ with $\underline{s}_1 \in [q]^d$. Note under this notation

$$\underline{y} = W\underline{y}_1 + \underline{v} = Wq\underline{y}_2 + W\underline{s}_1 + \underline{v} = Wq\underline{y}_2 + \underline{s}.$$

Writing G for $G_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta)$, we have

$$\begin{aligned}
G &= \sum_{\underline{y}_1\in\mathbb{Z}^d} e(\alpha|\underline{y}|^2 + \beta\underline{y}\cdot\underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right) \\
&= \sum_{\underline{s}_1\,(q)} \sum_{\underline{y}_2\in\mathbb{Z}^d} e(\alpha|\underline{y}|^2 + \beta\underline{y}\cdot\underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right) \\
&= \sum_{\underline{s}_1\,(q)} \sum_{\underline{y}_2\in\mathbb{Z}^d} e\left(\frac{a}{q}|Wq\underline{y}_2 + \underline{s}|^2 + \tau|\underline{y}|^2 + \beta\underline{y}\cdot\underline{x}\right)\varphi\left(\frac{\underline{y}}{\lambda}\right) \\
&= \sum_{\underline{s}_1\,(q)} e\left(\frac{a}{q}|\underline{s}|^2\right) \sum_{\underline{y}_2\in\mathbb{Z}^d} e(\tau|\underline{y}|^2 + \beta\underline{y}\cdot\underline{x})\varphi\left(\frac{\underline{y}}{\lambda}\right) \\
&= \sum_{\underline{s}_1\,(q)} e\left(\frac{a}{q}|\underline{s}|^2\right) \sum_{\underline{y}_2\in\mathbb{Z}^d} e(\beta\underline{y}\cdot\underline{x})\varphi_{\tau,\lambda}(\underline{y}) \\
&= \sum_{\underline{s}_1\,(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{x}\cdot\underline{s}\right) \sum_{\underline{y}_2\in\mathbb{Z}^d} e(Wq\beta\underline{x}\cdot\underline{y}_2)\varphi_{\tau,\lambda}(Wq\underline{y}_2 + \underline{s}) \\
&= \sum_{\underline{s}_1\,(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{x}\cdot\underline{s}\right) \sum_{\underline{y}_2\in\mathbb{Z}^d} e(Wq\beta\underline{x}\cdot\underline{y}_2)\varphi_{\tau,\lambda,Wq,\underline{s}}(\underline{y}_2) \\
&= \sum_{\underline{s}_1\,(q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{x}\cdot\underline{s}\right) \widehat{\varphi}_{\tau,\lambda,Wq,\underline{s}}(-Wq\beta\underline{x})
\end{aligned}$$

By Poisson summation, we have

$$\widehat{\varphi}_{\tau,\lambda,Wq,\underline{s}}(-Wq\beta\underline{x}) = \sum_{\underline{m}\in\mathbb{Z}^d} \widetilde{\varphi}_{\tau,\lambda,Wq,\underline{s}}(\underline{m} - Wq\beta\underline{x}),$$

and by equation (3.9), we have

$$\widetilde{\varphi}_{\tau,\lambda,Wq,\underline{s}}(\underline{\xi}) = W^{-d}q^{-d}e\left(\frac{\underline{\xi}\cdot\underline{s}}{Wq}\right)\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{\xi}}{Wq}\right).$$

Hence

$$\begin{aligned} G &= W^{-d}q^{-d}\sum_{\underline{s}_1\ (q)} e\left(\frac{a}{q}|\underline{s}|^2 + \beta\underline{x}\cdot\underline{s}\right) \\ &\quad\times \sum_{\underline{m}\in\mathbb{Z}^d} e\left(\frac{(\underline{m} - Wq\beta\underline{x})\cdot\underline{s}}{Wq}\right)\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - Wq\beta\underline{x}}{Wq}\right) \\ &= W^{-d}q^{-d}\sum_{\underline{s}_1\ (q)} e\left(\frac{a}{q}|\underline{s}|^2\right)\sum_{\underline{m}\in\mathbb{Z}^d} e\left(\frac{\underline{m}\cdot\underline{s}}{Wq}\right)\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - Wq\beta\underline{x}}{Wq}\right). \end{aligned}$$

□

Because of the involvement of W, a more delicate analysis is required than before.[31] We separate the sum into a main term M and an error term $\mathcal{E}$ as essentially only one term in the inner some is important:

$$M := W^{-d}q^{-d}\sum_{\underline{s}_1\ (q)} e\left(\frac{a}{q}|\underline{s}|^2\right)e\left(\frac{\underline{m}_0\cdot\underline{s}}{Wq}\right)\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m}_0 - Wq\beta\underline{x}}{Wq}\right),$$

$$\mathcal{E} := W^{-d}q^{-d}\sum_{\underline{s}_1\ (q)} e\left(\frac{a}{q}|\underline{s}|^2\right)\sum_{\underline{m}\neq\underline{m}_0} e\left(\frac{\underline{m}\cdot\underline{s}}{Wq}\right)\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m} - Wq\beta\underline{x}}{Wq}\right),$$

where $\underline{m}_0 = [-Wqby]$.[32]

We write

$$G_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta) = M + \mathcal{E}. \tag{3.26}$$

We are at least able to deal with $\mathcal{E}$ via essentially the same estimation tool as the previous section.

Lemma 3.36. *If there exists a* $\delta > 0$ *with* $W^2 < \lambda^\delta$ *and* $\lambda^{1-\theta-\delta/2} \geq 16$,[33] *then*

[31] In the first section we exchanged summation and were able to get a uniform bound on the Gaussian sum because of the coprimality of a and q. Here however after rewriting $\underline{s} = W\underline{s}_1 + \underline{v}$ there are complications as W may not be coprime to q.

[32] Recall if $\underline{y} \in \mathbb{R}^d$ then $[\underline{y}]$ denotes the nearest lattice point.

[33] Although it would be simpler to write the conditions $W < \lambda^\delta$ and $\lambda^{1-\theta-\delta} \geq 16$, we write it this way because elsewhere we want to use a δ with $W^2 < \lambda^\delta$.

for any integer $K > d$ there exists a C_K such that

$$\int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 |\mathcal{E}| \, d\beta \, d\alpha \lesssim_d C_K W^{K-d} q^K \lambda^{d-K-2} \log(\lambda).$$

Proof. Because this estimate has leeway for us later in the form of a power savings in λ, we may wastefully[34] bring in the absolute values to obtain

[34] We make no attempt to optimize constants.

$$\begin{aligned} |\mathcal{E}| &\le W^{-d} q^{-d} \sum_{\underline{s}_1 \, (q)} \sum_{\underline{m} \neq \underline{m}_0} \left| \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\underline{m} - W q \beta \underline{x}}{W q} \right) \right| \\ &\le W^{-d} \sum_{\underline{m} \neq \underline{m}_0} \left| \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\underline{m} - W q \beta \underline{x}}{W q} \right) \right|, \end{aligned} \tag{3.27}$$

where for convenience, we define

$$\mathcal{E}' := \sum_{\underline{m} \neq \underline{m}_0} \left| \widetilde{\varphi}_{\tau,\lambda} \left(\frac{\underline{m} - W q \beta \underline{x}}{W q} \right) \right|.$$

Let $\alpha \in \mathfrak{M}_{a/q}(\theta)$ and set $\underline{\xi} = \frac{\underline{m} - W q \beta \underline{x}}{W q}$ where $\underline{m} \neq \underline{m}_0$. As $\underline{m} \neq \underline{m}_0$ we have that $\max_{1 \le i \le d} |\xi_i| \ge \frac{1}{2Wq}$. Additionally, as we assumed $W^2 < \lambda^\delta$, we have

$$Wq \le W \lambda^\theta < \lambda^{\theta + \delta/2}.$$

Replacing q with Wq and θ with $\theta + \delta/2$ in Lemma 3.15 tells us that provided $\lambda^{1-\theta-\delta/2} \ge 16$ and $0 < |\tau| \le (Wq)^{-1} \lambda^{-2+\theta+\delta/2}$, then for any $K > 0$ there exists a constant C_K such that

$$|\widetilde{\varphi}_{\tau,\lambda}(\underline{\xi})| \lesssim_d C_K \lambda^d \min\left(1, |\tau|^{-1} \lambda^{-2} (\lambda \max_{1 \le i \le d} |\xi_i|)^{-K}\right). \tag{3.28}$$

The condition $\lambda^{1-\theta-\delta/2} \ge 16$ holds by assumption, and as α is in the major arc $\mathfrak{M}_{a/q}(\theta)$, we indeed have

$$|\tau| \le \frac{1}{q} \lambda^{-2+\theta} \le \frac{1}{Wq} \lambda^{-2+\theta+\delta/2},$$

so equation (3.28) holds.

We can now directly apply the same computation in Lemma 3.16 using Wq in place of q to get

$$\int\limits_{\mathfrak{M}_{a/q}(\theta)} \int\limits_0^1 \mathcal{E}' \, d\beta \, d\alpha \lesssim_d C_K (Wq)^K \lambda^{d-K-2} \log(\lambda)$$

Including the W^{-d} factor from equation (3.27) we achieve the claimed estimate. □

We now turn to the main term, though it will turn out that we need to identify a small exceptional set of $\underline{x}$'s for which we get a mildly worse estimate than expected. First we recall $\underline{s} = W\underline{s}_1 + \underline{v}$ and note

$$\begin{aligned} M &= (Wq)^{-d} \sum_{\underline{s}_1 \, (q)} e\left(\frac{a}{q}|\underline{s}|^2\right) e\left(\frac{\underline{m}_0 \cdot \underline{s}}{Wq}\right) \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\underline{m}_0 - Wq\beta\underline{x}}{Wq}\right) \\ &= (Wq)^{-d} \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\{-Wq\beta\underline{x}\}}{Wq}\right) \sum_{\underline{s}_1 \, (q)} e\left(\frac{a}{q}|\underline{s}|^2 + \frac{\underline{m}_0 \cdot \underline{s}}{Wq}\right) \\ &= (Wq)^{-d} \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\{-Wq\beta\underline{x}\}}{Wq}\right) e\left(\frac{\underline{m}_0 \cdot \underline{v}}{Wq}\right) \sum_{\underline{s}_1 \, (q)} e\left(\frac{a|\underline{s}|^2 + \underline{m}_0 \cdot \underline{s}_1}{q}\right), \end{aligned}$$

so we have

$$|M| \leq W^{-d} \left| \widetilde{\varphi}_{\tau,\lambda}\left(\frac{\{-Wq\beta\underline{x}\}}{Wq}\right) \right| S, \tag{3.29}$$

where S is the normalized Gaussian sum

$$S := q^{-d} \sum_{\underline{s}_1 \pmod q} e\left(\frac{a|W\underline{s}_1 + \underline{v}|^2 + \underline{m}_0 \cdot \underline{s}_1}{q}\right).$$

Lemma 3.37. *Let* $D = \gcd(W, q)$, *let* $q = q_1 D$, *and assume* $\gcd(a, q) = 1$. *If* $D \nmid \underline{m}_0$ *then* $S = 0$. *On the other hand, if* $D \mid \underline{m}_0$, *then*

$$|S| \lesssim_d q_1^{-d/2}.$$

Proof. We write $\underline{s}_1 = q_1 \underline{t} + \underline{r}$ with $\underline{t} \in [D]^d$ and $\underline{r} \in [q_1]^d$. Now we have

$$\begin{aligned} S &= q^{-d} \sum_{\underline{s}_1 \, (q)} e\left(\frac{a|W\underline{s}_1 + \underline{v}|^2 + \underline{m}_0 \cdot \underline{s}_1}{q}\right) \\ &= q^{-d} \sum_{\underline{r} \, (q_1)} \sum_{\underline{t} \, (D)} e\left(\frac{a|Wq_1\underline{t} + W\underline{r} + \underline{v}|^2 + q_1\underline{m}_0 \cdot \underline{t} + \underline{m}_0 \cdot \underline{r}}{q}\right) \end{aligned}$$

$$= q^{-d} \sum_{\underline{r}\,(q_1)} \sum_{\underline{t}\,(D)} e\left(\frac{a|W\underline{r}+\underline{v}|^2 + q_1\underline{m}_0\cdot\underline{t} + \underline{m}_0\cdot\underline{r}}{q}\right)$$
$$= q^{-d} \sum_{\underline{r}\,(q_1)} e\left(\frac{a|W\underline{r}+\underline{v}|^2 + \underline{m}_0\cdot\underline{r}}{q}\right) \sum_{\underline{t}\,(D)} e\left(\frac{\underline{m}_0\cdot\underline{t}}{D}\right)$$
$$= q_1^{-d} \sum_{\underline{r}\,(q_1)} e\left(\frac{a|W\underline{r}+\underline{v}|^2 + \underline{m}_0\cdot\underline{r}}{q}\right) \frac{1}{D^d} \sum_{\underline{t}\,(D)} e\left(\frac{\underline{m}_0\cdot\underline{t}}{D}\right).$$

This last term is the indicator function of the condition $D \mid \underline{m}_0$. Hence if $D \nmid \underline{m}_0$, then $S = 0$. Now suppose $\underline{m}_0 = D\underline{m}_0'$ with $\underline{m}_0' \in \mathbb{Z}^d$. In that case, we write $W = W_1 D$ and have

$$S = q_1^{-d} \sum_{\underline{r}\,(q_1)} e\left(\frac{a|W\underline{r}+\underline{v}|^2 + D\underline{m}_0'\cdot\underline{r}}{q}\right)$$
$$= q_1^{-d} e\left(\frac{a}{q}|\underline{v}|^2\right) \sum_{\underline{r}\,(q_1)} e\left(\frac{aWW_1|\underline{r}|^2 + (2aW_1\underline{v} + \underline{m}_0')\cdot\underline{r}}{q_1}\right)$$

As $\gcd(WW_1, q_1) = 1$ we may reorder the sum $\underline{r} \to WW_1\underline{r} \mod q_1$ and apply equation (3.10) to obtain

$$|S| \lesssim_d q_1^{-d/2}.$$

□

Lemma 3.38. *Let* $D = \gcd(W, q)$ *and let* $q = q_1 D$. *If* $D \mid [Wq\beta\underline{x}]$, *then*

$$|M| \lesssim_d q_1^{-d/2} \lambda^d W^{-d} \left(1 + \lambda^2|\tau| + \frac{\lambda}{Wq}\|Wq\beta\underline{x}\|\right)^{-d/2}.$$

If $D \nmid [Wq\beta\underline{x}]$, *then* $M = 0$.

Proof. In light of Lemma 3.37, we have $M = 0$ unless $D \mid \underline{m}_0$ which was defined as $\underline{m}_0 := [-Wq\beta\underline{x}]$. By Lemma 3.17 we have

$$\left|\widetilde{\varphi}_{\tau,\lambda}\left(\frac{\{-Wq\beta\underline{x}\}}{Wq}\right)\right| \lesssim \lambda^d \left(1 + \lambda^2|\tau| + \frac{\lambda}{Wq}\|Wq\beta\underline{x}\|\right)^{-d/2}. \quad (3.30)$$

Putting equation (3.29) together with Lemma 3.37 and equation (3.30), we obtain the claimed estimate.[35] □

[35] We have removed negative signs as $\|-\underline{y}\| = \|\underline{y}\|$ and $D \mid -\underline{y}$ is an equivalent condition to $D \mid \underline{y}$.

We will need to integrate M in β, so as we did for equation (3.21), we aim to understand the level sets of $\|Wq\beta\underline{x}\|$ with the condition that $D \mid [Wq\beta\underline{x}]$.

We will have better estimates for those $\underline{x}$ such that $\langle \underline{x} \rangle$ lies far away from lattice points. To capture those $\underline{x}$ we wish to consider, we define an exceptional set E_k.

Definition 3.39. *The exceptional set E_k is given by*

$$E_k := \left\{ \underline{x} \in \mathbb{Z}^d \;:\; \min_{\underline{y} \in \mathbb{Z}^d \setminus \langle \underline{x} \rangle} \mathrm{Dist}(\underline{y}, \langle \underline{x} \rangle) < \frac{k}{|\underline{x}|} \right\}.$$

Note that Lemma 3.20 shows $E_1 = \emptyset$. We now update lemmas 3.22 and 3.23 to include the parameters D and k.

Lemma 3.40. *Let $D \mid Wq$. If $\underline{x} \in \mathbb{Z}^d \setminus E_k$ is primitive, then*

$$\mathrm{meas} \left\{ \beta \in [0,1] \;:\; D \mid [Wq\beta\underline{x}],\; \|Wq\beta\underline{x}\| < \frac{Dk}{|\underline{x}|} \right\} \leq \frac{2k}{|\underline{x}|^2}.$$

Proof. Let $\underline{m} := [Wq\beta\underline{x}]$. As $D \mid \underline{m}$, we may write $\underline{m} = D\underline{m}'$ with $\underline{m}' \in \mathbb{Z}^d$ and compute

$$\mathrm{Dist}(\underline{m}', \langle \underline{x} \rangle) = \frac{1}{D}\mathrm{Dist}(\underline{m}, \langle \underline{x} \rangle) < \frac{k}{|\underline{x}|}.$$

As $\underline{x} \notin E_k$, this means $\underline{m}' \in \langle \underline{x} \rangle$ and so $\underline{m}$ is of the form $\ell D\underline{x}$. Therefore $Wq\beta\underline{x}$ lies in the radius $Dk|\underline{x}|^{-1}$ balls centered at $\ell D\underline{x}$ for $\ell = 0, \dots, Wq/D$. The measure of all such $Wq\beta\underline{x}$ is $2Wqk|\underline{x}|^{-1}$.[36] The result follows by scaling the measure by $Wq|\underline{x}|$. □

[36] We have combined the half-balls around $\underline{0}$ and $Wq\underline{x}$ into a single ball.

Lemma 3.41. *Let $D \mid Wq$ and let n be an integer with $1 \leq n < \frac{1}{k\sqrt{6\pi}}|\underline{x}|^{1/2}$. If $\underline{x} \in \mathbb{Z}^d \setminus E_k$ is primitive , then*

$$\mathrm{meas} \left\{ \beta \in [0,1] : D \mid [Wq\beta\underline{x}],\; T_{n-1} \leq \|Wq\beta\underline{x}\| < T_n \right\} \lesssim \frac{nk}{|\underline{x}|^2},$$

where $T_n = nDk|\underline{x}|^{-1}$.

Proof. The case $n = 1$ is handled by Lemma 3.40.

Set $\eta = \frac{1}{\sqrt{6\pi|\underline{x}|}}$. First note that if $\underline{y}_1$, $\underline{y}_2$, and $\underline{y}_3$ are integral vectors such that $D \mid \underline{y}_i$ and $\mathrm{Dist}(\underline{y}_i, \langle \underline{x} \rangle) < D\eta$, then $\underline{y}_i = D\underline{y}_i'$ where $\underline{y}_i' \in \mathbb{Z}^d$ and $\mathrm{Dist}(\underline{y}_i', \langle \underline{x} \rangle) < \eta$. However, Lemma 3.21 shows that $\underline{y}_1'$, $\underline{y}_2'$, and $\underline{y}_3'$ must be linearly dependent, and so in turn $\underline{y}_1$, $\underline{y}_2$, and $\underline{y}_3$ must be linearly dependent as well.

If the only such $[Wq\beta\underline{x}]$ are in $\langle\underline{x}\rangle$, then they are again of the form $\ell D\underline{x}$ for $\ell = 0, \ldots, Wq/D$. In this case by the same computation as in the previous lemma, we have

$$\text{meas}\,\{\beta \in [0,1] : D \mid [Wq\beta\underline{x}],\, T_{n-1} \leq \|Wq\beta\underline{x}\| < T_n\} \leq \frac{2k}{|\underline{x}|^2}.$$

Now we assume there is some primitive $\underline{x}^* \in \mathbb{Z}^d \setminus \langle\underline{x}\rangle$ within the η cylindrical neighborhood of $\{\beta\underline{x} \;:\; \beta \in [0,1]\}$.[37] For convenience let us assume $\underline{x}^*$ is the shortest such vector. As $\underline{x} \notin E_k$, we have $\text{Dist}(\underline{x}^*, \langle\underline{x}\rangle) \geq k|\underline{x}|^{-1}$. Additionally, all of the lattice points in the η-neighborhood of the line segment from $\underline{0}$ to $Wq\underline{x}$ can be expressed as $\ell_1\underline{x} + \ell_2\underline{x}^*$ with $0 \leq \ell_1 \leq Wq$ and $|\ell_2| \leq \eta|\underline{x}|k^{-1}$. If $\underline{m}' = \ell_1\underline{x} + \ell_2\underline{x}^*$ then $\text{Dist}(\underline{m}', \langle\underline{x}\rangle) = \ell_2\,\text{Dist}(\underline{x}^*, \langle\underline{x}\rangle)$.

[37] A priori we know there is a $\underline{x}^*$ in the η-neighborhood of $\{Wq\beta\underline{x} \;:\; \beta \in [0,1]\}$, but this $\underline{x}^*$ may then be shifted by a multiple of $\underline{x}$ to be in the η-neighborhood of $\{\beta\underline{x} \;:\; \beta \in [0,1]\}$.

We want to scale these by D, so in particular we note there are at most $2Wq/D$ lattice points $\underline{m}'$ in the η-neighborhood of the line segment from $\underline{0}$ to $(Wq/D)\underline{x}$ satisfying

$$(n-1)k|\underline{x}|^{-1} \leq \text{Dist}(\underline{m}', \langle\underline{x}\rangle) < nk|\underline{x}|^{-1},$$

where $1 \leq n \leq \eta|\underline{x}|k^{-1}$. Noting $nk|\underline{x}|^{-1} = T_n/D$, we scale by D to find there are at most $2Wq/D$ lattice points $\underline{m}$ with $D \mid \underline{m}$ where $\underline{m}$ is in between distance T_{n-1} and distance T_n from the line segment $\underline{0}$ to $Wq\underline{x}$.

By equation (3.19) we have

$$m_{n,\frac{Wq}{D},\frac{kD}{|\underline{x}|}} \lesssim \frac{Wqkn}{|\underline{x}|}.$$

Hence

$$\text{meas}\,\{\beta \in [0,1] : D \mid [Wq\beta\underline{x}],\, T_{n-1} \leq \|Wq\beta\underline{x}\| < T_n\} \lesssim \frac{nk}{|\underline{x}|^2}.$$

□

At this point, to condense notation, we use the following definitions:

$$C := 1 + \lambda^2|\tau|,$$

$$T_n := nDk|\underline{x}|^{-1},$$

and

$$T := T_N \text{ where } N = \left\lfloor \frac{1}{k}\sqrt{\frac{|\underline{x}|}{6\pi}} \right\rfloor.$$

We split the integration of M on the condition $\|Wq\beta\underline{x}\| < T$, and then further subdivide over T_n.

By Lemma 3.38 if $D = \gcd(W, q)$, where $q = q_1 D$ and $D \mid [Wq\beta\underline{x}]$, then

$$\int_0^1 |M| \lesssim_d q_1^{-d/2}\lambda^d W^{-d}[I_1^W + I_2^W], \tag{3.31}$$

where

$$I_1^W := \int\limits_{\|Wq\beta\underline{x}\|<T} \left(C + \frac{\lambda}{qW}\|Wq\beta\underline{x}\|\right)^{-d/2} d\beta,$$

and

$$I_2^W := \int\limits_{\|Wq\beta\underline{x}\|\geq T} \left(C + \frac{\lambda}{qW}\|Wq\beta\underline{x}\|\right)^{-d/2} d\beta.$$

Lemma 3.42. *If* $\underline{x} \in \mathbb{Z}^d \setminus E_k$ *is primitive with* $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$ *and* $d \geq 5$, *then*

$$I_1^W \lesssim_d k\lambda^{-2}C^{-d/2} + q_1^2 k^{-1} W^2 \lambda^{-2} C^{(4-d)/2}.$$

Proof. Splitting up the integral and applying Lemma 3.41 we get

$$\begin{aligned}
I_1^W &= \int\limits_{\|Wq\beta\underline{x}\|<T} \left(C + \frac{\lambda}{qW}\|Wq\beta\underline{x}\|\right)^{-d/2} d\beta \\
&= \sum_{n=1}^{N} \int\limits_{T_{n-1}\leq\|Wq\beta\underline{x}\|<T_n} \left(C + \frac{\lambda}{qW}\|Wq\beta\underline{x}\|\right)^{-d/2} d\beta \\
&\leq \sum_{n=1}^{N} \int\limits_{T_{n-1}\leq\|Wq\beta\underline{x}\|<T_n} \left(C + \frac{\lambda T_{n-1}}{qW}\right)^{-d/2} d\beta \\
&\lesssim_d \sum_{n=1}^{N} nk|\underline{x}|^{-2}\left(C + \frac{(n-1)Dk\lambda}{qW|\underline{x}|}\right)^{-d/2} \\
&\lesssim_d k|\underline{x}|^{-2}C^{-d/2} + k|\underline{x}|^{-2}\sum_{n=0}^{\infty} n\left(C + \frac{nk\lambda}{q_1 W|\underline{x}|}\right)^{-d/2} \\
&\lesssim_d k|\underline{x}|^{-2}C^{-d/2} + q_1^2 k^{-1} W^2 \lambda^{-2} C^{(4-d)/2} \\
&\lesssim_d k\lambda^{-2}C^{-d/2} + q_1^2 k^{-1} W^2 \lambda^{-2} C^{(4-d)/2},
\end{aligned}$$

where in the penultimate estimate, we have used equation (3.22) and in the last estimate we have used $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$. □

Lemma 3.43. *If* $\underline{x} \in \mathbb{Z}^d \setminus E_k$ *is primitive with* $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$*, then*

$$I_2^W \lesssim_d q_1^{d/2} W^{d/2} \lambda^{-d/4}.$$

Proof. We insert the bound on $\|Wq\beta\underline{x}\|$ and extend the integral to be over all $\beta \in [0,1]$:

$$\begin{aligned} I_2^W &= \int\limits_{\|Wq\beta\underline{x}\| \geq T} \left(C + \frac{\lambda}{qW} \|Wq\beta\underline{x}\| \right)^{-d/2} d\beta \\ &\lesssim_d \int\limits_0^1 \left(C + \frac{\lambda T}{qW} \right)^{-d/2} d\beta \\ &\lesssim_d \left(\frac{\lambda T}{qW} \right)^{-d/2}. \end{aligned}$$

Using the definition of T and applying $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$ yields the claimed estimate. □

As before, we now integrate I_1^W and I_2^W over the major arcs. Recall that we have set $C = 1 + \lambda^2|\tau|$.

Lemma 3.44. *If* $\underline{x} \in \mathbb{Z}^d \setminus E_k$ *is primitive with* $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$ *and* $d \geq 7$*, then*

$$\int\limits_{\mathfrak{M}_{a/q}(\theta)} I_1^W \, d\alpha \lesssim_d k\lambda^{-4} + q_1^2 k^{-1} W^2 \lambda^{-4}.$$

Proof. We apply Lemma 3.42 and integrate. Using exactly the same process as Lemma 3.26, we increase the domain of integration, gain a factor of λ^{-2} by substitution, and require $d \geq 7$ to get convergence in the integral. □

Lemma 3.45. *If* $\underline{x} \in \mathbb{Z}^d \setminus E_k$ *is primitive with* $c_1\lambda \leq |\underline{x}| \leq c_2\lambda$*, then*

$$\int\limits_{\mathfrak{M}_{a/q}(\theta)} I_2^W \, d\alpha \lesssim_d q^{-1} q_1^{d/2} W^{d/2} \lambda^{-2+\theta-d/4}.$$

Proof. By Lemma 3.43

$$\begin{aligned} \int\limits_{\mathfrak{M}_{a/q}(\theta)} I_2^W \, d\alpha &\lesssim_d \int\limits_{\mathfrak{M}_{a/q}(\theta)} q_1^{d/2} W^{d/2} \lambda^{-d/4} \, d\alpha \\ &\lesssim_d q^{-1} q_1^{d/2} W^{d/2} \lambda^{-2+\theta-d/4}, \end{aligned}$$

where we have used the length of the major arc. $\square$

Putting everything together and applying the observation that the main term balances when $k = W$, we have the following.

Theorem 3.46. *Let* $0 < \theta < 1$. *Let* $\underline{x} \in \mathbb{Z}^d$ *be primitive with* $c_1\lambda \le |\underline{x}| \le c_2\lambda$. *If* $\delta > 0$ *with* $W^2 < \lambda^\delta$ *and* $\lambda^{1-\theta-\delta/2} \ge 16$, *then*

$$|I_{\mathfrak{M}^W(\theta)}| \lesssim_d \begin{cases} \lambda^{d-4}W^{2-d} & \text{for } \underline{x} \in \mathbb{Z}^d \setminus E_W, \\ \lambda^{d-4}W^{3-d} & \text{for } \underline{x} \in E_W \end{cases}$$

provided

$$d \ge \max\left(9,\ \frac{8+8\theta-4\delta}{1-\delta}\right).$$

Proof. By equation (3.25), equation (3.26), and equation (3.31)

$$\begin{aligned} |I_{\mathfrak{M}^W(\theta)}| &\le \int_{\mathfrak{M}(\theta)} \int_0^1 |G_{\underline{x},\lambda,W,\underline{v}}(\alpha,\beta)|\, d\beta\, d\alpha \\ &\le \sum_{1\le q<\lambda^\theta} \sum_{\substack{0\le a<q,\\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 |M+\mathcal{E}|\, d\beta\, d\alpha \\ &\lesssim_d S_1 + S_2 + S_3, \end{aligned}$$

where

$$S_1 := \sum_{1\le q<\lambda^\theta} \sum_{\substack{0\le a<q,\\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} q_1^{-d/2}\lambda^d W^{-d} I_1^W\, d\alpha,$$

$$S_2 := \sum_{1\le q<\lambda^\theta} \sum_{\substack{0\le a<q,\\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} q_1^{-d/2}\lambda^d W^{-d} I_2^W\, d\alpha,$$

and

$$S_3 := \sum_{1\le q<\lambda^\theta} \sum_{\substack{0\le a<q,\\ (a,q)=1}} \int_{\mathfrak{M}_{a/q}(\theta)} \int_0^1 |\mathcal{E}|\, d\beta\, d\alpha.$$

We first address the main term S_1. Note for $k \le W$ the second term in the estimate from Lemma 3.44 dominates and we have

$$\int_{\mathfrak{M}_{a/q}(\theta)} I_1^W\, d\alpha \lesssim_d q_1^2 k^{-1} W^2 \lambda^{-4}.$$

Recalling $q = q_1 D$ where $D = \gcd(W, q)$ and letting $\phi(n)$ denote the totient function, we compute

$$\begin{aligned} S_1 &\lesssim_d k^{-1}\lambda^{d-4}W^{-d+2}\sum_{1\le q<\lambda^\theta}\sum_{\substack{0\le a<q,\\(a,q)=1}} q_1^{(4-d)/2} \\ &\lesssim_d k^{-1}\lambda^{d-4}W^{-d+2}\sum_{q=1}^{\infty} q_1^{(4-d)/2}\phi(q) \\ &\lesssim_d k^{-1}\lambda^{d-4}W^{-d+2}\sum_{q=1}^{\infty} q_1^{(6-d)/2}\phi(D) \\ &\lesssim_d k^{-1}\lambda^{d-4}W^{-d+2}\sum_{q_1=1}^{\infty} q_1^{(6-d)/2}\sum_{D|W}\phi(D) \\ &\lesssim_d k^{-1}\lambda^{d-4}W^{-d+3}\sum_{q_1=1}^{\infty} q_1^{(6-d)/2} \\ &\lesssim_d k^{-1}\lambda^{d-4}W^{-d+3} \end{aligned}$$

where the penultimate inequality follows from the general identity

$$\sum_{d|n}\phi(d) = n,$$

and the final inequality follows as $d \ge 9$. In particular

$$S_1 \lesssim_d \begin{cases} \lambda^{d-4}W^{2-d} & \text{if } k = W, \\ \lambda^{d-4}W^{3-d} & \text{if } k = 1. \end{cases}$$

From the estimate in Lemma 3.45:

$$\begin{aligned} S_2 &\lesssim_d \lambda^{d-2+\theta-d/4}W^{-d/2}\sum_{1\le q<\lambda^\theta}\sum_{\substack{0\le a<q,\\(a,q)=1}} q^{-1} \\ &\lesssim_d \lambda^{d-2+\theta-d/4}W^{-d/2}\sum_{1\le q<\lambda^\theta} 1 \\ &\lesssim_d \lambda^{d-2+2\theta-d/4}W^{-d/2}. \end{aligned}$$

We have assumed $d \ge (8 + 8\theta - 4\delta)(1-\delta)^{-1}$ which is equivalent to $d \ge 8 + 8\theta + (d-4)\delta$, which pairs exactly with $W^2 < \lambda^\delta$ to give

$$\lambda^{d-2+2\theta-d/4}W^{-d/2} < \lambda^{d-4}W^{2-d},$$

so

$$S_2 \lesssim_d \lambda^{d-4} W^{2-d}.$$

For S_3, we use Lemma 3.36:

$$\begin{aligned} S_3 &\lesssim_d C_K \lambda^{d-K-2} \log(\lambda) W^{K-d} \sum_{1 \le q < \lambda^\theta} \sum_{\substack{0 \le a < q, \\ (a,q)=1}} q^K \\ &\lesssim_d C_K \lambda^{d-K-1} W^{K-d} \sum_{1 \le q < \lambda^\theta} q^{K+1} \\ &\lesssim_d C_K \lambda^{d-(K+1)(1-\theta)} W^{K-d} \end{aligned}$$

We have assumed $\theta < 1$, so we can choose a sufficiently large $K = K(\theta, \delta)$ to obtain

$$S_3 \lesssim_d \lambda^{d-4} W^{2-d},$$

because we can again use $W^2 < \lambda^\delta$ to make the power savings in λ dominate the extra powers of W.

Adding the estimates for S_1, S_2, and S_3 gives the desired result. □

3.3.3 Theorem on Upper Bounds in a Fixed Residue Class

Theorem 3.47. *Let $\underline{x} \in \mathbb{Z}^d$ be primitive with $c_1\lambda \le |\underline{x}| \le c_2\lambda$. If $d \ge 13$ and $W^{2000} < \lambda$ then*

$$|S_{\underline{x},W,\underline{v}}(\lambda, \mu)| \lesssim_d \begin{cases} \lambda^{d-4} W^{2-d} & \text{for } \underline{x} \in \mathbb{Z}^d \setminus E_W, \\ \lambda^{d-4} W^{3-d} & \text{for } \underline{x} \in E_W, \end{cases}$$

where E_W is the exceptional set

$$E_W = \left\{ \underline{x} \in \mathbb{Z}^d \ : \ \min_{\underline{y} \in \mathbb{Z}^d \setminus \langle \underline{x} \rangle} \mathrm{Dist}(\underline{y}, \langle \underline{x} \rangle) < \frac{W}{|\underline{x}|} \right\}.$$

Proof. Combining minor arc estimate Corollary 3.33 and major arc estimate Theorem 3.46 in equation (3.23) gives

$$\begin{aligned} |S_{\underline{x},W,\underline{v}}(\lambda, \mu)| &\le |I_{\mathfrak{m}^W(\theta)}| + |I_{\mathfrak{M}^W(\theta)}| \\ &\lesssim_{\varepsilon,d} \begin{cases} \lambda^{d - \frac{d(\theta-\delta)}{2} + \varepsilon} + \lambda^{d-4} W^{2-d} & \text{if } \underline{x} \in \mathbb{Z}^d \setminus E_W, \\ \lambda^{d - \frac{d(\theta-\delta)}{2} + \varepsilon} + \lambda^{d-4} W^{3-d} & \text{if } \underline{x} \in E_W, \end{cases} \end{aligned}$$

where for the minor arc estimate we need $W^2 < \lambda^\delta$ and for the major arc estimate we additionally need

$$\lambda^{1-\theta-\delta/2} \geq 16 \qquad \text{and} \qquad d \geq \max\left(9,\, \frac{8+8\theta-4\delta}{1-\delta}\right).$$

To get the claimed bound, we also need to choose $\varepsilon > 0$ such that

$$\lambda^{d-\frac{d(\theta-\delta)}{2}+\varepsilon} \leq \lambda^{d-4} W^{2-d}.$$

A basic calculation using $W^2 < \lambda^\delta$ shows that it is possible to find such an ε if $d > \frac{8-2\delta}{\theta-2\delta}$.

Choosing $\theta = 0.618$ and $\delta = 0.001$ is good enough to allow $d = 13$. The $\lambda^{1-\theta-\delta/2} \geq 16$ requirement is satisfied for all large enough λ. We can of course increase the implied constant so that the estimate is valid for all small λ too. □

Chapter 4

Distance Trees of Triangles in $\mathbb{Z}^d$

4.1 Distance Trees of Triangles

In the model setting of finite fields, our results concerned general k-degenerate graphs given the weak condition that they could be placed in $2k$-general position. The $2k$-general position assumption was useful because when estimating the upper bound in the full space with some vertices fixed, we wanted to reorder the graph, possibly doubling its degeneracy. The situation in $\mathbb{Z}^d$ is much more delicate. The upper bound we obtained applies only to 2-degenerate graphs, as a number of our arguments about primitive vectors are not so easily generalizable to a notion of primitive planes or primitive hyperplanes. Additionally, in the version of the upper bound that uses W in preparation for the W-trick, we made use of the fact that the distance between the centers of our two spheres was on the same scale as the radii of the spheres. This leads us to look for a class of graphs to which these methods can apply.

Distance trees of triangles, also known as 2-trees, provide an answer. A tree of triangles is exactly a graph which can be constructed by starting from a single edge and repeatedly adding a vertex with edges to some pair of adjacent vertices already in the graph. This is the way we shall always use trees of triangles—thinking of them as being constructed by iteratively completing triangles. An alternative way of identifying trees of triangles is that they are those graphs which are composed of edge-connected triangles such that if you define a new graph with a vertex for each triangle with an edge for adjacent triangles, the result is a tree. As we are always adding a pair of edges to an existing edge when building a copy of $\lambda\Gamma$, the radii of the spheres we want and the distance apart of their centers are all on the scale λ. The other crucial property of distance trees of triangles is that we can always reorder the graph to start with any edge and maintain 2-degeneracy, i.e. we can construct all the triangles starting from any edge. This is not in general a property of 2-degenerate graphs.[38]

[38] E.g. K_5 with an edge removed is 2-degenerate, but it is not possible to move the last vertex in the ordering, as all other vertices have degree 3.

The theorems we prove in this section are of the form that if Γ is a distance tree of triangles, then for large enough λ there exist isometric copies of $q\lambda\Gamma$ in any subset of $\mathbb{Z}^d$ having positive upper Banach density. This of course will hold for all subgraphs of trees of triangles—for example trees.

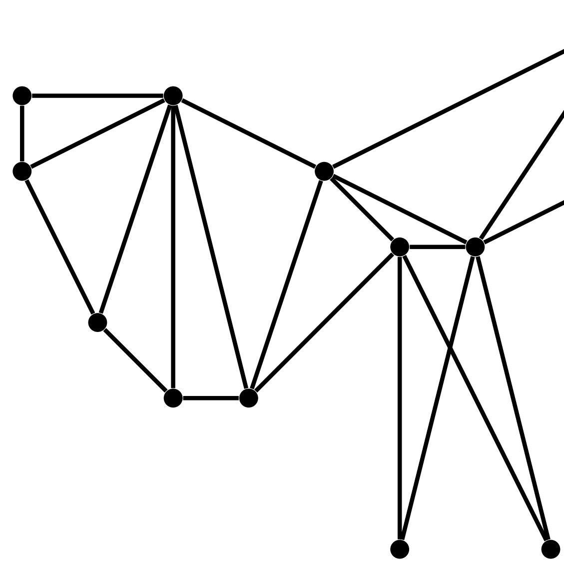

Figure 4.1: A Tree of Triangles.

4.2 The L^2-Estimate

This section takes the problem of finding integral solutions to a system of Diophantine equations into the problems of finding real solutions and p-adic solutions.[39] We then count the number of third points of a triangle that can be chosen from a particular modular class given the other two points and the edge lengths are fixed using a mean-squared estimate.

[39] A slight lie as we care about solutions in $(\mathbb{Z}/p^{\ell}\mathbb{Z})^d$ even if they don't lift to a p-adic solution.

4.2.1 Hensel's Lemma

We give a general version of Hensel's lemma with a goal of lifting solutions to systems of r equations in d variables. Our main reference is Greenberg [Gre69],[40] though useful treatments are also given in [Cas18],[Cona], and [Conb].

[40] Though the version of Hensel's lemma will differ from Greenberg because our goal is to count the number of solutions at each level.

In this section we work with respect to a fixed prime p. To limit notational clutter, we use the shorthand $\equiv_n$ to mean $\equiv \pmod{p^n}$ as in [Cas18] and the notation $[N]$ to mean the set $\{0, 1, \ldots, N-1\} \subset \mathbb{Z}$.

Lemma 4.1 (Hensel). *Let $\mathcal{F}(\underline{x})$ be a system of r integral polynomials in d variables. Let $\underline{x}_n \in [p^n]^d$ and suppose $\mathrm{Jac}_{\mathcal{F}}(\underline{x}_n) = p^m M$ where M has full rank modulo p. If $n \geq m+1$ and $\mathcal{F}(\underline{x}_n) \equiv_{n+m} \underline{0}$, then there are exactly p^{d-r} choices of $\underline{x}_{n+1} \in [p^{n+1}]^d$ with $\mathcal{F}(\underline{x}_{n+1}) \equiv_{n+m+1} \underline{0}$ and $\underline{x}_{n+1} \equiv_n \underline{x}_n$. Moreover, $\mathrm{Jac}_{\mathcal{F}}(\underline{x}_{n+1}) = p^m M'$ where M' has full rank modulo p.*

Proof. The proof relies on the following two elementary polynomial identities which follow from binomial expansion.

If $f : \mathbb{R}^d \to \mathbb{R}$ is a polynomial with integer coefficients, then

$$f(\underline{x} + c\underline{y}) = f(\underline{x}) + c\nabla f(\underline{x}) \cdot \underline{y} + c^2 g(\underline{x}, \underline{y}), \tag{4.1}$$

where g is a polynomial with integer coefficients. If $\mathcal{F} : \mathbb{R}^d \to \mathbb{R}^r$ is a system of polynomials with integer coefficients, then

$$\mathcal{F}(\underline{x} + c\underline{y}) = \mathcal{F}(\underline{x}) + c\mathrm{Jac}_{\mathcal{F}}(\underline{x})\underline{y} + c^2 \mathcal{G}(\underline{x}, \underline{y}), \tag{4.2}$$

where $\mathcal{G}$ is a system of polynomials with integer coefficients.

We are searching for solutions of the form $\underline{x}_{n+1} \equiv_n \underline{x}_n$. We can write all such $\underline{x}_{n+1}$ as $\underline{x}_n + p^n \underline{h}$ with $\underline{h} \in [p]^d$. Defining $\underline{t} \in \mathbb{Z}^r$ with $\mathcal{F}(\underline{x}_n) = p^{n+m}\underline{t}$ and recalling we have assumed $2n \geq n+m+1$, an application of equation (4.2) shows

$$\mathcal{F}(\underline{x}_n + p^n \underline{h}) \equiv_{n+m+1} \mathcal{F}(\underline{x}_n) + p^n \mathrm{Jac}_{\mathcal{F}}(\underline{x}_n)\underline{h} \equiv_{n+m+1} p^{n+m}\left(\underline{t} + M\underline{h}\right).$$

As M is full rank modulo p by assumption, there are exactly p^{d-r} choices of $\underline{h} \in [p]^d$ such that $\underline{t} + M\underline{h} \equiv_1 \underline{0}$.

Finally, applying equation (4.1) with such a choice of $\underline{h}$ to $\partial f_j / \partial x_i$ shows

$$\begin{aligned}\frac{\partial f_j}{\partial x_i}(\underline{x}_n + p^n \underline{h}) &\equiv_{m+1} \frac{\partial f_j}{\partial x_i}(\underline{x}_n) + p^n \nabla \frac{\partial f_j}{\partial x_i}(\underline{x}_n) \cdot \underline{h} + p^{2n} g(\underline{x}_n, \underline{h}) \\ &\equiv_{m+1} \frac{\partial f_j}{\partial x_i}(\underline{x}_n),\end{aligned}$$

as $n \geq m+1$.

Hence $\mathrm{Jac}_{\mathcal{F}}(\underline{x}_{n+1}) \equiv_{m+1} \mathrm{Jac}_{\mathcal{F}}(\underline{x}_n)$, so $\mathrm{Jac}_{\mathcal{F}}(\underline{x}_{n+1}) = p^m M'$ where M' is full rank modulo p. □

Corollary 4.2. *Let $\mathcal{F}(\underline{x})$ be a system of r integral polynomials in d variables. Let $\underline{x}_n \in [p^n]^d$ and suppose $\mathrm{Jac}_{\mathcal{F}}(\underline{x}_n) = p^m M$ where M has full rank modulo p. If $n \geq m+1$ and $\mathcal{F}(\underline{x}_n) \equiv_{n+m} \underline{0}$, then for $K \geq n+m$*

$$\#\{\underline{x} \in [p^K]^d \,:\, \mathcal{F}(\underline{x}) \equiv_K \underline{0},\ \underline{x} \equiv_n \underline{x}_n\} = p^{(K-n)(d-r)+md}.$$

Proof. Suppose that $\underline{x}_N \in [p^N]^d$ with $\mathcal{F}(\underline{x}_N) \equiv_N 0$, $\mathrm{Jac}_{\mathcal{F}}(\underline{x}_N) = p^m M$ where M has full rank modulo p, and $N \geq m+1$. We now consider all elements $\underline{x} \in [p^{N+m}]^d$ with $\underline{x} \equiv_N \underline{x}_N$, i.e. those elements of the form $\underline{x} = \underline{x}_N + p^N \underline{y}$ with $\underline{y} \in [p^m]^d$. From equation (4.2) we have

$$\mathcal{F}(\underline{x}_N + p^N \underline{y}) \equiv_{N+m} \mathcal{F}(\underline{x}_N). \tag{4.3}$$

Observe that setting $N = n$ in equation (4.3) implies there are p^{md} solutions to $\mathcal{F}(\underline{x}) \equiv_{n+m} \underline{0}$ with $\underline{x} \equiv_n \underline{x}_n$.[41]

Under the condition $\mathcal{F}(\underline{x}_n) \equiv_{n+m} \underline{0}$, we can use Lemma 4.1 iteratively to get exactly $p^{(K-n)(d-r)}$ choices $\underline{x}_{K-m} \in [p^{K-m}]^d$ with $\underline{x}_{K-m} \equiv_n \underline{x}_n$ and $\mathcal{F}(\underline{x}_{K-m}) \equiv_K \underline{0}$. Plugging $N = K-m$ into equation (4.3) shows that there are therefore exactly p^{md} solutions in $[p^K]^d$ for each choice of $\underline{x}_{K-m}$. In total this gives

$$\#\{\underline{x} \in [p^K]^d \,:\, \mathcal{F}(\underline{x}) \equiv_K \underline{0},\ \underline{x} \equiv_n \underline{x}_n\} = p^{(K-n)(d-r)+md}.$$

□

[41] Equation (4.3) also shows if $\mathcal{F}(\underline{x}_n) \equiv_n \underline{0}$ but $\mathcal{F}(\underline{x}_n) \not\equiv_{n+m} \underline{0}$, then there are no solutions to $\mathcal{F}(\underline{x}) \equiv_{n+m} \underline{0}$ with $\underline{x} \equiv_n \underline{x}_n$. If there are no solutions modulo p^{n+m} then of course there cannot be solutions modulo p^K for $K > n+m$.

4.2.2 Systems with Modular Restrictions

We will use results stemming from the seminal work of Birch [Bir62], in particular those developed in [MT16]. We will apply the so-called W-trick of Green and Tao [GT08] and count solutions to systems of equations just among a fixed residue class modulo W. The W-trick is used to eliminate issues arising from small primes.

We shall always consider the W to take the specific form

$$W := 2\prod_{p<w} p,$$

where w is a slow-growing function to be chosen later.

Definition 4.3. *Let $N \in \mathbb{N}$, and let $\mathcal{F}$ be a system of forms in d variables.*

$$\mathcal{R}_N(W, \underline{s}, \underline{v}) := \#\{\underline{x} \in [N]^d \,:\, \underline{x} \equiv \underline{s} \pmod{W},\ \mathcal{F}(\underline{x}) = \underline{v}\}.$$

Definition 4.4. *If $\mathcal{F}$ is a system of forms in* d *variables, we set*

$$\mathrm{Rank}(\mathcal{F}) := \mathrm{codim}(V_{\mathcal{F}}^*),$$

where $V_{\mathcal{F}}^*$ *is the singular variety containing those* $\underline{z} \in \mathbb{C}^d$ *where* $\mathrm{Jac}_{\mathcal{F}}(\underline{z})$ *drops rank.*

Definition 4.5. *Let $\mathcal{F}$ be a system of r forms in* d *variables and let* $\underline{t} \in \mathbb{R}^r$. *We define* $\omega_{\mathcal{F}}(\underline{t})$ *to be the Gelfand-Leray measure*[42] *of the set* $\{\underline{x} \in \mathbb{R}^d \,:\, \mathcal{F}(\underline{x}) = \underline{t}\}$.

[42] See section 10 of [CLM21] for a description of the Gelfand-Leray measure. It is the viewpoint we will take on the singular integral of Birch [Bir62].

A particular case of Proposition 1.1 in [MT16] is the following fact which we cite below.

Proposition 4.6. *If $\mathcal{F}$ is a system of r homogeneous quadratic forms satisfying*

$$\mathrm{Rank}(\mathcal{F}) > 2r(r+1),$$

then there exist $\delta = \delta(r) > 0$ *and* $\beta = \beta(r) > 0$ *such that if* $W \leq N^{\beta}$, *then*

$$\mathcal{R}_N(W, \underline{s}, \underline{v}) = N^{d-2r} W^{-d} \omega_{\mathcal{F}}\left(\frac{\underline{v}}{N^2}\right) \prod_p \sigma_p(W, \underline{s}, \underline{v}) + O(N^{d-2r-\delta}),$$

where

$$\sigma_p(W, \underline{s}, \underline{v}) = \lim_{\ell \to \infty} \sigma_p^{\ell}(W, \underline{s}, \underline{v}),$$

$$\sigma_p^{\ell}(W, \underline{s}, \underline{v}) = \frac{\#\{\underline{x} \in [p^{\ell}]^d \,:\, \mathcal{F}(W\underline{x} + \underline{s}) \equiv \underline{v}\ (p^{\ell})\}}{p^{\ell(d-r)}}.$$

Additionally a mild adaptation of Lemma 2.1 in [MT16] gives the following identity for W and p such that $p^m \mid\mid W$:

$$\sigma_p(W, \underline{s}, \underline{v}) = \begin{cases} \sigma_p(\underline{v}) & \text{if } m = 0, \\ \sigma_p(p^m, \underline{s}, \underline{v}) & \text{if } m \geq 1. \end{cases} \tag{4.4}$$

Lemma 4.7. *Let $\mathcal{F}$ be a system of r quadratic forms in* d *variables. Suppose that* $\underline{s} \in [p]^d$ *and* $\mathrm{Jac}_{\mathcal{F}}(\underline{s})$ *has full rank modulo p. If* $\mathcal{F}(\underline{s}) \equiv \underline{v} \pmod{p}$, *then*

$$\sigma_p(p, \underline{s}, \underline{v}) = p^r.$$

Proof. Corollary 4.2 with $n = 1$ and $m = 0$ gives for $\ell \geq 1$

$$\#\{\underline{y} \in [p^{\ell}]^d \,:\, \mathcal{F}(\underline{y}) \equiv \underline{v}\ (p^{\ell}),\ \underline{y} \equiv \underline{s}\ (p)\} = p^{(\ell-1)(d-r)}.$$

For each such $\underline{y}$, there are p^d options $\underline{x} \in [p^\ell]^d$ such that $p\underline{x} + \underline{s} = \underline{y}$. Hence

$$\sigma_p^\ell(p, \underline{s}, \underline{v}) = \frac{p^d p^{(\ell-1)(d-r)}}{p^{\ell(d-r)}} = p^r,$$

so $\sigma_p(p, \underline{s}, \underline{v}) = p^r$. □

Lemma 4.7 will be sufficient for our purposes whenever $p \neq 2$. However, because we want to address a system of quadratic forms, every solution modulo 2 will be singular.[43] The following lemma is adapted for this case.

[43] That is, 2 will always factor out of the Jacobian matrix.

Lemma 4.8. *Let $\mathcal{F}$ be a system of r quadratic forms in d variables. Suppose $\underline{s} \in [4]^d$ and $\mathrm{Jac}_{\mathcal{F}}(\underline{s}) = 2M$ where M has full rank modulo 2. If $\mathcal{F}(\underline{s}) \equiv \underline{v} \pmod 8$, then*

$$\sigma_2(4, \underline{s}, \underline{v}) = 2^d 4^r.$$

Proof. By Corollary 4.2 with $n = 2$ and $m = 1$, we have that for $\ell \geq 3$

$$\#\{\underline{y} \in [2^\ell]^d \ : \ \mathcal{F}(\underline{y}) \equiv \underline{v} \ (\mathrm{mod}\ 2^\ell), \ \underline{y} \equiv \underline{s} \ (\mathrm{mod}\ 4)\} = 2^{(\ell-2)(d-r)+d}.$$

As in the proof of the previous lemma, for each such $\underline{y}$, there are 2^{2d} options $\underline{x} \in [2^\ell]^d$ such that $4\underline{x} + \underline{s} = \underline{y}$. Hence

$$\sigma_2^\ell(4, \underline{s}, \underline{v}) = \frac{2^{2d} 2^{(\ell-2)(d-r)+d}}{2^{\ell(d-r)}} = 2^d 2^{2r},$$

so $\sigma_2(4, \underline{s}, \underline{v}) = 2^d 4^r$. □

Corollary 4.9. *Let $\mathcal{F}$ be a system of r quadratic forms in d variables and set $W = 2 \prod_{p<w} p$. Suppose $\underline{s}$ is a nonsingular solution to $\mathcal{F}(\underline{s}) \equiv \underline{v} \ (\mathrm{mod}\ p)$ for $2 < p < w$, and $\mathcal{F}(\underline{s}) \equiv \underline{v} \ (\mathrm{mod}\ 8)$ with $\mathrm{Jac}_{\mathcal{F}}(\underline{s}) = 2M$ where M has full rank modulo 2. Then*

$$\prod_{p<w} \sigma_p(W, \underline{s}, \underline{v}) = 2^d W^r.$$

Proof. The is direct computation from equation (4.4) and Lemmas 4.7 - 4.8. □

Lemma 4.10. *If $W = 2 \prod_{p<w} p$ and $\mathcal{F}$ is a system of r quadratic forms in d variables, then*

$$\prod_{p \geq w} \sigma_p(W, \underline{s}, \underline{v}) = 1 + o_{w \to \infty}(1).$$

Proof. This is the note following Equation (2.30) in [MT16]. The idea is that $\sigma_p(\underline{v}) = 1 + O(p^{-2})$ and of course the tail of $\sum p^{-2}$ goes to 0. $\square$

Lemma 4.11. *Let $\mathcal{F}$ be a system of r homogeneous quadratic forms in d variables satisfying the rank condition of Proposition 4.6. Set $W = 2\prod_{p<w} p$ and let $\underline{s}$ satisfy the conditions of Corollary 4.9. Let $\lambda^2 \in \mathbb{N}$ and let $\underline{t}^* \in \mathbb{N}^r$. If $W < \lambda^\beta$ where the β is that appearing in Proposition 4.6, then*

$$\#\{\underline{x} \in \mathbb{Z}^d \ : \ \mathcal{F}(\underline{x}) = \lambda^2 \underline{t}^*,\ \underline{x} \equiv \underline{s}\,(W)\} = 2^d \lambda^{d-2r} W^{r-d} \omega_{\mathcal{F}}(\underline{t}^*)(1+\mathcal{E}),$$

where $\mathcal{E} = o_{w\to\infty}(1) + o_{\lambda\to\infty}(1)$.

Proof. This is an application of Proposition 4.6 with $N = \lambda^2$ and $\underline{v} = \lambda^2 \underline{t}^*$, along with an application of Corollary 4.9 and Lemma 4.10 to obtain

$$\prod_p \sigma_p(W, \underline{s}, \underline{v}) = 2^d W^r (1 + o_{w\to\infty}(1)).$$

$\square$

Definition 4.12. *We define the systems*

$$\mathcal{F}_0(\underline{x}) := (|\underline{x}|^2),$$

$$\mathcal{F}_1(\underline{x}, \underline{y}) := (|\underline{y}|^2, |\underline{x} - \underline{y}|^2),$$

$$\mathcal{F}_2(\underline{x}, \underline{y}) := (|\underline{x}|^2, |\underline{y}|^2, |\underline{x} - \underline{y}|^2),$$

$$\mathcal{F}_3(\underline{x}, \underline{y}, \underline{z}) := (|\underline{x}|^2, |\underline{y}|^2, |\underline{x} - \underline{y}|^2, |\underline{z}|^2, |\underline{x} - \underline{z}|^2),$$

where $\underline{x}, \underline{y}, \underline{z} \in \mathbb{Z}^d$.

The system we are really interested in is $\mathcal{F}_1$ as it will correspond to counting points given the fixed edge $\underline{x}$. The other systems are more regular and can be addressed via Lemma 4.11.

Remark 4.13. The rank condition in Proposition 4.6 for $\mathcal{F}_0$, $\mathcal{F}_2$, and $\mathcal{F}_3$ in fact boils down to $d > r(r+1)$, so we need $d > 60$ to address $\mathcal{F}_3$. However, the later work of S.L. Rydin Myerson [Mye18] only requires $\mathrm{Rank}(\mathcal{F}) > 8r$ for quadratic forms, so by that work $d > 40$ is good enough to address $\mathcal{F}_3$. This will be the dominant restriction on the dimension we require.[44] Additionally, the $\beta = \beta(r)$ from Proposition 4.6 appearing in Lemma 4.11 is explicit in [MT16], and can be computed as $1/97$, $1/769$, and $1/2881$ for $\mathcal{F}_0$, $\mathcal{F}_2$, and $\mathcal{F}_3$ respectively. Whenever we need to, we will capture this condition with $W^{3000} < \lambda$.

[44] The community expectation is that the asymptotic formula here actually works for $\mathcal{F}_3$ when $d > 10$.

Remark 4.14. Part of the conditions in Corollary 4.9 are that $\underline{s}$ is nonsingular for $p \neq 2$ and $\mathrm{Jac}_{\mathcal{F}}(\underline{s}) = 2M$ with M full rank modulo 2. For $\mathcal{F}_0$, this is satisfied if $\underline{s} \in [W]^d$ is primitive. For $\mathcal{F}_2$ with $\underline{s} = (\underline{u}, \underline{v})$ and for $\mathcal{F}_3$ with $\underline{s} = (\underline{u}, \underline{v}, \underline{v})$ where $\underline{u}, \underline{v} \in [W^d]$, the conditions are met exactly when $\underline{u}$, $\underline{v}$, and $\underline{v} - \underline{u}$ are primitive. The other condition we need is that $\underline{s}$ is a solution to $\mathcal{F}(\underline{s}) \equiv \underline{t} \pmod 8$ and $\mathcal{F}(\underline{s}) \equiv \underline{t} \pmod p$ for $3 \leq p < w$. We capture these conditions with the assumption that $\mathcal{F}(\underline{s}) \equiv \underline{v} \pmod{2W}$, noting that $W := 2\prod_{p<w} W$, so $8 \mid 2W$.

Definition 4.15. *Let $\underline{t} \in \mathbb{Z}^2$ and let $\underline{x} \in \mathbb{Z}^d$.*

$$\mathcal{N}_{\lambda,\underline{t},W,\underline{v}}(\underline{x}) := \#\{\underline{y} \in \mathbb{Z}^d \ : \ \mathcal{F}_1(\underline{x}, \underline{y}) = \lambda^2 \underline{t},\ \underline{y} \equiv \underline{v} \pmod W\}.$$

We reiterate the point that $\mathcal{N}_{\lambda,\underline{t},W,\underline{v}}(\underline{x})$ cannot be directly computed via Lemma 4.11 as $\underline{x}$ is fixed, and if $\mathcal{F}_1(\underline{x}, \underline{y})$ is viewed as a system of just the variable $\underline{y}$, then it fails to be homogeneous. Instead, we show $\mathcal{N}_{\lambda,\underline{t},W,\underline{v}}(\underline{x})$ is as big as expected most of the time with a mean squared error estimate.

Definition 4.16. *Let $t_0 \in \mathbb{Z}$ and let $\underline{t} \in \mathbb{Z}^2$. We define E to be sum of square differences*

$$E := \sum_{\substack{|\underline{x}|^2 = \lambda^2 t_0, \\ \underline{x} \equiv \underline{u}\,(W)}} |\mathcal{N}_{\lambda,\underline{t},W,\underline{v}}(\underline{x}) - 2^d \lambda^{d-4} W^{2-d} \omega_{\mathcal{F}_1}(\underline{t})|^2.$$

The expression E can be understood by expanding into three sums and then recognizing the result as counts of graphs, the most complicated of which is a diamond graph. Our final dimensional constraint will come from the dimension needed to properly count such graphs.

Lemma 4.17. *Let $t_0 \in \mathbb{Z}$, $\underline{t} \in \mathbb{Z}^2$ and set $W = 2\prod\limits_{p<w} p$. Let $\underline{u}$, $\underline{v} \in [W]^d$ with $\underline{u}$, $\underline{v}$, and $\underline{v} - \underline{u}$ primitive. Suppose $d > 40$ and $\mathcal{F}_2(\underline{u}, \underline{v}) \equiv \lambda^2(t_0, \underline{t}) \pmod{2W}$. Then when $W < \lambda^{3000}$,*

$$E = 2^{3d} \lambda^{3d-10} W^{5-3d} (o_{w\to\infty}(1) + o_{\lambda\to\infty}(1)).$$

Proof. We first note that by Remarks 4.13 and 4.14, we may apply Lemma 4.11.

We expand the sum into the three terms

$$S_1 = \sum_{\substack{|\underline{x}|^2 = \lambda^2 t_0, \\ \underline{x} \equiv \underline{u}\,(W)}} \mathcal{N}^2_{\lambda,\underline{t},W,\underline{v}}(\underline{x}),$$

$$S_2 = -2^{d+1}\lambda^{d-4}W^{2-d}\omega_{\mathcal{F}_1}(\underline{t}) \sum_{\substack{|\underline{x}|^2=\lambda^2 t_0,\\ \underline{x}\equiv\underline{u}\,(W)}} \mathcal{N}_{\lambda,\underline{t},W,\underline{v}}(\underline{x}),$$

$$S_3 = 2^{2d}\lambda^{2d-8}W^{4-2d}\omega^2_{\mathcal{F}_1}(\underline{t}) \sum_{\substack{|\underline{x}|^2=\lambda^2 t_0,\\ \underline{x}\equiv\underline{u}\,(W)}} 1.$$

We first examine S_3. By Lemma 4.11 with $r = 1$ and d variables, we have

$$\sum_{\substack{|\underline{x}|^2=\lambda^2 t_0,\\ \underline{x}\equiv\underline{u}\,(W)}} 1 = 2^d\lambda^{d-2}W^{1-d}\omega_{\mathcal{F}_0}(t_0)\left(1+\mathcal{E}_3\right),$$

where $\mathcal{E}_3 = o_{w\to\infty}(1) + o_{\lambda\to\infty}(1)$. so

$$S_3 = 2^{3d}\lambda^{3d-10}W^{5-3d}\omega_{\mathcal{F}_1}(\underline{t})^2\omega_{\mathcal{F}_0}(t_0)\left(1+\mathcal{E}_3\right).$$

Turning to S_2, by Lemma 4.11 with $r = 3$ and $2d$ variables, we have

$$\sum_{\substack{|\underline{x}|^2=\lambda^2 t_0,\\ \underline{x}\equiv\underline{u}\,(W)}} \mathcal{N}_{\lambda,\underline{t},W,\underline{v}}(\underline{x}) = 2^{2d}\lambda^{2d-6}W^{3-2d}\omega_{\mathcal{F}_2}(t_0,\underline{t})\left(1+\mathcal{E}_2\right),$$

with $\mathcal{E}_2 = o_{w\to\infty}(1) + o_{\lambda\to\infty}(1)$. Thus

$$S_2 = -2^{3d+1}\lambda^{3d-10}W^{5-3d}\omega_{\mathcal{F}_1}(\underline{t})\omega_{\mathcal{F}_2}(t_0,\underline{t})\left(1+\mathcal{E}_2\right).$$

Finally, by Lemma 4.11 with $r = 5$ and $3d$ variables

$$S_1 = 2^{3d}\lambda^{d-10}W^{5-d}\omega_{\mathcal{F}_3}(\underline{t}_0,\underline{t},\underline{t})\left(1+\mathcal{E}_1\right),$$

with $\mathcal{E}_1 = o_{w\to\infty}(1) + o_{\lambda\to\infty}(1)$.

To prove the lemma, it remains to show that the main terms of these sums cancel, i.e. we need to show

$$\omega_{\mathcal{F}_3}(t_0,\underline{t},\underline{t}) - 2\omega_{\mathcal{F}_1}(\underline{t})\omega_{\mathcal{F}_2}(t_0,\underline{t}) + \omega_{\mathcal{F}_1}(\underline{t})^2\omega_{\mathcal{F}_0}(t_0) = 0.$$

In fact, by Fubini's theorem indicating the variables, Equation (48) in [CLM21], we have

$$\omega_{\mathcal{F}_3}(t_0,\underline{t},\underline{t}) = \omega_{\mathcal{F}_1}(\underline{t})\omega_{\mathcal{F}_2}(t_0,\underline{t}) = \omega_{\mathcal{F}_1}(\underline{t})^2\omega_{\mathcal{F}_0}(t_0).$$

□

We now turn to counting the number of solutions when $\underline{y} \in A \subset \mathbb{Z}^d$. We fix $\underline{x}_1, \underline{x}_2 \in \mathbb{Z}^d$ and consider the number of points in A which are a specified distance away.

Definition 4.18. *Let $A \subset \mathbb{Z}^d$ and $\underline{t} \in \mathbb{Z}^2$. For $X = (\underline{x}_1, \underline{x}_2)$ with $\underline{x}_1, \underline{x}_2 \in \mathbb{Z}^d$ define*

$$\mathcal{N}^A_{\lambda,\underline{t},W,\underline{v}}(X) := \#\{\underline{y} \in A \,:\, \mathcal{F}_1(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1) = \lambda^2 \underline{t},\ \underline{y} \equiv \underline{v} \,(\mathrm{mod}\, W)\}.$$

Definition 4.19. *Let $A \subset Q_N$ with density $\alpha := \frac{|A|}{|Q_N|}$, and let $\underline{t} \in \mathbb{Z}^2$.*

$$E^A_{W,\underline{v}} := \sum_{\substack{\underline{x}_1, \underline{x}_2 \in Q_N, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \ (\mathrm{mod}\ W), \\ \underline{x}_2 \equiv \underline{u}_2 \ (\mathrm{mod}\ W)}} |\mathcal{N}^A_{\lambda,\underline{t},W,\underline{v}}(X) - 2^d \alpha \lambda^{d-4} W^{2-d} \omega_{\mathcal{F}_1}(\underline{t})|^2,$$

where $X = (\underline{x}_1, \underline{x}_2)$.

Note that this includes a sum over Q_N that was not in definition of E. This is because in E the basepoint $\underline{x}_1$ was irrelevant. In contrast, here we need to count only points in A and the location of the basepoint matters. However, summing over Q_N allows us to take advantage of uniformity in the distribution of A. We now recall the definition of the $U^1_{m,L}(Q)$ norm from Chapter 1 and as in [LM22]. [45]

For a lattice cube $Q \subset \mathbb{Z}^d$ and a function $f \,:\, Q \to [-1, 1]$

$$\|f\|_{U^1_{m,L}(Q)} := \left(\frac{1}{|Q|} \sum_{\underline{t} \in \mathbb{Z}^d} |f * \chi_{m\mathcal{C}_L}(\underline{t})|^2 \right)^{1/2},$$

where $\chi_{m\mathcal{C}_L}$ is the normalized characteristic function of the m-spaced cube with L lattice points on each side $m\mathcal{C}_L = m\{1, \ldots, L\}^d$.

Lemma 4.20. *Let $A \subset Q_N \subset \mathbb{Z}^d$ with density $\alpha := \frac{|A|}{|Q_N|}$. Let $t_0 \in \mathbb{Z}$, $\underline{t} \in \mathbb{Z}^2$ and set $W = 2 \prod_{p<w} p$. Let $\underline{u}_1, \underline{u}_2, \underline{v} \in [W]^d$ with $\underline{u}_2 - \underline{u}_1$, $\underline{v} - \underline{u}_1$, and $\underline{v} - \underline{u}_2$ primitive. Suppose $d > 40$ and*

$$\mathcal{F}_2(\underline{u}_2 - \underline{u}_1, \underline{v} - \underline{u}_1) \equiv \lambda^2 (t_0, \underline{t}) \,(\mathrm{mod}\, 2W).$$

Fix $0 < \eta \ll 1$ and set $m := \mathrm{lcm}\{1 \le n \le C_{t,d}\eta^{-10}\}$ where $C_{t,d}$ is a constant depending on t_0, $\underline{t}$, and d.[46] Let $1 \ll L \ll \eta^{10}\lambda$ and assume A is $(W^{-d/2}\eta, L)$-uniformly distributed with respect to Wm. If $W^{3000} \le \lambda \ll N$,

[45] Our version is different in an unimportant way. Namely, the spacing parameter m is the same, but the window size parameter L as we define it is scaled by m from the version in [LM22].

[46] The particular value of $C_{t,d}$ is unimportant, and it can be chosen large enough to suit the needs of the proof.

then

$$E^A_{W,\underline{v}} \le 2^{4d}\alpha^2 N^d \lambda^{3d-10} W^{5-4d}\mathcal{E}$$

with

$$\mathcal{E} = o_{w\to\infty}(1) + o_{\lambda\to\infty}(1) + o_{N/\lambda^3\to\infty}(1) + o_{\eta\to 0}(1).$$

Proof. Let $\Lambda := W\mathbb{Z}^d + \underline{v}$, and let

$$1_{\mathcal{F},\underline{t}}(\underline{x}) := \begin{cases} 1 & \text{if } \mathcal{F}(\underline{x}) = \underline{t}, \\ 0 & \text{otherwise.} \end{cases}$$

Let f_A be the balance function of A, i.e. $f_A := 1_A - \alpha 1_{Q_N}$. Then

$$\begin{aligned} \mathcal{N}^A_{\lambda,\underline{t},W,\underline{v}}(X) &= \sum_{\underline{y}\in\Lambda} 1_A(\underline{y}) 1_{\mathcal{F}_1,\lambda^2\underline{t}}(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1) \\ &= \sum_{\underline{y}\in\Lambda} (\alpha 1_{Q_N}(\underline{y}) + f_A(\underline{y})) 1_{\mathcal{F}_1,\lambda^2\underline{t}}(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1) \\ &= \alpha \sum_{\underline{y}\in Q_N\cap\Lambda} 1_{\mathcal{F}_1,\lambda^2\underline{t}}(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1) + F(X) \\ &= \alpha \mathcal{N}_{\lambda,\underline{t},W,\underline{v}-\underline{u}_1}(\underline{x}_2 - \underline{x}_1) + \mathcal{E}(X) + F(X) \end{aligned} \tag{4.5}$$

where

$$F(X) := \sum_{\underline{y}\in\Lambda} f_A(\underline{y}) 1_{\mathcal{F}_1,\lambda^2\underline{t}}(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1),$$

and $\mathcal{E}(X)$ is capturing the error from the undercount which happens in the first term if $\underline{x}_1$ or $\underline{x}_2$ is too close to the boundary of Q_N. To handle this error, we set $N_1 := N - \lambda^2(t_0 + t_1 + t_2)$ and note that if $\underline{x}_1 \in Q_{N_1}$ then there are no boundary issues and $\mathcal{E}(X) = 0$, and in fact we can drop the condition $\underline{x}_2 \in Q_N$ as this is already forced by $|\underline{x}_2 - \underline{x}_1| = \lambda^2 t_0$ for $\underline{x}_1 \in Q_{N_1}$.

We write

$$E^A_{W,\underline{v}} = \mathcal{E} + \sum_{\substack{\underline{x}_1\in Q_{N_1}, \underline{x}_2\in\mathbb{Z}^d, \\ |\underline{x}_2-\underline{x}_1|^2=\lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod{W}, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod{W}}} |\mathcal{N}^A_{\lambda,\underline{t},W,\underline{v}}(X) - 2^d\alpha\lambda^{d-4}W^{2-d}\omega_{\mathcal{F}_1}(\underline{t})|^2,$$

where $\mathcal{E}$ is the error from restricting the sum to $\underline{x}_1 \in Q_{N_1}$.

For the sum, we use the same computation as in Lemma 4.17, insert equation (4.5) with $\mathcal{E}(X) = 0$, and examine the resulting error terms. We expand into the three terms

$$S_1 = \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \underline{x}_2 \in \mathbb{Z}^d, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod W, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod W}} \left[\alpha \mathcal{N}_{\lambda, \underline{t}, W, \underline{v} - \underline{u}_1}(\underline{x}_2 - \underline{x}_1) + F(X)\right]^2,$$

$$S_2 = -\, 2^{d+1} \alpha \lambda^{d-4} W^{2-d} \omega_{\mathcal{F}_1}(\underline{t}) \cdots$$
$$\cdots \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \underline{x}_2 \in \mathbb{Z}^d, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod W, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod W}} \left[\alpha \mathcal{N}_{\lambda, \underline{t}, W, \underline{v} - \underline{u}_1}(\underline{x}_2 - \underline{x}_1) + F(X)\right],$$

$$S_3 = 2^{2d} \alpha^2 \lambda^{2d-8} W^{4-2d} \omega_{\mathcal{F}_1}^2(\underline{t}) \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \underline{x}_2 \in \mathbb{Z}^d, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod W, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod W}} 1.$$

For a fixed $\underline{x}_1$, the main terms are the main terms from Lemma 4.17 with an extra factor of α^2, and so cancel up to the error term from that lemma. Up to an $O_d(N^d - N_1^d)$ error, there are $(2N)^d W^{-d}$ choices for $\underline{x}_1 \in Q_{N_1}$ satisfying $\underline{x}_1 \equiv \underline{u}_1 \pmod W$. This leaves what will be an error term from S_1 and S_2, i.e.

$$E^A_{W,\underline{v}} = 2^{4d} \alpha^2 N^d \lambda^{3d-10} W^{5-4d} (o_{w \to \infty}(1) + o_{\lambda \to \infty}(1)) + \mathcal{E} + \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2, \tag{4.6}$$

where $\mathcal{E}$ as described earlier is the error from using $\underline{x}_1 \in Q_{N_1}$

$$\mathcal{E} := \sum_{\substack{\underline{x}_1 \in Q_N \setminus Q_{N_1}, \underline{x}_2 \in Q_N, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod W, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod W}} |\mathcal{N}^A_{\lambda, \underline{t}, W, \underline{v}}(X) - 2^d \alpha \lambda^{d-4} W^{2-d} \omega_{\mathcal{F}_1}(\underline{t})|^2,$$

$\mathcal{E}_0$ is the error from overcounting the number of $\underline{x}_1 \equiv \underline{u}_1$ in Q_{N_1}

$$\mathcal{E}_0 = \lambda^{3d-10} O_d(N^d - N_1^d),$$

$\mathcal{E}_1$ is the non-main term after expanding S_1

$$\mathcal{E}_1 := \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \underline{x}_2 \in \mathbb{Z}^d, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod{W}, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod{W}}} \left[2\alpha \mathcal{N}_{\lambda, \underline{t}, W, \underline{v} - \underline{u}_1}(\underline{x}_2 - \underline{x}_1) + F(X)\right] F(X),$$

and $\mathcal{E}_2$ is the non-main term after expanding S_2

$$\mathcal{E}_2 := -2^{d+1}\alpha \lambda^{d-4} W^{2-d} \omega_{\mathcal{F}_1}(\underline{t}) \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \underline{x}_2 \in \mathbb{Z}^d, \\ |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0, \\ \underline{x}_1 \equiv \underline{u}_1 \pmod{W}, \\ \underline{x}_2 \equiv \underline{u}_2 \pmod{W}}} F(X).$$

Let the shorthand t denote dependence on t_0 and $\underline{t}$. For $\mathcal{E}$ and $\mathcal{E}_0$, the savings in N will subsume any other considerations. By Theorem 3.29 we have the wasteful estimate $\mathcal{N}^A_{\lambda, \underline{t}, W, \underline{v}}(X) \lesssim_{t,d} \lambda^{d-4}$.[47] Note $(\lambda^{d-4})^2 \leq \lambda^{3d-10}$. Hence

[47] Under the condition $\underline{x}_1 \neq \underline{x}_2$ which is forced by $t_0 > 0$.

$$\mathcal{E} + \mathcal{E}_0 \lesssim_{t,d} (N^d - N_1^d)\lambda^{3d-10} \lesssim_{t,d} N^d \lambda^{3d-10}(\lambda^2/N), \qquad (4.7)$$

where we have recalled the definition $N_1 = N - \lambda^2(t_0 + t_1 + t_2)$. We record the contribution of this error with $o_{N/\lambda^3 \to \infty}(1)$ where we use the extra power of λ to cover for $2^{4d}\alpha^2 W^{5-4d}$ in the main term.

We want to find upper bounds on $|\mathcal{E}_1|$ and $|\mathcal{E}_2|$. To do this, we pull in the absolute value and drop the modular conditions on $\underline{x}_1$ and $\underline{x}_2$.[48] We will expand out the sums in $\mathcal{E}_1$ and $\mathcal{E}_2$. First note that the inner sum of $\mathcal{E}_1$ becomes

[48] Dropping a modular condition conceptually loses a W^d power, but we can tolerate such losses by eventually by requiring A to be sufficiently uniform with respect to W.

$$\sum_{\underline{y}_1, \underline{y}_2 \in \Lambda} \left[2\alpha + f_A(\underline{y}_1)\right] f_A(\underline{y}_2) 1_{\mathcal{F}_3, \lambda^2(t_0, \underline{t}, \underline{t})}(\underline{x}_2 - \underline{x}_1, \underline{y}_1 - \underline{x}_1, \underline{y}_2 - \underline{x}_1).$$

As we do for $\underline{x}_1$ and $\underline{x}_2$, we can drop the modular condition of $\underline{y}_1$ and make the uniform bound $|2\alpha + f_A(\underline{y}_1)| \leq 3$ to obtain

$$|\mathcal{E}_1| \lesssim \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \\ \underline{x}_2, \underline{y}_1 \in \mathbb{Z}^d}} \left| \sum_{\underline{y}_2 \in \Lambda} f_A(\underline{y}_2) 1_{\mathcal{F}_3, \lambda^2(t_0, \underline{t}, \underline{t})}(\underline{x}_2 - \underline{x}_1, \underline{y}_1 - \underline{x}_1, \underline{y}_2 - \underline{x}_1) \right|,$$

and

$$|\mathcal{E}_2| \lesssim_{t,d} \lambda^{d-4} \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \\ \underline{x}_2 \in \mathbb{Z}^d}} \left| \sum_{\underline{y} \in \Lambda} f_A(\underline{y}) 1_{\mathcal{F}_2, \lambda^2(t_0, \underline{t})}(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1) \right|.$$

From here, we handle $\mathcal{E}_1$ and $\mathcal{E}_2$ in essentially the same way as outlined in the proof of Lemma 6.1 in [LM22] using estimates from [Mag09], albeit with a wrinkle thrown in by Λ. To mirror the notation in those papers, we use introduce the following:[49]

[49] Though we do not normalize here.

$$\sigma_t(\underline{y}) := \begin{cases} 1 & \text{if } |\underline{y}|^2 = \underline{t}, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\sigma_{\underline{t}}^{\underline{x}}(\underline{y}) := \sigma_{t_1}(\underline{y}) \sigma_{t_2}(\underline{y} + \underline{x}).$$

We start by bounding the simpler sum, $\mathcal{E}_2$. By reindexing $\underline{x}_2 \to \underline{x}_2 + \underline{x}_1$ and $\underline{y} \to \underline{x}_1 - \underline{y}$, and observing $\sigma_t(-\underline{y}) = \sigma_t(\underline{y})$ we have

$$\begin{aligned}
|\mathcal{E}_2| &\lesssim_{t,d} \lambda^{d-4} \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \\ \underline{x}_2 \in \mathbb{Z}^d}} \left| \sum_{\underline{y} \in \mathbb{Z}^d} 1_\Lambda(\underline{y}) f_A(\underline{y}) 1_{\mathcal{F}_2, \lambda^2(t_0, \underline{t})}(\underline{x}_2 - \underline{x}_1, \underline{y} - \underline{x}_1) \right| \\
&\lesssim_{t,d} \lambda^{d-4} \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \\ \underline{x}_2 \in \mathbb{Z}^d}} \left| \sum_{\underline{y} \in \mathbb{Z}^d} 1_\Lambda(\underline{x}_1 - \underline{y}) f_A(\underline{x}_1 - \underline{y}) 1_{\mathcal{F}_2, \lambda^2(t_0, \underline{t})}(\underline{x}_2, -\underline{y}) \right| \\
&\lesssim_{t,d} \lambda^{d-4} \sum_{\substack{\underline{x}_1 \in Q_{N_1}, \\ \underline{x}_2 \in \mathbb{Z}^d}} \sigma_{\lambda^2 t_0}(\underline{x}_2) \left| \sum_{\underline{y} \in \mathbb{Z}^d} 1_\Lambda(\underline{x}_1 - \underline{y}) f_A(\underline{x}_1 - \underline{y}) \sigma_{\lambda^2 \underline{t}}^{\underline{x}_2}(\underline{y}) \right| \\
&\lesssim_{t,d} \lambda^{d-4} \sum_{\underline{x}_2 \in \mathbb{Z}^d} \sigma_{\lambda^2 t_0}(\underline{x}_2) \sum_{\underline{x}_1 \in Q_{N_1}} \left| 1_\Lambda f_A * \sigma_{\lambda^2 \underline{t}}^{\underline{x}_2}(\underline{x}_1) \right|. \qquad (4.8)
\end{aligned}$$

From here, we use Cauchy-Schwarz twice, with the first use on the functions

$$\sigma_{\lambda^2 t_0}(\underline{x}_2)^{1/2} \qquad \text{and} \qquad \sigma_{\lambda^2 t_0}(\underline{x}_2)^{1/2} \sum_{\underline{x}_1 \in Q_{N_1}} \left| 1_\Lambda f_A * \sigma_{\lambda^2 \underline{t}}^{\underline{x}_2}(\underline{x}_1) \right|,$$

the second use on the functions

$$1_{Q_{N_1}}(\underline{x}_1) \qquad \text{and} \qquad \left| 1_\Lambda f_A * \sigma_{\lambda^2 \underline{t}}^{\underline{x}_2}(\underline{x}_1) \right|,$$

and finally apply Plancherel.[50] We first note the standard estimate

$$\sum_{\underline{x}_2 \in \mathbb{Z}^d} \sigma_{\lambda^2 t_0}(\underline{x}_2) \lesssim_{t,d} \lambda^{d-2}.$$

Thus

$$\begin{aligned} |\mathcal{E}_2|^2 &\lesssim_{t,d} \lambda^{2d-8} \lambda^{d-2} \sum_{\underline{x}_2 \in \mathbb{Z}^d} \sigma_{\lambda^2 t_0}(\underline{x}_2) \left(\sum_{\underline{x}_1 \in Q_{N_1}} \left| 1_\Lambda f_A * \sigma^{\underline{x}_2}_{\lambda^2 \underline{t}}(\underline{x}_1) \right| \right)^2 \\ &\lesssim_{t,d} \lambda^{3d-10} N_1^d \sum_{\underline{x}_2 \in \mathbb{Z}^d} \sigma_{\lambda^2 t_0}(\underline{x}_2) \sum_{\underline{x}_1 \in \mathbb{Z}^d} \left| 1_\Lambda f_A * \sigma^{\underline{x}_2}_{\lambda^2 \underline{t}}(\underline{x}_1) \right|^2 \\ &\lesssim_{t,d} \lambda^{3d-10} N^d \int_{\mathbb{T}} \left| \widehat{1_\Lambda f_A}(\underline{\xi}) \right|^2 H_{\lambda^2 (t_0, \underline{t})}(\underline{\xi}) \, d\underline{\xi}, \end{aligned} \tag{4.9}$$

where

$$H_{(t_0, \underline{t})}(\underline{\xi}) := \sum_{\underline{x}_2 \in \mathbb{Z}^d} \sigma_{t_0}(\underline{x}_2) \left| \widehat{\sigma}^{\underline{x}_2}_{\underline{t}}(\underline{\xi}) \right|^2.$$

We make the observation

$$\begin{aligned} \widehat{1_\Lambda f_A}(\underline{\xi}) &= \sum_{\underline{y} \in \mathbb{Z}^d} 1_\Lambda(\underline{y}) f_A(\underline{y}) e^{-2\pi i \underline{y} \cdot \underline{\xi}} \\ &= \sum_{\underline{y}_1 \in \mathbb{Z}^d} f_A(W \underline{y}_1 + \underline{v}) e^{-2\pi i (W \underline{y}_1 + \underline{v}) \cdot \underline{\xi}} \\ &= e^{-2\pi i \underline{v} \cdot \underline{\xi}} \sum_{\underline{y}_1 \in \mathbb{Z}^d} g_{\underline{v}}(\underline{y}_1) e^{-2\pi i \underline{y}_1 \cdot W \underline{\xi}} \\ &= e^{-2\pi i \underline{v} \cdot \underline{\xi}} \widehat{g}_{\underline{v}}(W \underline{\xi}), \end{aligned} \tag{4.10}$$

where $g_{\underline{v}} := 1_{R_{W,\underline{v}}(A)} - \alpha 1_{Q_{N/W}}$ and $R_{W,\underline{v}}(A)$ is the W-scaled, $\underline{v}$-restricted set of A.[51]

Plugging equation (4.10) into equation (4.9), we obtain

$$|\mathcal{E}_2|^2 \lesssim_{t,d} \lambda^{3d-10} N^d \int_{\mathbb{T}} \left| \widehat{g}_{\underline{v}}(W \underline{\xi}) \right|^2 H_{\lambda^2 (t_0, \underline{t})}(\underline{\xi}) \, d\underline{\xi}. \tag{4.11}$$

We may now recall the m and L as given in the statement of this lemma, and apply the trick of Lemma 6.1 in [LM22] to write

[50] Writing a $1/2$ power on an indicator function is a bit silly, but we include it for familiarity.

[51] The function $g_{\underline{v}}$ is not quite the balance function of $R_{W,\underline{v}}(A)$ because $R_{W,\underline{v}}(A)$ may not have precisely density α, though by uniformity its density must be close to α.

$$|\mathcal{E}_2|^2 \lesssim_{t,d} \lambda^{3d-10} N^d \int_{\mathbb{T}} \left|\widehat{g}_{\underline{v}}(W\underline{\xi})\right|^2 H_{\lambda^2(t_0,\underline{t})}(\underline{\xi}) \widehat{\chi}_{m,L}(W\xi)^2 \, d\underline{\xi}$$
$$+ \lambda^{3d-10} N^d \int_{\mathbb{T}} \left|\widehat{g}_{\underline{v}}(W\underline{\xi})\right|^2 H_{\lambda^2(t_0,\underline{t})}(\underline{\xi})(1 - \widehat{\chi}_{m,L}(W\xi)^2) \, d\underline{\xi}. \tag{4.12}$$

By Plancherel and the definition of the $U^1_{m,L}$ norm, we have that the first term above is $\lesssim_d \lambda^{3d-10} N^{2d} \|g_{\underline{v}}\|^2_{U^1_{m,L}(Q_{N/W})} \sup_{\underline{\xi}\in\mathbb{T}} |H_{\lambda^2(t_0,\underline{t})}(\underline{\xi})|$. We note that when $\lambda \lesssim_t |\underline{x}_2| \lesssim_t \lambda$, [52]

$$\left|\widehat{\sigma}^{\underline{x}_2}_{\underline{t}}(\underline{\xi})\right| \le \sum_{\underline{y}\in\mathbb{Z}^d} \sigma_{t_1}(\underline{y})\sigma_{t_2}(\underline{y} + \underline{x}_2) \lesssim_{t,d} \lambda^{d-4}, \tag{4.13}$$

so

$$|H_{\lambda^2(t_0,\underline{t})}(\underline{\xi})| \lesssim_{t,d} \lambda^{2d-8} \sum_{\underline{x}_2\in\mathbb{Z}^d} \sigma_{t_0}(\underline{x}_2) \lesssim_{t,d} \lambda^{3d-10}.$$

We note that $\widehat{\chi}_{m,L}(W\xi) = \widehat{\chi}_{Wm,L}(\xi)$. By equation 6.11 in [LM22] and Plancherel, provided $L \lesssim_t \eta^{10}\lambda$, the second term of equation (4.12) is bounded by $C_{t,d}\lambda^{3d-10}N^d\eta^2\lambda^{3d-10}\sum_{\underline{x}\in\mathbb{Z}^d} |g_{\underline{v}}(\underline{x})|^2$. [53,54]

Bounding $\sum_{\underline{x}\in\mathbb{Z}^d} |g_{\underline{v}}(\underline{x})|^2$ trivially by N^d, we apply our calculations to equation (4.12) and report

$$|\mathcal{E}_2|^2 \lesssim_{t,d} N^{2d}\lambda^{6d-20} \left(\|g_{\underline{v}}\|^2_{U^1_{m,L}(Q_{N/W})} + \eta^2\right).$$

This of course yields

$$|\mathcal{E}_2| \lesssim_{t,d} N^{d}\lambda^{3d-10} \left(\|g_{\underline{v}}\|_{U^1_{m,L}(Q_{N/W})} + \eta\right). \tag{4.14}$$

We now turn to estimating $|\mathcal{E}_1|$. Fortunately, we can proceed with with the same method of calculation as was used for $|\mathcal{E}_2|$, so we omit the details of the manipulations. Reindexing and reordering yields the $\mathcal{E}_1$ version of equation (4.8):

$$|\mathcal{E}_1| \lesssim \sum_{\underline{x}_2,\underline{y}_1\in\mathbb{Z}^d} \sigma_{\lambda^2 t_0}(\underline{x}_2)\sigma^{\underline{x}_2}_{\lambda^2\underline{t}}(-\underline{y}_1) \sum_{\underline{x}_1\in Q_{N_1}} \left|1_\Lambda f_A * \sigma^{\underline{x}_2}_{\lambda^2\underline{t}}(\underline{x}_1)\right|.$$

[52] A condition forced by the $\sigma_{\lambda^2 t_0}(\underline{x}_2)$ term in H.

[53] We have used that the normalization factor for H is $\lesssim_{t,d} \lambda^{3d-10}$.

[54] The η gain is explained in [Mag09] and represents the fact that H is small away from rationals with small denominator.

Cauchy-Schwarz and Plancherel return the $\mathcal{E}_1$ version of equation (4.11)

$$|\mathcal{E}_1|^2 \lesssim_{t,d} \lambda^{2d-6} N^d \int_{\mathbb{T}} \left| \widehat{1_\Lambda f_A}(\underline{\xi}) \right|^2 H_{\lambda^2(t_0,\underline{t},\underline{t})}(\underline{\xi})\, d\underline{\xi},$$

where

$$H_{(t_0,\underline{t},\underline{t})}(\underline{\xi}) := \sum_{\underline{x}_2,\underline{y}_1 \in \mathbb{Z}^d} \sigma_{t_0}(\underline{x}_2)\sigma_{\underline{t}}^{\underline{x}_2}(-\underline{y}_1) \left| \widehat{\sigma}_{\underline{t}}^{\underline{x}_2}(\underline{\xi}) \right|^2.$$

Again using equation (4.13), we have

$$|H_{\lambda^2(t_0,\underline{t},\underline{t})}(\underline{\xi})| \lesssim_{t,d} \lambda^{2d-8} \sum_{\underline{x}_2,\underline{y}_1 \in \mathbb{Z}^d} \sigma_{t_0}(\underline{x}_2)\sigma_{\underline{t}}^{\underline{x}_2}(-\underline{y}_1) \lesssim_{t,d} \lambda^{4d-14},$$

and the estimates in [Mag09] once again allow us to use the $\widehat{\chi}_{m,L}(W\xi)^2$ trick to obtain

$$|\mathcal{E}_1|^2 \lesssim_{t,d} N^{2d}\lambda^{6d-20} \left(\|g_{\underline{v}}\|^2_{U^1_{m,L}(Q_{N/W})} + \eta^2 \right),$$

whence

$$|\mathcal{E}_1| \lesssim_{t,d} N^d \lambda^{3d-10} \left(\|g_{\underline{v}}\|_{U^1_{m,L}(Q_{N/W})} + \eta \right). \tag{4.15}$$

Finally, note that A was assumed to be $(W^{-d/2}\eta, L)$-uniformly distributed with respect to Wm, so in fact by Lemma 1.13 we have $\|g_{\underline{v}}\|_{U^1_{m,L}(Q_{N/W})} < \eta$.

Combining estimates equation (4.7), equation (4.14), and equation (4.15) in equation (4.6) finishes off the lemma. □

The estimate $E^A_{W,\underline{v}}$ shows that $\mathcal{N}^A_{\lambda,\underline{t},W,\underline{v}}(X) \approx 2^d \alpha \lambda^{d-4} W^{2-d} \omega_{\mathcal{F}_1}(\underline{t})$ for most X, i.e. we have an analogy to the finite field exceptional set as given in Theorem 2.21.

Corollary 4.21. *Let $A \subset Q_N$. Let $t_0 \in \mathbb{Z}$, $\underline{t} \in \mathbb{Z}^2$ and set $W = 2\prod_{p<w} p$. Let $\underline{u}_1, \underline{u}_2, \underline{v} \in [W]^d$ with $\underline{u}_2 - \underline{u}_1$, $\underline{v} - \underline{u}_1$, and $\underline{v} - \underline{u}_2$ primitive. Suppose $d > 40$ and*

$$\mathcal{F}_2(\underline{u}_2 - \underline{u}_1, \underline{v} - \underline{u}_1) \equiv \lambda^2(t_0, \underline{t}) \pmod{2W}.$$

Fix $0 < \eta \ll 1$ and set $m := \operatorname{lcm}\{1 \le n \le C_{t,d}\eta^{-10}\}$. Suppose that $1 \ll L \ll \eta^{10}\lambda$ and assume A is $(W^{-d/2}\eta, L)$-uniformly distributed with respect to Wm. Additionally, let $W^{3000} \le \lambda \ll N$.

Define

$$\mathcal{X} := \{(\underline{x}_1, \underline{x}_2) \in Q_N^2 \,:\, |\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0,\ \underline{x}_1 \equiv \underline{u}_1\,(W),\ \underline{x}_2 \equiv \underline{u}_2\,(W)\}.$$

Fix $\rho > 0$. If $X \in \mathcal{X} \setminus B_\rho$ then

$$\mathcal{N}^A_{\lambda,\underline{t},W,\underline{v}}(X) \geq (1-\rho)2^d\alpha\lambda^{d-4}W^{2-d}\omega_{\mathcal{F}_1}(\underline{t}),$$

where

$$|B_\rho| < 2^{2d}\rho^{-2}N^d\lambda^{d-2}W^{1-2d}\omega^{-2}_{\mathcal{F}_1}(\underline{t})\mathcal{E},$$

with

$$\mathcal{E} = o_{w\to\infty}(1) + o_{\lambda\to\infty}(1) + o_{N/\lambda^3\to\infty}(1) + o_{\eta\to 0}(1).$$

Proof. We make the simple observation that if $\{a_n\}$ is a sequence of real numbers with

$$\sum_n a_n^2 = T,$$

then for a cutoff $C > 0$,

$$\#\{n \,:\, |a_n| > C\} < \frac{T}{C^2}.$$

We apply this observation with $E^A_{W,\underline{v}}$ in Lemma 4.20 by setting

$$T = 2^{4d}\alpha^2N^d\lambda^{3d-10}W^{5-4d}\mathcal{E},$$

$$C = \rho 2^d\alpha\lambda^{d-4}W^{2-d}\omega_{\mathcal{F}_1}(\underline{t}).$$

This yields $|B_\rho| < TC^{-2}$, i.e.

$$|B_\rho| < 2^{2d}\rho^{-2}N^d\lambda^{d-2}W^{1-2d}\omega^{-2}_{\mathcal{F}_1}(\underline{t})\mathcal{E}$$

as desired.

□

4.3 Isometric Copies of Distance Trees of Triangles in Subsets of $\mathbb{Z}^d$

The ideas of the exceptional set argument in the finite fields setting form the core of this section. Within a set $A \subset Q_N$ which is sufficienly uniformly distributed, we inductively count the number of isometric copies of $\lambda\Gamma$. This section demonstrates that most of the time we are able to add a vertex $\underline{y} \in A$ to Γ subject to two edge conditions and a modulo W restriction in about as many ways as expected.

We recall the definition of the exceptional set E_W from Chapter 3:

$$E_W = \left\{ \underline{y} \in \mathbb{Z}^d \;:\; \min_{\underline{x} \in \mathbb{Z}^d \setminus \langle \underline{y} \rangle} \mathrm{Dist}(\underline{x}, \langle \underline{y} \rangle) < \frac{W}{|\underline{y}|} \right\}.$$

Sometimes we wish to count with respect to a basepoint, so additionally define

$$E'_W := \{(\underline{x}_1, \underline{x}_2) \;:\; \underline{x}_1, \underline{x}_2 \in \mathbb{Z}^d,\ \underline{x}_2 - \underline{x}_1 \in E_W\}.$$

It will be important to bound the number of points in E_W which lie on a lattice sphere. We get estimates in Lemmas 4.23 and 4.26.

Lemma 4.22. *If $\underline{y} \in E_W$ then there exists a nonzero vector $\underline{y}^* \in \mathbb{Z}^d$ such that $|\underline{y}^*| < W$ and $\underline{y}^* \cdot \underline{y} = 0$.*

Proof. Let $\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{Z}^d$ and let $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n)$ denote the n-dimensional volume of the parallelepiped generated by $\underline{x}_1, \ldots, \underline{x}_n$. Then

$$\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n) = \mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_{n-1})\, \mathrm{Dist}(\underline{x}_n, \langle \underline{x}_1, \ldots, \underline{x}_{n-1} \rangle). \quad (4.16)$$

As $\underline{y} \in E_W$, there exists some $\underline{x} \in \mathbb{Z}^d$ with $\mathrm{Dist}(\underline{x}, \langle \underline{y} \rangle) < W|\underline{y}|^{-1}$. Let $\underline{e}_1, \ldots, \underline{e}_{d-3}$ be distinct standard basis vectors which are not in $\langle \underline{y}, \underline{x} \rangle$. Then by equation (4.16)

$$\mathrm{Vol}(\underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3}) < W. \quad (4.17)$$

Let $\underline{m} \in \mathbb{Z}^d \setminus \langle \underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3} \rangle$. Then $\underline{m}$ together with $\underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3}$ generates a non-degenerate d-dimensional parallelepiped, so

$$1 \leq \mathrm{Vol}(\underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3}, \underline{m}).$$

which by equation (4.16) and equation (4.17) implies

$$\frac{1}{W} < \mathrm{Dist}(\underline{m}, \langle \underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3} \rangle). \quad (4.18)$$

Let $\underline{y}^* \in \mathbb{Z}^d$ be the[55] primitive vector in $\langle \underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3} \rangle^\perp$. We have

[55] Unique up to sign.

$$\mathrm{Dist}(\underline{m}, \langle \underline{y}, \underline{x}, \underline{e}_1, \ldots, \underline{e}_{d-3} \rangle) = \frac{\underline{y}^* \cdot \underline{m}}{|\underline{y}^*|}.$$

However, as $\underline{y}^*$ is primitive, there exists $\underline{m} \in \mathbb{Z}^d$ such that $\underline{y}^* \cdot \underline{m} = 1$. Equation (4.18) then shows $|\underline{y}^*| < W$. □

Lemma 4.23. *Let $t \in \mathbb{N}$ and let $\lambda \in \sqrt{\mathbb{N}}$. If $d \geq 13$ then*

$$\#\{\underline{y} \in E_W \ : \ |\underline{y}|^2 = t\lambda^2\} \lesssim_{t,d} \lambda^{d-3} W^d.$$

Proof. By Lemma 4.22, there is a primitive vector $\underline{y}^* \in \mathbb{Z}^d$ such that $|\underline{y}^*| < W$ and $\underline{y}^* \cdot \underline{y} = 0$. By a volume estimate, there are $\lesssim_d W^d$ choices of $\underline{y}^* \in \mathbb{Z}^d$ with $|\underline{y}^*| < W$ and by Theorem 3.29 for each choice of $\underline{y}^*$ there are $\lesssim_{t,d} \lambda^{d-3}|\underline{y}^*|^{-1}$ choices for $\underline{y}$ with $|\underline{y}^2| = t\lambda^2$ and $\underline{y}^* \cdot \underline{y} = 0$. Putting the estimates together gives

$$\#\{\underline{y} \in E_W \ : \ |\underline{y}|^2 = t\lambda^2\} \lesssim_{t,d} \lambda^{d-3} W^d.$$

$\square$

Corollary 4.24. *Let $t \in \mathbb{N}$ and let $\lambda \in \sqrt{\mathbb{N}}$. If $d \geq 13$ then*

$$\#\{(\underline{x}_1, \underline{x}_2) \in E'_W \cap Q_N \ : \ |\underline{x}_2 - \underline{x}_1|^2 = t\lambda^2\} \lesssim_{t,d} N^d \lambda^{d-3} W^d.$$

Proof. This follow directly from Lemma 4.23 after noting that there are $\lesssim_d N^d$ options for $\underline{x}_1 \in Q_N$. $\square$

Remark 4.25. What we would like to do next is to prove

$$|\{\underline{y} \in E_W : \ |\underline{y}|^2 = t_1\lambda^2, \underline{y} \cdot \underline{x} = \mu\}| \leq C_{t_1,d,W}\, \lambda^{d-4.5},$$

uniformly for $c_1\lambda < |\underline{x}| < c_2\lambda$.[56] This would imply by induction that there are few isometric copies of the distance graph $\lambda\Gamma_0$ with an exceptional edge in E_W. Since E_W consists of vectors orthogonal to some short vector $\underline{y}^*$ with $|\underline{y}^*| < W$ as shown in Lemma 4.22, it would be enough to show the above inequality for $\underline{y} \in \langle \underline{y}^* \rangle^\perp \cap \mathbb{Z}^d$, for a fixed $\underline{y}^*$.

[56] Any estimate of the form $\lambda^{d-4-\varepsilon}$ would be sufficient.

Let's assume for simplicity that $\underline{y}^* = (0, \ldots, 0, 1)$. Then we'd like to have that

$$|\{\underline{y}' \in \mathbb{Z}^{d-1} : \ |\underline{y}'|^2 = t_1\lambda^2, y' \cdot x' = \mu\}| \leq C_{t_1,W,d}\, \lambda^{d-4.5}.$$

By Theorem 3.29, we have that

$$|\{\underline{y}' \in \mathbb{Z}^{d-1} : \ |\underline{y}'|^2 = t_1\lambda^2, \underline{y}' \cdot x' = \mu\}| \leq C_{t_1,W,d}\, \lambda^{d-4}|\underline{x}''|^{-1},$$

where $\underline{x}''$ is the primitive part of $\underline{x}'$. Thus the estimate holds outside the set (going back to $\mathbb{Z}^d$),

$$D_\lambda := \{\underline{x} \in \mathbb{Z}^d : \ \underline{x} = \ell\, \underline{y}^* + k\underline{x}'', \ |\underline{x}''| \leq \lambda^{1/2}, \ \ell, k \in \mathbb{Z}\}.$$

However, it is easy to see that the intersection of D_λ with any sphere $S_{\underline{z},t,\lambda}$ of dimension $d-2$ and radius $\approx \lambda$ is roughly bounded by $\lambda^{(d-2)/2+2}$, as for fixed k and ℓ, the intersection is contained in the intersection of a spherical cap of radius $R = |k|\lambda^{1/2}$ and the lattice $k\mathbb{Z}^d$, and there are at most λ^2 choices for k and ℓ.

In dimensions $d \geq 14$, this is less then $\lambda^{d-4.5}$ and this will imply that there are few isometric copies of the distance graph $\lambda\Gamma_0$ with an exceptional edge in D_λ.[57] Then one can further argue that there are few isometric copies with an exceptional edge in E_W as we may assume that none of the edges are in D_λ, see Lemma 4.34. We proceed now to make the heuristics above precise.

[57] We needed $d \geq 14$ to apply Theorem 3.29 in $d-1$ dimensions.

Lemma 4.26. *Let* $\underline{t} \in \mathbb{N}^2$ *and let* $\lambda \in \sqrt{\mathbb{N}}$. *Let* $d \geq 14$ *and let* $\underline{x} \in \mathbb{Z}^d$ *be primitive with* $c_1\lambda < |\underline{x}| < c_2\lambda$. *Set*

$$S_{\underline{x},\lambda,\underline{t}} := \{\underline{y} \in \mathbb{Z}^d \ : \ \mathcal{F}_1(\underline{x},\underline{y}) = \lambda^2\underline{t}\}.$$

There is a set $D_{W,\lambda}$ *with the properties that if* $\underline{x} \notin D_{W,\lambda}$, *then*

$$\#\left\{\underline{y} \in S_{\underline{x},\lambda,\underline{t}} \ : \ \underline{y} \in E_W \text{ or } \underline{y} - \underline{x} \in E_W\right\} \lesssim_{\underline{t},d,W} \lambda^{d-4.5},$$

and

$$\max_{\substack{\underline{z}\in\mathbb{Z}^d,\\ c_1\lambda<|\underline{z}|<c_2\lambda,\\ \underline{z}\text{ primitive}}} \#\left\{\underline{y} \in S_{\underline{z},\lambda,\underline{t}} \ : \ \underline{y} \in D_{W,\lambda} \text{ or } \underline{y} - \underline{z} \in D_{W,\lambda}\right\} \lesssim_{\underline{t},d,W} \lambda^{d-4.5}.$$

Proof. Motivated by Remark 4.25, for a given sufficiently large $\lambda \in \mathbb{N}$, we define the derived exceptional set

$$D_{W,\lambda} := \bigcup_{\substack{\underline{y}^*\in\mathbb{Z}^d,\\ 0<|\underline{y}^*|<W.\\ \underline{y}^*\text{ primitive}}} D_{\underline{y}^*,W,\lambda}$$

where

$$D_{\underline{y}^*,W,\lambda} := \{\ell\underline{y}^* + k\underline{x}' \ : \ |\underline{x}'| \leq C_W\lambda^{1/2}, \ \underline{x}' \text{ prim}, \ k,\ell \in \mathbb{Z}, \ |k| \leq C_W\lambda\},$$

with C_W a sufficiently large constant depending on Γ, d, and W.

We first estimate

$$\#\left\{\underline{y} \in S_{\underline{z},\lambda,\underline{t}} \ : \ \underline{y} \in D_{W,\lambda} \text{ or } \underline{y} - \underline{z} \in D_{W,\lambda}\right\}.$$

Without loss of generality, let $\underline{y} \in D_{\underline{y}^*,W,\lambda} \cap S_{\underline{z},\lambda,\underline{t}}$ with $c_1\lambda < |\underline{z}| < c_2\lambda$.[58]

[58] Either $\underline{y}$ or $\underline{y}-\underline{z}$ is in $D_{W,\lambda}$, but if it is the latter, we may interchange the roles of $\underline{z}$ and $\underline{y}-\underline{z}$. This assumption necessitates doubling the resulting bounds, but constant factors will be absorbed in any case.

Since $|k| \le C_W \lambda$ we have that $|\ell| \le C_W\, \lambda^{3/2}$ as $k|\underline{x}'| \le C_W \lambda^{3/2}$ and we have $|\underline{y}| \le C\lambda$. Thus there are at most $C_W\, \lambda^{5/2}$ choices for k and ℓ.

Fix $\underline{y}^*$, k, and ℓ. Then the intersection

$$\{\ell \underline{y}^* + k\underline{x}' \ :\ |\underline{x}'| \le C_W \lambda^{1/2}\} \cap S_{\underline{z},\lambda,\underline{t}},$$

is either empty or contains a point $\underline{y}_0$, and then it is contained in

$$(\underline{y}_0 + B_{C_W k \lambda^{1/2}}) \cap S_{\underline{z},\lambda,\underline{t}} \cap (\underline{y}_0 + k\mathbb{Z}^d),$$

where B_R denotes the ball of radius R centered at the origin. Since spherical caps on the $d-2$-dimensional sphere $S_{\underline{z},\lambda,\underline{t}}$ of radius $R = k$ centered at lattice points $\underline{y} = \underline{y}_0 + k\mathbb{Z}^d$ are disjoint, the above intersection is bounded by $(C_W k \lambda^{1/2})^{d-2} k^{-(d-2)} = C_W \lambda^{(d-2)/2}$. Noting that there are at most $C_W \lambda^{5/2}$ choices for y^* and $\ell, k \in \mathbb{Z}$, and that $(d+3)/2 \le d - 4.5$ when $d \ge 14$,[59] we have

$$\max_{\substack{\underline{z}\in\mathbb{Z}^d,\\ c_1\lambda<|\underline{z}|<c_2\lambda,\\ \underline{z}\text{ primitive}}} \#\left\{\underline{y} \in S_{\underline{z},\lambda,\underline{t}} \ :\ \underline{y} \in D_{W,\lambda} \text{ or } \underline{y} - \underline{z} \in D_{W,\lambda}\right\} \lesssim_{\underline{t},d,W} \lambda^{d-4.5}.$$

[59] We could of course do better than an exponent of $d-4.5$, but we keep this one as it will allow us to combine this error with other errors later in the argument, e.g. in Lemma 4.34.

Now we estimate

$$\#\left\{\underline{y} \in S_{\underline{x},\lambda,\underline{t}} \ :\ \underline{y} \in E_W \text{ or } \underline{y} - \underline{x} \in E_W\right\},$$

given $\underline{x} \notin D_{W,\lambda}$ and $c_1\lambda < |\underline{x}| < c_2\lambda$.

As before, we immediately assume $\underline{y} \in E_W$ as the other case is the same.[60] Fix $\underline{y}^* \in \mathbb{Z}^d$ such that $\underline{y}^*$ is primitive $0 < |\underline{y}^*| < W$ and $\underline{y}^* \cdot \underline{y} = 0$. Then write

$$\mathbb{Z}^d \cap \langle \underline{y}^* \rangle^\perp = \mathbb{Z}\, \underline{b}_1 \oplus \cdots \oplus \mathbb{Z}\, \underline{b}_{d-1},$$

[60] Again, this technically doubles the resulting bound, but the constant is absorbed.

so we can write $\underline{y}$ uniquely in the form $\underline{y} = m_1 \underline{b}_1 + \ldots + m_{d-1}\underline{b}_{d-1}$, with $\underline{m} = (m_1, \ldots, m_{d-1}) \in \mathbb{Z}^{d-1}$. Then in variables $\underline{m} = (m_1, \ldots, m_{d-1})$, we have

$$|\underline{y}|^2 = \sum_{i,j=1}^{d-1} \underline{b}_i \cdot \underline{b}_j\, m_i m_j =: A_B\, \underline{m} \cdot \underline{m},$$

where A_B is a positive definite matrix with entries $\underline{b}_i \cdot \underline{b}_j$. Since $0 < |\underline{y}^*| < W$ we may choose the lattice basis vectors $\underline{b}_i$ so that $|\underline{b}_i| \le C_W$ for $i = 1, \ldots, d-1$. There are finitely many choices of the basis vectors $\underline{b}_i$, so by compactness there exists a constant $M_W > 0$ such that for all $\underline{m} \in \mathbb{Z}^{d-1}$

$$M_W^{-1}|\underline{m}|^2 \leq A_B\underline{m} \cdot \underline{m} \leq M_W|\underline{m}|^2,$$

uniformly for $0 < |\underline{y}^*| < W$ and for all of the possible corresponding bases $B = (\underline{b}_1, \ldots, \underline{b}_{d-1})$ satisfying $|\underline{b}_i| \leq C_W$. Consider now the linear equation

$$\mu = \underline{y} \cdot \underline{x} = \sum_{i=1}^{d-1} m_i\, \underline{b}_i \cdot \underline{x} =: \underline{m} \cdot \underline{\chi},$$

where $\underline{\chi} = (\underline{b}_1 \cdot \underline{x}, \ldots, \underline{b}_{d-1} \cdot \underline{x})$. Thus to prove the lemma, it is enough to show that if $\underline{x} \notin D_{y^*,W,\lambda}$, then

$$\#\{\underline{m} \in \mathbb{Z}^{d-1} : A_B\underline{m} \cdot \underline{m} = t_1\lambda^2,\ \underline{m} \cdot \underline{\chi} = \mu\} \leq C_{t_1,d,W}\ \lambda^{d-4.5}.$$

By Theorem 3.29 we have that

$$\#\{\underline{m} \in \mathbb{Z}^{d-1} : A_B\underline{m} \cdot \underline{m} = t_1\lambda^2,\ \underline{m} \cdot \underline{\chi} = \mu\} \leq C_{t_1,d,W}\ \lambda^{d-4}\, |\underline{\chi}'|^{-1}, \tag{4.19}$$

where $\underline{\chi}'$ denotes the primitive part of the vector $\underline{\chi}$. We make the following elementary observation: Suppose $\underline{\chi} = k\underline{\chi}'$. Then there exists $\ell \in \mathbb{Z}$ such that $\underline{x} = \ell\underline{y}^* + k\underline{x}'$ for some $\ell \in \mathbb{Z}$ and $\underline{x}' \in \mathbb{Z}^d$.

Since $\mathbb{Z}^d \cap \langle \underline{y}^* \rangle^\perp = \mathbb{Z}\,\underline{b}_1 \oplus \cdots \oplus \mathbb{Z}\,\underline{b}_{d-1}$ there exists an integer vector $\underline{b}_d$ so that $\mathbb{Z}^d = \mathbb{Z}\,\underline{b}_1 \oplus \cdots \oplus \mathbb{Z}\,\underline{b}_{d-1} \oplus \mathbb{Z}\,\underline{b}_d$. Since we assume that $\underline{y}^*$ is primitive and $\underline{b}_1, \ldots, \underline{b}_{d-1}$ form a lattice basis of $\langle y^* \rangle^\perp \cap \mathbb{Z}^d$ we must have that $\underline{b}_d \cdot \underline{y}^* = 1$. Choose $\ell \in \mathbb{N}$ so that $\ell \equiv \underline{b}_d \cdot \underline{x} \pmod k$, then $\underline{b}_i \cdot (\underline{x} - \ell\underline{y}^*) \equiv 0 \pmod k$ for all $i = 1, \ldots, d$. Since $\underline{b}_1, \ldots, \underline{b}_d$ form a basis of $\mathbb{Z}^d$ this implies that $\underline{z} \cdot (\underline{x} - \ell\underline{y}^*) \equiv \underline{0} \pmod k$ for all $\underline{z} \in \mathbb{Z}^d$ thus $\underline{x} \equiv \ell\underline{y}^* \pmod k$ i.e. $\underline{x} = \ell\underline{y}^* + k\underline{x}'$, for some vector $\underline{x}' \in \mathbb{Z}^d$.

Fix $\underline{y}^*$ and write $\underline{\chi} = k\underline{\chi}'$ with $\underline{\chi}'$ being the primitive part of $\underline{\chi}$. We have $|k| \leq C_W\lambda$ as $|k| \leq |\underline{\chi}|$. If $|\underline{\chi}'| \geq c_W\lambda^{1/2}$ then equation (4.19) gives us

$$\#\{\underline{m} \in \mathbb{Z}^{d-1} : A_B\underline{m} \cdot \underline{m} = t_1\lambda^2,\ \underline{m} \cdot \underline{\chi} = \mu\} \leq C_{t_1,d,W}\ \lambda^{d-4.5}$$

Now assume for the sake of contradiction that $\underline{x} \notin D_{\underline{y}^*,W,\lambda}$, but $|\underline{\chi}'| < c_W\lambda^{1/2}$ for some sufficiently small constant $c_W > 0$.

Then, by previous observation, we have that $\underline{x} = \ell\underline{y}^* + k\underline{x}'$ which implies that $\underline{\chi}' = (\underline{b}_1 \cdot \underline{x}', \ldots, \underline{b}_{d-1} \cdot \underline{x}')$ as recall that $\underline{\chi} = (\underline{b}_1 \cdot \underline{x}, \ldots, \underline{b}_{d-1} \cdot \underline{x})$. Thus by our assumption

$$|\underline{\chi}'|^2 = \sum_{i=1}^{d-1} |\underline{b}_i \cdot \underline{x}'|^2 \le c_W^2 \, \lambda.$$

Note that $\underline{x} = \ell \underline{y}^* + k\underline{x}' = (\ell + km)\underline{y}^* + k\underline{x}''$ for any $m \in \mathbb{Z}$ where $\underline{x}'' = \underline{x}' - m\underline{y}^*$. Choose m so that $|\underline{x}'' \cdot \underline{y}^*| = |\underline{x}' \cdot \underline{y}^* - m|\underline{y}^*|^2|$ is minimized. Since $|\underline{y}^*|^2 < W^2$, it is guaranteed that there is some $\underline{x}''$ with $|\underline{x}'' \cdot \underline{y}^*| < W^2$, so $\underline{x}''$ and $\underline{y}^*$ are almost orthogonal.

We will show that $|\underline{x}''| \le \lambda^{1/2}$ which implies $\underline{x} \in D_{\underline{y}^*,W,\lambda}$ in contradiction to our assumption on $\underline{x}$. If we assume $|\underline{x}''| > \lambda^{1/2}$ then writing $\underline{x}'' = t\underline{y}^* + \underline{z}$ with $\underline{z} \cdot \underline{y}^* = 0$ we have that

$$|\underline{z}| \ge |\underline{x}''| - |t\underline{y}^*| \ge \lambda^{1/2} - W^2 \ge \frac{1}{2}\lambda^{1/2}.$$

Note that $\underline{b}_i \cdot \underline{x}' = \underline{b}_i \cdot \underline{x}'' = \underline{b}_i \cdot \underline{z}$ for $i = 1, \ldots, d-1$. However, since $\underline{z} \in \mathbb{R}\,\underline{b}_1 \oplus \cdots \oplus \mathbb{R}\,\underline{b}_{d-1}$ and $\underline{b}_1, \ldots, \underline{b}_{d-1}$ are linearly independent, it follows by compactness and the fact that there are only a bounded number of choices for the basis $B = (\underline{b}_1, \ldots, \underline{b}_{d-1})$, that

$$|\underline{z}|^2 \le C_W \sum_{i=1}^{d-1} |\underline{b}_i \cdot \underline{z}|^2 < \lambda/4,$$

choosing $c_W > 0$ sufficiently small, contradicting the fact that $|\underline{z}| \ge \frac{1}{2}\lambda^{1/2}$. Thus indeed $|\underline{x}''| \le \lambda^{1/2}$ and $\underline{x} \in D_{\underline{y}^*,W,\lambda}$. □

Remark 4.27. In the previous lemma, we have used a generalization of Theorem 3.29 to write equation (4.19) where A_B was a positive definite matrix with eigenvalues between M_W^{-1} and M_W for some constant $M_W > 0$. We note here that this generalization follows from some minor updates to pieces of the Chapter 3 argument. First, the minor arcs estimate goes through by appealing to Lemma 4.3 in Birch [Bir62], where the estimate is obtained for a general form. Secondly, writing $A_B = U^T DU$ where U is orthogonal and D is diagonal, allows a reduction to the case of just the diagonal matrix of eigenvalues for estimates in Lemmas 3.15 and 3.17. This leads to a change by a factor of M_W which is captured by the constant $C_{t_1,d,W}$ in equation (4.19).

The author would like to thank Ákos Magyar for providing a detailed description of the approach to Lemma 4.26 and Remarks 4.25 and 4.27.

The upper bounds in Chapter 3 require $\underline{y}$ primitive, so we make the following general definition for distance graphs.

Definition 4.28. *A distance graph $\Gamma \subset \mathbb{Z}^d$ is primitive if the vectors corresponding to each edge of Γ are all primitive.*

We now set notation for two distance graphs to be isometric:

Definition 4.29. *For two distance graphs $\Gamma, \Gamma' \subset \mathbb{Z}^d$, we say $\Gamma' \sim \Gamma$ if there is a bijection between Γ' and Γ which preserves edge lengths.*

Definition 4.30.

$$\mathcal{B}(\underline{x}) := \{\underline{y} \in S_{\underline{x},W,\underline{v}}(\lambda, \mu) \ : \ \underline{y} \text{ or } \underline{y} - \underline{x} \text{ non-primitive.}\}$$

The following estimate allows us to consign non-primitive isometric copies of $\lambda\Gamma$ to an error term.

Lemma 4.31. *Let $W = 2\prod_{p<w} p$. Let $\underline{x} \notin E_W$ be primitive with $c_1\lambda < |\underline{x}| < c_2\lambda$ and $\underline{x} \equiv \underline{u} \pmod{W}$. Let $\underline{v} \in [W]^d$ such that $\underline{v}$ and $\underline{u} - \underline{v}$ are primitive. If $d \geq 13$, and $W^{3000} < \lambda$, then*

$$|\mathcal{B}(\underline{x})| = \lambda^{d-4} W^{2-d} o_{w\to\infty}(1).$$

Proof. Without loss of generality, assume that $\underline{y} \in \mathcal{B}(\underline{x})$ is non-primitive.[61] Set $g := \gcd(y_1, \ldots, y_d)$. The equation $|\underline{y}|^2 = \lambda^2$ implies $g^2 \mid \lambda^2$ and as $\underline{y} \equiv \underline{v} \pmod{W}$ with $\underline{v}$ primitive, we have necessarily $\gcd(g, W) = 1$.[62] However, $\underline{y}$ is non-primitive so $g > 1$, i.e. the prime divisors of g are all greater than or equal to w. Now observe $g^{-1}\underline{y}$ is an integer vector lying in a smaller sphere:

$$g^{-1}\underline{y} \in S_{\underline{x},W,g^{-1}\underline{v}}(g^{-1}\lambda, g^{-1}\mu),$$

where $g^{-1}\underline{v}$ should be interpreted modulo W.[63] We can bound $|\mathcal{B}(\underline{x})|$ by summing over the estimate from Theorem 3.47 for $S_{\underline{x},W,\underline{v}}(g^{-1}\lambda, \mu)$ where $g^2 \mid \lambda^2$. The theorem says that when $\underline{x} \notin E_W$ and when $W^{2000} < g^{-1}\lambda$, we have

$$|S_{\underline{x},W,\underline{v}}(g^{-1}\lambda, \mu)| \lesssim_d (g^{-1}\lambda)^{d-4} W^{2-d}.$$

Hence

$$|\mathcal{B}(\underline{x})| \leq \sum_{\substack{g^2|\lambda^2,\\ w\leq g}} |S_{\underline{x},W,\underline{v}}(g^{-1}\lambda, \mu)|$$

[61] By the definition of $\mathcal{B}(\underline{x})$, one of $\underline{y}$ or $\underline{x} - \underline{y}$ is non-primitive, but we can freely interchange the roles of $\underline{y}$ and $\underline{x} - \underline{y}$.

[62] The condition $g^2 \mid \lambda^2$ looks silly, but is what we want because $\lambda \in \sqrt{\mathbb{N}}$.

[63] This is valid as $\gcd(g, W) = 1$.

$$
\begin{aligned}
&\leq \sum_{\substack{g^2|\lambda^2,\\ w\leq g<\lambda W^{-2000}}} |S_{\underline{x},W,\underline{v}}(g^{-1}\lambda,\mu)| + \mathcal{E}\\
&\leq c_d\lambda^{d-4}W^{2-d} \sum_{\substack{g^2|\lambda^2,\\ w\leq g<\lambda W^{-2000}}} \frac{1}{g^{d-4}} + \mathcal{E},\\
&\leq c_d\lambda^{d-4}W^{2-d} o_{w\to\infty}(1),
\end{aligned}
$$

where

$$
\mathcal{E} = \sum_{\substack{g^2|\lambda^2,\\ \lambda W^{-2000}\leq g}} |S_{\underline{x},W,\underline{v}}(g^{-1}\lambda,\mu)|.
$$

Making the crude observation that $\mathcal{E}$ is counting distinct lattice points inside a ball of radius W^{2000}, we get $\mathcal{E} \lesssim_d W^{2000d}$. This is lower order than $\lambda^{d-4}W^{2-d}$ when we have assumed $W^{3000} < \lambda$. □

Definition 4.32. *Let Γ be a distance graph with n vertices. Let $U = (\underline{u}_1, \ldots, \underline{u}_n)$ be a tuple with $\underline{u}_i \in [W]^d$ for $i = 1, \ldots, n$.*

$$
\mathcal{N}_{N,W,U}(\Gamma) := \#\{\Gamma' \subset Q_N : \Gamma' \sim \Gamma, \Gamma' \textit{ primitive}, V(\Gamma') \equiv U \pmod W\},
$$

where $V(\Gamma')$ is the ordered tuple of the vertices of Γ' and $Q_N = [-N, N]^d$.

Definition 4.33. *Let Γ be a distance graph with n vertices. Let $U = (\underline{u}_1, \ldots, \underline{u}_n)$ be a tuple with $\underline{u}_i \in [W]^d$. We say U is primitive with respect to Γ if $\underline{u}_j - \underline{u}_i$ is primitive whenever (v_i, v_j) is an edge in Γ.*

Lemma 4.34. *Let $\Gamma \subset \mathbb{Z}^d$ be a distance tree of triangles with n vertices and e edges. Let $W = 2\prod_{p<w} p$ and let $U = (\underline{u}_1, \ldots, \underline{u}_n)$ with $\underline{u}_i \in [W]^d$ and where U is primitive with respect to Γ. Let $1 \lesssim_{\Gamma,d,W} \lambda \ll N$.[64] If $d \geq 13$ then there is a large constant[65] $c = c(d, \Gamma)$ such that*

$$
\mathcal{N}_{N,W,U}(\lambda\Gamma) \leq c^n N^d \lambda^{(n-1)d-2e} W^{e-nd}.
$$

[64] A suitable value of the implied lower bound constant for λ may be chosen from the proof.

[65] It is important that c is independent of W.

Proof. For convenience, we suppress the subscripts in $\mathcal{N}_{N,W,U}(\Gamma)$. We will also need to track distance graphs lying on exceptional edges, so using the set $D_{\lambda,W}$ from Lemma 4.26, we define

$$
\begin{aligned}
D_\lambda(\Gamma) := \#\{\Gamma' \subset Q_N \ :& \ \Gamma' \sim \Gamma,\ \Gamma' \text{ primitive},\ V(\Gamma') \equiv U \pmod W),\\
& \Gamma' \text{ has an edge in } E_W \text{ and an edge in } D_{W,\lambda}\}.
\end{aligned}
$$

$$
\begin{aligned}
E_\lambda(\Gamma) := \#\{\Gamma' \subset Q_N \ :& \ \Gamma' \sim \Gamma,\ \Gamma' \text{ primitive},\ V(\Gamma') \equiv U \pmod W),\\
& \Gamma' \text{ has an edge in } E_W,\ \text{ but no edge in } D_{W,\lambda}\}.
\end{aligned}
$$

We will proceed via induction. If Γ has a single vertex, then the only constraint is $V(\Gamma') \equiv \underline{u}_1 \pmod{W}$, so

$$\mathcal{N}(\lambda\Gamma) \lesssim_d N^d W^{-d}. \tag{4.20}$$

We now fix $c = c(d, \Gamma)$ to be more than twice each of the implied constants appearing in equation (4.20), Lemma 4.23, and Theorem 3.47. Additionally we choose c to be more than twice $2^{d+1}\omega_{\mathcal{F}_0}(t_0)$ where t_0 is the square of the edge length between v_1 and v_2 in Γ. This last is so we can apply Lemma 4.11 when moving from one vertex to two vertices.

If Γ has two vertices and a single edge, we fix the starting vertex v_1 and use Lemma 4.11 with $r = 1$ to obtain $\leq \frac{c}{2}\lambda^{d-2}W^{1-d}$ options for v_2.[66] By Lemma 4.23, at most $\frac{c}{2}\lambda^{d-3}W^d$ of these are such that the edge from v_1 to v_2 lies in the exceptional set E_W.[67] Multiplying by the number of choices for v_1, we have

$$\mathcal{N}(\lambda\Gamma) \leq c^2 N^d \lambda^{d-2} W^{1-2d},$$

$$E_\lambda(\lambda\Gamma) \leq c^2 N^d \lambda^{d-3},$$

and

$$D_\lambda(\lambda\Gamma) \leq c^2 N^d \lambda^{d-3}.$$

[66] The rank condition for $\mathcal{F}_0$ is satisfied when $d > 6$.

[67] In our case, we have an additional modulo W restriction, so one should expect the truth to be that this upper bound is λ^{d-3}. However, the extra powers of W are dwarfed by the power savings in λ and we do not pursue this precision.

For induction, assume that if Γ is a distance tree of triangles with $2 \leq k < n$ vertices and e edges, then

$$\mathcal{N}(\lambda\Gamma) \leq c^k N^d \lambda^{(k-1)d-2e} W^{e-kd},$$

$$E_\lambda(\lambda\Gamma) \leq C_{k,\Gamma,d,W} N^d \lambda^{(k-1)d-2e-0.5},$$

and

$$D_\lambda(\lambda\Gamma) \leq C_{k,\Gamma,d,W} N^d \lambda^{(k-1)d-2e-0.5},$$

where $C_{k,\Gamma,d,W}$ depends on k, Γ, d, and W, and for $k = 2, \ldots, n$, we assume we have chosen λ such that

$$2WC_{k,\Gamma,d,W} \leq \lambda^{0.5}.$$

Let $2 \leq k < n$ and let Γ be a distance tree of triangles with $k+1$ vertices. We can build Γ by adding v_{k+1} adjacent to two prior vertices v_i and v_j which are themselves joined by a primitive edge. By Theorem 3.47 if that edge is not in the exceptional set, there are $\leq \frac{c}{2}\lambda^{d-4}W^{2-d}$ choices for v_{k+1} of which by Lemma 4.26 we have $\lesssim_{\Gamma,d,W} \lambda^{d-4.5}$ with an exceptional edge to v_i or v_j, and also $\lesssim_{\Gamma,d,W} \lambda^{d-4.5}$ where the edge to v_i or v_j is in $D_{W,\lambda}$. On the other hand, Theorem 3.47 says that if the v_i to v_j edge is in the exceptional set E_W, then there are $\leq \frac{c}{2}\lambda^{d-4}W^{3-d}$ choices for v_{k+1}. Letting Γ_0 denote Γ restricted to its first k vertices and noting that the number of edges of Γ is $e+2$ where e is the number of edges of Γ_0, we have

$$\begin{aligned}
\mathcal{N}(\lambda\Gamma) &\leq \frac{c}{2}\lambda^{d-4}W^{2-d}\mathcal{N}(\lambda\Gamma_0) + \frac{c}{2}\lambda^{d-4}W^{3-d}\left[E_\lambda(\lambda\Gamma_0) + D_\lambda(\lambda\Gamma_0)\right] \\
&\leq \frac{c^{k+1}}{2}N^d\lambda^{kd-2(e+2)}W^{e+2-(k+1)d}\left[1 + 2WC_{k,\Gamma,d,W}\lambda^{-0.5}\right] \\
&\leq c^{k+1}N^d\lambda^{kd-2(e+2)}W^{e+2-(k+1)d}
\end{aligned}$$

as $2WC_{k,\Gamma,d,W}\lambda^{-0.5} \leq 1$ by assumption, and

$$\begin{aligned}
E_\lambda(\lambda\Gamma) &\lesssim_{\Gamma,d,W} \lambda^{d-4.5}\mathcal{N}(\lambda\Gamma_0) + \frac{c}{2}\lambda^{d-4}W^{3-d}E(\lambda\Gamma_0) \\
&\lesssim_{\Gamma,d,W} N^d\lambda^{kd-2(e+1)-0.5},
\end{aligned}$$

and

$$\begin{aligned}
D_\lambda(\lambda\Gamma) &\lesssim_{\Gamma,d,W} \lambda^{d-4.5}\mathcal{N}(\lambda\Gamma_0) + \frac{c}{2}\lambda^{d-4}W^{3-d}E(\lambda\Gamma_0) \\
&\lesssim_{\Gamma,d,W} N^d\lambda^{kd-2(e+1)-0.5}.
\end{aligned}$$

□

In analogy to Definition 2.23, we define a count where some of the vertices are fixed.

Definition 4.35. *Let Γ be a distance graph with n vertices. Let $X = (\underline{x}_1, \ldots, \underline{x}_k)$ with $\underline{x}_i \in Q_N$, let $I = (i_1, \ldots, i_k)$ be an index set, and let $U = (\underline{u}_1, \ldots, \underline{u}_n)$ be a tuple with $\underline{u}_i \in [W]^d$ for $i = 1, \ldots, n$.*

$$\begin{aligned}
\mathcal{N}_{I,N,W,U}(\Gamma, X) := \#\{\Gamma' \subset Q_N \ :\Gamma' \sim \Gamma,\ \Gamma' \textit{ primitive}, \\
V(\Gamma') \equiv U \pmod{W},\ V_I(\Gamma') = X\}
\end{aligned}$$

where $V(\Gamma')$ is the ordered tuple of the vertices of Γ' and $V_I(\Gamma')$ is the ordered tuple of those vertices with indices in I.

The following lemma is the analogue of Lemma 2.24. However, we have replaced the weak requirement that a k-degenerate distance graph Γ be in $2k$-general position with the strong requirement that Γ is a tree of triangles. The reason for this is that we need estimates after a reordering of the graph, but in the integer lattice we only have estimates for 2-degenerate graphs, while in finite fields we could handle a doubling of the degeneracy.

Lemma 4.36. *Let $\Gamma \subset \mathbb{Z}^d$ be a distance tree of triangles with n vertices and e edges. Let $W = 2\prod_{p<w} p$ and let $U = (\underline{u}_1, \dots, \underline{u}_n)$ with $\underline{u}_i \in [W]^d$ and where U is primitive with respect to Γ. Let $X = (\underline{x}_1, \underline{x}_2)$ with $\underline{x}_1, \underline{x}_2 \in Q_N$ and let $I = (i_1, i_2)$ be an index set for which there is an edge between v_{i_1} and v_{i_2} in Γ. Let $1 \lesssim_{\Gamma,d,W} \lambda \ll N$. If $d \geq 13$ then there are large constants*[68] $c = c(d, \Gamma)$ *and $C_{\Gamma,d,W}$ such that*

[68] Again, it is crucial that c is independent of W.

$$\mathcal{N}_{I,N,W,U}(\lambda\Gamma, X) \leq \begin{cases} c^{n-2}\lambda^{(n-2)d-2(e-1)}W^{e-1-(n-2)d} & \text{if } X \notin E'_W, \\ C_{\Gamma,d,W}\lambda^{(n-2)d-2(e-1)} & \text{if } X \in E'_W. \end{cases}$$

Proof. As we did in the previous lemma, we suppress the subscripts N, W, and U in $\mathcal{N}_{I,N,W,U}(\lambda\Gamma, X)$ and define

$$\begin{aligned} D_{\lambda,I}(\Gamma, X) := \#\{\Gamma' \subset Q_N : \; & \Gamma' \sim \Gamma,\ \Gamma' \text{ primitive},\ V(\Gamma') \equiv U \pmod{W}, \\ & V_I(\Gamma') = X,\ \Gamma' \text{ has an edge in } E_W, \\ & \Gamma' \text{ has an edge in } D_{W,\lambda}\}. \end{aligned}$$

$$\begin{aligned} E_{\lambda,I}(\Gamma, X) := \#\{\Gamma' \subset Q_N : \; & \Gamma' \sim \Gamma,\ \Gamma' \text{ primitive},\ V(\Gamma') \equiv U \pmod{W}, \\ & V_I(\Gamma') = X,\ \Gamma' \text{ has an edge in } E_W, \\ & \Gamma' \text{ has no edge in } D_{W,\lambda}\}, \end{aligned}$$

where $D_{W,\lambda}$ is the set from Lemma 4.26.

As in the proof of Lemma 2.24, we immediately consider a reordering of Γ where v_{i_1} and v_{i_2} are the first and second vertices. As Γ is a tree of triangles, it is possible to choose a reordering in a way so that we can still construct the graph by adding two sides of a triangle at each step. Let $\widetilde{\Gamma}$ be this reordered graph.[69]

[69] For a general 2-degenerate graph, it is not possible to maintain 2-degeneracy after a reordering. The smallest counterexample is K_5 with one edge removed.

The reordering does not affect the count, so letting $[2]$ denote the index tuple $(1, 2)$, we have

$$\mathcal{N}_I(\lambda\Gamma, X) = \mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}, X),$$

where we have implicitly also reordered the modular conditions U to $\widetilde{U}$.

Note that if $\widetilde{\Gamma}$ has two vertices, then

$$\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}, X) \leq 1.$$

There may be no valid isometric copy of $\lambda\widetilde{\Gamma}$ if X does not satisfy the modular restrictions or if the two points of X are not the correct distance apart. However, if there is a solution, it is of course unique as X is fixed.

From here, we follow the ideas of the previous proof. We again start by fixing $c \in \mathbb{R}$ such that c is at least twice as large as the implied constant appearing in Theorem 3.47.[70]

[70] This is a subset of the requirements in Lemma 4.34, so we may as well take the same c as appears there.

Let $\widetilde{\Gamma}$ have $n \geq 2$ vertices and e edges. If $X \in E'_W$, then we add a vertex with two edges $n-2$ times, and each time we apply the exceptional set upper bound $\frac{c}{2}\lambda^{d-4}W^{3-d}$ for a choice of next vertex, obtaining

$$\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}, X) \leq \left(c\lambda^{d-4}W^{3-d}\right)^{n-2} \leq C_{\Gamma,d,W}\lambda^{(n-2)(d-2)-2(e-1)},$$

where we have observed that as Γ is a tree of triangles with more than 2 vertices, we have $e = 2n-3$ and hence $(n-2)(d-4) = (n-2)(d-2) - 2(e-1)$.

Now we examine the case $X \notin E'_W$. We note in this case that when $\widetilde{\Gamma}$ has two vertices, we have

$$D_{\lambda,[2]}(\lambda\widetilde{\Gamma}, X) = E_{\lambda,[2]}(\lambda\widetilde{\Gamma}, X) = 0.$$

If $\widetilde{\Gamma}$ has three vertices, then by Theorem 3.47 there are $\leq \frac{c}{2}\lambda^{d-4}W^{2-d}$ choices for v_3 of which by Lemma 4.26 at most $\lesssim_{\Gamma,d,W} \lambda^{d-4.5}$ come with an edge in the exceptional set E_W and at most $\lesssim_{\Gamma,d,W} \lambda^{d-4.5}$ come with an edge in $D_{\lambda,W}$, so

$$\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}, X) \leq c\lambda^{d-4}W^{2-d},$$

$$E_{\lambda,[2]}(\lambda\widetilde{\Gamma}, X) \lesssim_{\Gamma,d,W} \lambda^{d-4.5},$$

and

$$D_{\lambda,[2]}(\lambda\widetilde{\Gamma}, X) \lesssim_{\Gamma,d,W} \lambda^{d-4.5}.$$

For the sake of induction, we now assume that if $\widetilde{\Gamma}$ has k vertices and e edges with $3 \leq k < n$, then

$$\mathcal{N}_{[2]}(\lambda\Gamma) \leq c^{k-2}\lambda^{(k-2)d-2(e-1)}W^{e-1-(k-2)d},$$
$$E_{\lambda,[2]}(\lambda\Gamma) \leq C_{k,\Gamma,d,W}\lambda^{(k-2)d-2(e-1)-0.5},$$

and

$$D_{\lambda,[2]}(\lambda\Gamma) \leq C_{k,\Gamma,d,W}\lambda^{(k-2)d-2(e-1)-0.5},$$

where $C_{k,\Gamma,d,W}$ depends on k, Γ, d, and W, and for $k = 3, \dots, n$, we assume we have chosen λ such that

$$2WC_{k,\Gamma,d,W} \leq \lambda^{0.5}.$$

Let $3 \leq k < n$ and let Γ be a distance tree of triangles with $k+1$ vertices. Let $\widetilde{\Gamma}_0$ again denote $\widetilde{\Gamma}$ restricted to its first k vertices and let e denote the number of edges of $\widetilde{\Gamma}_0$, i.e. $\widetilde{\Gamma}$ has $e+2$ edges. With exactly the same argument as the previous lemma, we have by Theorem 3.47 and Lemma 4.26

$$\begin{aligned}\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}) \leq & \frac{c}{2}\lambda^{d-4}W^{2-d}\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}_0) \\ & \quad + \frac{c}{2}\lambda^{d-4}W^{3-d}\left[E_{\lambda,[2]}(\lambda\widetilde{\Gamma}_0) + D_{\lambda,[2]}(\lambda\widetilde{\Gamma}_0)\right] \\ & \leq \frac{c^{k-1}}{2}\lambda^{(k-1)d-2(e+1)}W^{e+1-(k-1)d}\left[1 + 2WC_{k,\Gamma,d,W}\lambda^{-0.5}\right] \\ & \leq c^{k-1}\lambda^{(k-1)d-2(e+1)}W^{e+1-(k-1)d}\end{aligned}$$

as $2WC_{k,\Gamma,d,W}\lambda^{-0.5} \leq 1$ by assumption. Similarly

$$\begin{aligned}E_{\lambda,[2]}(\lambda\widetilde{\Gamma}) & \lesssim_{\Gamma,d,W} \lambda^{d-4.5}\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}_0) + \frac{c}{2}\lambda^{d-4}W^{3-d}E_{\lambda,[2]}(\lambda\widetilde{\Gamma}_0) \\ & \lesssim_{\Gamma,d,W} \lambda^{(k-2)d-2(e-1)-0.5},\end{aligned}$$

and

$$\begin{aligned}D_{\lambda,[2]}(\lambda\widetilde{\Gamma}) & \lesssim_{\Gamma,d,W} \lambda^{d-4.5}\mathcal{N}_{[2]}(\lambda\widetilde{\Gamma}_0) + \frac{c}{2}\lambda^{d-4}W^{3-d}D_{\lambda,[2]}(\lambda\widetilde{\Gamma}_0) \\ & \lesssim_{\Gamma,d,W} \lambda^{(k-2)d-2(e-1)-0.5}.\end{aligned}$$

□

Definition 4.37. *Let $A \subset Q_N$. Let Γ be a distance graph with n vertices. Let $U = (\underline{u}_1, \dots, \underline{u}_n)$ be a tuple with $\underline{u}_i \in [W]^d$ for $i = 1, \dots, n$.*

$$\mathcal{N}^A_{N,W,U}(\Gamma) := \#\{\Gamma' \subset A : \Gamma' \sim \Gamma, \Gamma' \text{ primitive}, V(\Gamma') \equiv U \pmod{W}\},$$

where $V(\Gamma')$ is the ordered tuple of the vertices of Γ' and $Q_N = [-N,N]^d$.

Lemma 4.38. *Let Γ be a distance tree of triangles with n vertices in $\mathbb{Z}^d$ where $d > 40$. Let $W = 2\prod_{p<w} p$. Then for every $\lambda \in \sqrt{\mathbb{N}}$, there exists some $U(\lambda) = (\underline{u}_1, \dots, \underline{u}_n)$ with $\underline{u}_i \in [W]^d$ such that if $e = (v_i, v_j) \in E(\Gamma)$ with length $\sqrt{t_e}$, then $\underline{u}_j - \underline{u}_i$ is primitive and $|\underline{u}_j - \underline{u}_i|^2 \equiv \lambda^2 t_e \pmod{2W}$.*

Proof. It is sufficient to find primitive solutions modulo 4 and modulo p for each $3 \leq p < w$ because we can then piece such solutions together to find a primitive solution in $[W]^d$ via the Chinese remainder theorem. For $p \geq 3$, we can turn to our work in finite fields, though we shall have to use an ad hoc method for $p = 2$.

Fix $p \geq 3$. We will generate a solution $U_p = (\underline{u}_{1,p}, \ldots, \underline{u}_{n,p})$ with $\underline{u}_{i,p} \in [p]$ iteratively. The condition that the solution is primitive is exactly that $\underline{u}_{1,i} \neq \underline{u}_{j,p}$ whenever v_i and v_j are adjacent in Γ. First we set $\underline{u}_{1,p} = 0$. Now, suppose the length of the edge between v_1 and v_2 in Γ is $\sqrt{t_0}$, so we look for a solution to $|\underline{u}_{2,p}|^2 \equiv \lambda^2 t_0 \pmod{p}$. When $\lambda^2 t_0$ is not divisible by p, our estimate on the number of points on a sphere in finite fields, Corollary 2.9 says there are at least $p^{d-1} - 2p^{(d-1)/2}$ valid choices for $|\underline{u}_{2,p}|$. This guarantees a solution for $p \geq 3$ when $d \geq 2$. However, it is possible that p divides $\lambda^2 t_0$, so we need to generate a more general estimate. Fortunately, it is not a very difficult adjustment in the proof. Writing Corollary 2.9 with the distance $t = 0$ and center $\underline{u}_{1,p} = 0$, we have

$$|S_0(0)| = p^{-1} \sum_{y \in \mathbb{F}_p} \sigma_0(y) = p^{d-1} \hat{\sigma}_0(0).$$

To estimate $\hat{\sigma}_0(0)$, we use the computation in Lemma 2.8. All of the equations go through up to equation (2.4), which says

$$\hat{\sigma}_t(r) = 1_{\{0\}}(r) + p^{-d} G(\eta, \chi)^d \sum_{s \in \mathbb{F}_p^*} \eta^d(s) \chi(-st - |r|^2 (4s)^{-1}).$$

Using the normal Gaussian sum calculation $|G(\eta, \chi)| = p^{1/2}$ and the trivial estimate on the inner sum[71] yields

$$|\hat{\sigma}_0(0)| \geq 1 - p^{1-d/2},$$

which in turn gives

$$|S_0(0)| \geq p^{d-1} - p^{d/2-1}.$$

This shows there are (nontrivial) solutions for $\underline{u}_{2,p}$ with $p \geq 3$ when $d \geq 2$. Finally, for $k \geq 3$, we need to show the existence of a $\underline{u}_{k,p}$ so that $|\underline{u}_{k,p} - \underline{u}_{i,p}|^2 \equiv \lambda^2 t_i \pmod{p}$ and $|\underline{u}_{k,p} - \underline{u}_{j,p}|^2 \equiv \lambda^2 t_j \pmod{p}$ with $\underline{u}_{k,p} \neq \underline{u}_{i,p}$ and $\underline{u}_{k,p} \neq \underline{u}_{j,p}$ where i and j are the indices of the backwards neighbors of v_k in Γ. As Γ is a tree of triangles, v_i and v_j share an edge and must already be distinct. We can make exactly the same trivial estimate edit to equation equation (2.5), to find that for any $(t_i, t_j) \in \mathbb{F}_p^2$ we have

$$|S_{(t_i,t_j)}(\underline{u}_{i,p}, \underline{u}_{j,p})| \geq p^{d-2} - p^{d/2-1}.$$

[71] When $t \neq 0$ we could do better than the trivial estimate.

When $p \geq 3$, the choice $d \geq 4$ is good enough to guarantee a solution for $\underline{u}_{k,p}$ which is distinct from $\underline{u}_{i,p}$ and $\underline{u}_{j,p}$.

Finally, we show the existence of a solution modulo 8 with a direct construction which works when $d \geq 21$.[72] We first note that it is possible to find a solution modulo 2 using $d = 3$. To see this, note that if v_k has a backwards edge to v_i and v_j, it is sufficient to check that $|\underline{u}_{k,2} - \underline{u}_{i,2}|^2 \equiv \lambda^2 t_i \pmod 2$, as if so the condition $|\underline{u}_{k,2} - \underline{u}_{j,2}|^2 \equiv \lambda^2 t_j \pmod 2$ is forced by virtue of the fact that v_i, v_j, and v_k form a triangle in $\mathbb{Z}^d$, where we have the relation $t_0 + t_1 - t_2 \equiv 0 \pmod 2$. To make sure $\underline{u}_{k,2} \neq \underline{u}_{i,2}$ and $\underline{u}_{k,2} \neq \underline{u}_{j,2}$ we need the sphere of radius (squared) 1 to have at least 2 distinct points and the sphere of radius 0 to have at least 3 distinct points. This is true for $d \geq 3$.

[72] There should be simpler solutions, but this is sufficient for our purposes since the dimension requirement is $d > 40$ in any case

We next move to finding a nontrivial solution modulo 4. For any edge of length (squared) t_e, write $t_e = s_e \mod 2$. The idea is that we can use 3 coordinates to write down a modulo 2 solution which works up to parity for a graph with edge lengths (squared) s_e, and concatenate that with a solution for a graph which only has edge lengths 0 and 2.[73] We have a solution modulo 2 with already, and so it remains to show we can build any tree of triangles modulo 4 where the edge lengths are all 0 and 2.[74] We can do this with 6 coordinates with a direct construction.

[73] It's possible $s_e = 1$ and we use a corresponding edge $(1, 1, 1)$ which has length 3 modulo 4, but this is not actually a problem as the only thing we need is the same parity.

[74] And the $\underline{u}_{k,2} \neq \underline{u}_{i,2}$ and $\underline{u}_{k,2} \neq \underline{u}_{j,2}$ requirement is already met by the modulo 2 solution.

To showcase the idea, suppose we are trying to find a copy of the triangle in $\mathbb{Z}^2$ with with coordinate $(0, 0)$, $(1, 0)$, and $(3, 1)$. It has side lengths 1, $\sqrt{5}$, and $\sqrt{10}$. We write these as $1 \equiv 0 + 1 \pmod 4$, $5 \equiv 0 + 1 \pmod 4$, $10 \equiv 2 + 0 \pmod 4$. First we find a copy of a triangle with (squared) edge lengths 1, 1, and 0 in $(\mathbb{Z}/2\mathbb{Z})^3$, say with the coordinates $(0, 0, 0)$, $(1, 0, 0)$, and $(1, 1, 1)$. The distances (squared) between these coordinates are 1, 3, and 2 modulo 4. This tells us to look for a copy of a triangle with (squared) edge lengths $1 - 1 \equiv 0 \pmod 4$, $5 - 3 \equiv 2 \pmod 4$, and $10 - 2 \equiv 0 \pmod 4$ in $(\mathbb{Z}/4\mathbb{Z})^6$, which we describe how to do in the following paragraphs, and then we concatenate the solutions. One valid solution in $(\mathbb{Z}/4\mathbb{Z})^6$ is

$$(0, 0, 0, 0, 0, 0), (1, 1, 1, 1, 0, 0), (1, 0, 0, 0, 1, 0).$$

Concatenating the solutions, we find a copy of the original triangle in $(\mathbb{Z}/4\mathbb{Z})^9$ with coordinates

$$(0, 0, 0 \mid 0, 0, 0, 0, 0, 0), (1, 0, 0 \mid 1, 1, 1, 1, 0, 0), (0, 1, 0 \mid 1, 0, 0, 0, 1, 0).$$

Let e denote 0 or 2 and let o denote 1 or 3. Suppose first that $|\underline{u}_{j,2} - \underline{u}_{i,2}| \equiv 2 \pmod 4$. Then shifting $\underline{u}_{i_2}$ to $\underline{0}$, we let $\underline{u}_{i,2} = (0, 0, 0, 0, 0, 0)$ and, with a possible reordering of the coordinates, write $\underline{u}_{j,2} = (o, o, e, e, e, e)$. We show with our construction we can avoid the other case $\underline{u}_{j,2} = (o,o,o,o,o,o)$.[75] There are 4 possibilities for pairs of edges of length 0 or 2 connecting $\underline{u}_{k,2}$ to $\underline{u}_{i,2}$ and $\underline{u}_{j,2}$, and we can check that the choices

$$\underline{u}_{k,2} \in \{(0,1,1,1,1,0),(1,1,1,1,0,0),(0,0,1,1,0,0),(0,1,1,0,0,0)\}$$

cover each of those possibilities.

On the other hand, suppose that $|\underline{u}_{j,2} - \underline{u}_{i,2}| \equiv 0 \pmod 4$. Again we shift $\underline{u}_{i,2}$ to $\underline{0}$ and after reordering the coordinates write $\underline{u}_{j,2} = (o,o,o,o,e,e)$. We will show that we can avoid the case $\underline{u}_{j,2} = (e,e,e,e,e,e)$. In the case $\underline{u}_{j,2} = (o,o,o,o,e,e)$ we check that the choices of

$$\underline{u}_{k,2} \in \{(1,1,0,0,1,1),(1,0,0,0,1,0),(1,1,1,0,1,0),(1,1,0,0,0,0)\}$$

cover all of the possible edge length requirements. We note that for each of our choices of $\underline{u}_{k,2}$ at least one coordinate is the same parity and one coordinate is different parity from both $\underline{0}$ and $\underline{u}_{j,2}$. Therefore, neither of the cases $\underline{u}_{j,2} = (o,o,o,o,o,o)$ or $\underline{u}_{j,2} = (e,e,e,e,e,e)$ can occur as we build the graph, and the only two cases are those we described. This concludes the search for solutions modulo 4.

Finally we need to find solutions modulo 8, this time with the requirement $\underline{u}_{k,2} \in (\mathbb{Z}/4\mathbb{Z})^2$. Thankfully there is a relatively cheap trick available to make use of the modulo 4 solutions we already have available. This time we take an edge of length (squared) t_e and write it as $t_e = r_e \mod 4$. We use 9 coordinates to write down a solution for a graph with edges r_e modulo 4, and then are left with constructing all graphs with edges of length 0 or 4. This can be done by adapting our solution for constructing all graphs with edges of length 0 or 2. The trick is to take a graph with only edges of length 0 or 2 and concatenate a copy of itself. This has the effect of doubling the number of coordinates, but also doubles the edge lengths while not increasing the values of coordinates. Thus we get a solution in $\{0,1,2,3\}^{12}$ for any graph with edges of length (squared) 0 and 4 in $\mathbb{Z}/8\mathbb{Z}$. Concatenating this solution with the 9-coordinate solution for graphs with edges r_e gives us the solution we desire in $9+12=21$ dimensions. □

Theorem 4.39. *Let Γ be a distance tree of triangles in $\mathbb{Z}^d$ with n vertices and e edges. Let $\omega(v_i)$ denote $\omega_{\mathcal{F}_0}(t_0)$ for $i=2$ and denote $\omega_{\mathcal{F}_1}(\underline{t}(v_i))$ for $i \geq 3$ where t_0 is the square of the length of the edge between v_1 and v_2 and $\underline{t}(v_i) = (t_1, t_2)$ is the squares of the lengths of the two backwards edges from v_i for $i \geq 3$.*

Let $A \subset Q_N \subset \mathbb{Z}^d$ where $d > 40$. Let $\alpha := \frac{|A|}{|Q_N|}$ be the density of A with $\alpha \geq \beta$ for some $\beta > 0$.[76] There exist $w = w(\Gamma, \beta, d) \in \mathbb{N}$, $\eta = \eta(\Gamma, \beta, d, w)$

[75] The only way $x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 \equiv 2 \pmod 4$ is if two or six of the variables are odd.

[76] We set a lower bound of β for α because in application we allow N to grow, and we do not want the other parameters to depend on N.

with $\eta > 0$*, and* $m = m(\Gamma, \beta, d, w, \eta) \in \mathbb{N}$ *such that if* A *is* (η, L)*-uniform with respect to* m *for some* $L \in \mathbb{N}$*, then there exists a* $\lambda_0 = \lambda_0(\Gamma, \beta, d, w, \eta, L)$ *such that for every* $\lambda \in \sqrt{\mathbb{N}}$ *with* $\lambda_0 \leq \lambda \leq N^{1/4}$

$$\mathcal{N}^A_{N,W,U}(\lambda\Gamma) \geq 2^{n(d-1)}\alpha^n N^d \lambda^{(n-1)d-2e} W^{e-nd} \prod_{i=2}^{n} \omega(v_i),$$

for some $U = (\underline{u}_1, \ldots, \underline{u}_n)$ *with* $\underline{u}_i \in [W]^d$ *where* $W := 2\prod_{p<w} p$*.*

Proof. This proof stems from the proof of Theorem 2.25 in the finite field setting, though here there are more parameters and sources of error. The key idea is that we can add vertices one at a time in about as many ways as expected once we have chosen our parameters conveniently, i.e. the argument is essentially

$$\mathcal{N}^A_{N,W,U_{k+1}}(\lambda\Gamma) \approx 2^d \alpha \lambda^{d-4} W^{2-d} \omega(v_{k+1}) \mathcal{N}^A_{N,W,U_k}(\lambda\Gamma_0),$$

where Γ has $k+1$ vertices and Γ_0 is Γ restircted to its first k vertices.

The first source of error is the vertices that we might add which come with non-primitive edges. This set is $\mathcal{B}(\underline{x})$ and controlled via Lemma 4.31.

The second and third sources of error are intertwined. There is the set of edges B_ρ for which we might have fewer than the expected number of options of vertex to add to complete a triangle. To deal with these edges, we count the total number of copies of $\lambda^2\Gamma$ which could possibly have been built in Q_N with these edges regardless of the restriction $\lambda^2\Gamma' \subset A$. This quantity is $\mathcal{N}_{I,N,W,U}(\Gamma, X)$. However, bounds for $\mathcal{N}_{I,N,W,U}(\Gamma, X)$ depend on whether the edge $X = (x_1, x_2)$ is in the exceptional set E'_W. Fortunately, the size of $E'_W \cap B_\rho$ is not too large.

Control on $|B_\rho|$ comes from Lemma 4.21, control on $\mathcal{N}_{I,N,W,U}(\Gamma, X)$ is from Lemma 4.36, and Corollary 4.24 yields a sufficient bound on $|B_\rho \cap E'_W|$.

We first set parameters. We do not lose too much by fixing a value for ρ, so we set $\rho = 1/5$.

Now we choose $w = w(\Gamma, \beta, d) \in \mathbb{N}$ such the following hold.[77]

[77] The choices are made exactly so that each error term will be at most a tenth of the main term in the induction after recalling $\beta \leq \alpha$.

1. If $f(w)$ is the $o_{w\to\infty}(1)$ term in Lemma 4.31 then for $k = 2, \ldots, n$

$$f(w) < \frac{1}{10} 2^d \beta \omega(v_k).$$

2. If c is the constant $c = c(\Gamma, d)$ appearing in Lemma 4.36, and if $g(w)$ is the $o_{w\to\infty}(1)$ term in Corollary 4.21 then for $k = 2, \dots, n$

$$c^{k-1} 2^{2d} \rho^{-2} \omega(v_{k+1})^{-2} g(w) < \frac{1}{40} 2^{kd} \beta^{k+1} \prod_{i=2}^{k+1} \omega(v_i).$$

Notice as Γ was assumed to be a distance graph in $\mathbb{Z}^d$ all of the Gelfand-Leray measures $\omega(v_i)$ are positive, as they correspond to the real solutions of the system of quadratic forms describing the triangles in Γ.

Next choose $\eta_0 = \eta_0(\Gamma, \beta, d)$ such that if c is the constant $c = c(\Gamma, d)$ appearing in Lemma 4.36, and $f(\eta)$ is the $o_{\eta\to 0}(1)$ term in Corollary 4.21 then for $k = 2, \dots, n$

$$c^{k-1} 2^{2d} \rho^{-2} \omega(v_{k+1})^{-2} f(\eta_0) < \frac{1}{40} 2^{kd} \beta^{k+1} \prod_{i=2}^{k+1} \omega(v_i).$$

Set $\eta = W^{-d/2}\eta_0$ and set $m = W \operatorname{lcm}\{1 \le n \le C_{\Gamma,d}\eta_0^{-10}\}$, where $C_{\Gamma,d} = \max_{i\in[3,n]} C_{\underline{t}(v_i),d}$ where $C_{\underline{t}(v_i),d}$ is the constant appearing in Lemma 4.20.

Finally, we choose $\lambda_0 = \lambda_0(\Gamma, \beta, d, w, \eta, L) \in \sqrt{\mathbb{N}}$ satisfying

1. $W^{3000} \le \lambda_0$.

2. $L \ll \eta^{10}\lambda_0$.

3. λ_0 is large enough that if c is the constant $c = c(\Gamma, d)$ appearing in Lemma 4.36, and $f(\lambda)$ is the $o_{\lambda\to\infty}(1)$ term in Corollary 4.21 then for $k = 2, \dots, n$

$$c^{k-1} 2^{2d} \rho^{-2} \omega(v_{k+1})^{-2} f(\lambda) < \frac{1}{40} 2^{kd} \beta^{k+1} \prod_{i=2}^{k+1} \omega(v_i).$$

   whenever $\lambda \ge \lambda_0$.

4. Given that $N \ge \lambda^4$, we have $N/\lambda^3 \ge \lambda_0$. If $g(N/\lambda^3)$ is the $o_{N/\lambda^3\to\infty}(1)$ term in Corollary 4.21, then let λ_0 large enough so that for $k = 2, \dots, n$

$$c^{k-1} 2^{2d} \rho^{-2} \omega(v_{k+1})^{-2} g(x) < \frac{1}{40} 2^{kd} \beta^{k+1} \prod_{i=2}^{k+1} \omega(v_i).$$

   whenever $x \ge \lambda_0$, c is the constant $c = c(\Gamma, d)$ appearing in Lemma 4.36.

5. If c_1 is the implied constant from Corollary 4.24 and $C_{\Gamma_{k+1},d,W}$ is the constant appearing in Lemma 4.36 for $X \in E'_W$ where Γ_{k+1} denotes Γ restricted to its first $k+1$ vertices, then for $k = 2, \dots, n$

$$\lambda_0^{-1} c_1 C_{\Gamma_{k+1},d,W} W^d < \frac{1}{10} 2^{(k+1)(d-1)} \beta^{k+1} W^{2k-1-(k+1)d} \prod_{i=2}^{k+1} \omega(v_i).$$

6. If $C_{\Gamma_k,d,W}$ is the implied constant in the lower bound of λ in Lemma 4.36 where Γ_k denotes Γ restricted to its first k vertices, then $C_{\Gamma_k,d,W} \leq \lambda_0$ for $k = 2, \dots, n$.

To complete our preparations, we note Lemma 4.38 guarantees the existence of some $U = (\underline{u}_1, \dots, \underline{u}_n)$ with $\underline{u}_i \in [W]^d$ which satisfies the necessary conditions to apply Corollary 4.21 whenever adding a vertex to Γ_0 in the coming induction.

From here, for convenience, we drop the subscripts N, W, and U as they remain fixed throughout (where we allow U to mean possibly just the first k modular restrictions of U).

Notice that there must be some equivalence class $\underline{u}_0 \in [W]^d$ such that $|A \cap (W\mathbb{Z}^d + \underline{u})| \geq |A| W^{-d}$. Translating all elements of U by a vector does not affect the properties we need from U, so we can set $\underline{u}_1 = \underline{u}_0$. Now, if Γ is the one vertex graph, we have

$$\mathcal{N}^A(\lambda\Gamma) \geq \alpha N^d W^{-d}.$$

If Γ has two vertices and one edge, then following a watered-down version of Lemma 4.20 and Corollary 4.21 gives $(1-\rho) 2^d \alpha \lambda^{d-2} W^{1-d} \omega(v_2)$ next points other than for a bad set B of size less than $2^d \rho^{-2} N^d W^{-d} \omega(v_2)^{-2} \mathcal{E}$ where $\mathcal{E}$ is bounded by the $\mathcal{E}$ in Corollary 4.21.[78] The only other error we must consider is risk that $\underline{x}_2 - \underline{x}_1$ is non-primitive. Let $\mathcal{B}$ be that set of points giving non-primitive edges. Again, a watered-down version of Lemma 4.31 demonstrates $|\mathcal{B}| = \lambda^{d-2} W^{1-d} o_{w\to\infty}(1)$. By the assumptions we have already made on w, we have $|\mathcal{B}| \leq \frac{1}{10} 2^d \alpha \lambda^{d-2} W^{1-d} \omega(v_2)$, and by the assumptions we have made on w, η, and λ_0 we have $|B| \leq \frac{1}{10} \alpha^2 N^d W^{-d} \omega(v_2)$ where $\rho = 1/5$. Noting the general upper bound for any $\underline{x}_1 \in \mathbb{Z}^d$ there are $\lesssim_d \lambda^{d-2} W^{1-d}$ points $\underline{x}_2$ on a sphere of radius λ satisfying $\underline{x}_2 \equiv \underline{u}_2 \pmod W$, we have

$$\begin{aligned}\mathcal{N}^A(\lambda\Gamma) &\geq \left(\frac{4}{5} - \frac{1}{10} - \frac{1}{10}\right) 2^{2d} \alpha^2 N^d \lambda^{d-2} W^{2-d} \omega(v_2) \\ &\geq 2^{2(d-1)} \alpha^2 N^d \lambda^{d-2} W^{2-d} \omega(v_2)\end{aligned}$$

[78] Let $\mathcal{N}^A(\underline{x}_1)$ be the number of $\underline{x}_2 \in A$ with $\underline{x}_2 \equiv \underline{u}_2\,(W)$ and $|\underline{x}_2 - \underline{x}_1|^2 = \lambda^2 t_0$. The entire process of the proof goes through with fewer sums and the result is once again bounded by η and $\|g_{\underline{u}_2}\|_{U^1_{m,L}(Q_{N/W})}$ with smaller error terms.

If Γ has k vertices and e edges with $2 \le k < n$, we assume

$$\mathcal{N}^A(\lambda\Gamma) \ge 2^{k(d-1)}\alpha^k N^d \lambda^{(k-1)d-2e} W^{e-kd} \prod_{i=2}^{k} \omega(v_i).$$

Let Γ be a distance tree of triangles on $k+1$ vertices and let Γ_0 be that graph restricted to its first k vertices. Using Corollary 4.21 with $\rho = 1/5$ we obtain

$$\mathcal{N}^A(\lambda\Gamma) \ge \left(\frac{4}{5}2^d\alpha\lambda^{d-4}W^{2-d}\omega(v_k) - \mathcal{E}_1\right)\mathcal{N}^A(\lambda\Gamma_0) - \mathcal{E}_2 - \mathcal{E}_3, \quad (4.21)$$

where

$$\mathcal{E}_1 := \sup_{\substack{\underline{x} \text{ primitive}, \\ |\underline{x}| = c_\Gamma\lambda}} |\mathcal{B}(\underline{x})|,$$

$$\mathcal{E}_2 := |B_\rho \setminus E'_W| \sup_{X \in B_\rho \setminus E'_W} \mathcal{N}_{I_{k+1}}(\lambda\Gamma, X),$$

and

$$\mathcal{E}_3 := |B_\rho \cap E'_W| \sup_{X \in B_\rho \cap E'_W} \mathcal{N}_{I_{k+1}}(\lambda\Gamma, X),$$

where I_{k+1} is the index pair of the neighbors of v_{k+1} in Γ_{k+1}.

By Lemma 4.31, and condition 1 in the definition of w

$$\mathcal{E}_1 \le \frac{1}{10}2^d\alpha\lambda^{d-4}W^{2-d}\omega(v_{k+1}). \quad (4.22)$$

For $\mathcal{E}_2$ we use the bound on $|B_\rho|$ from Corollary 4.21 and take the bound for $X \notin E'_W$ from Lemma 4.36, which is applicable because of Condition 6 on λ_0. The definitions of w (condition 2), η, and λ_0 (conditions 3 and 4) give

$$\mathcal{E}_2 \le \frac{1}{10}2^{(k+1)(d-1)}\alpha^{k+1}N^d\lambda^{kd-2(e+2)}W^{e+2-(k+1)d}\prod_{i=2}^{k+1}\omega(v_i) \quad (4.23)$$

For $\mathcal{E}_3$, we use the bound for $X \in E'_W$ from Lemma 4.36 and Corollary 4.24 to bound $|B_\rho \cap E'_W|$, the former of which is applicable by Condition 6 on λ_0. Condition 5 on λ_0 then gives

$$\mathcal{E}_3 \le \frac{1}{10}2^{(k+1)(d-1)}\alpha^{k+1}N^d\lambda^{kd-2(e+2)}W^{e+2-(k+1)d}\prod_{i=2}^{k+1}\omega(v_i). \quad (4.24)$$

Plugging equations 4.22, 4.23, and 4.24 into equation (4.21) and observing $\frac{4}{5} - \frac{1}{10} - \frac{1}{10} - \frac{1}{10} = \frac{1}{2}$ yields

$$\mathcal{N}^A(\lambda\Gamma) \geq 2^{(k+1)(d-1)}\alpha^{k+1}N^d\lambda^{kd-2(e+2)}W^{e+2-(k+1)d}\prod_{i=2}^{k+1}\omega(v_i).$$

□

4.4 Application to Sets with Positive Upper Banach Density

We now show Theorem 4.39 implies a result for sets $A \subset \mathbb{Z}^d$ which are η-uniformly distributed, which in turn implies a result for sets $A \subset \mathbb{Z}^d$ with positive upper Banach density.

Theorem 4.40. *Let $A \subset \mathbb{Z}^d$ with $\delta^*(A) > 0$ and $d > 40$. Let Γ be a distance tree of triangles in $\mathbb{Z}^d$. Then there exist $\eta > 0$ and $m \in \mathbb{N}$ such that if A is η-uniformly distributed with respect to m, then there is a $\lambda_0 = \lambda_0(\Gamma, A, d) \in \sqrt{\mathbb{N}}$ such that A contains an isometric copy of $\lambda\Gamma$ for every $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.*

Proof. By Lemma 1.10 if A is $\frac{1}{16}\eta^4$-uniformly distributed with respect to m then there exist $L = L(A, d, \eta, m)$ and $N_0 = N_0(A, d, \eta, m, L)$ such that for any $N \geq N_0$ there is a $\underline{t}_N$ such that $A_N := (A - \underline{t}_N) \cap \mathcal{C}_N$ has density $\frac{|A_N|}{|\mathcal{C}_N|} \geq \delta^*(A)$ and A_N is (η, L)-uniformly distributed with respect to m.

In particular, let w, η, and m be those given by Theorem 4.39 where we have set $\beta = \delta^*(A)$. Using the L we have chosen using Lemma 1.10, we find there exists $\lambda_0 = \lambda_0(\Gamma, \delta^*(A), d, w, \eta, L) = \lambda_0(\Gamma, A, d)$ from Theorem 4.39 such that for $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda_0 \leq \lambda \leq (N/3)^{1/4}$ there are at least

$$2^{n(d-1)}\delta^*(A)^n\left(\frac{N}{3}\right)^d\lambda^{(n-1)d-2e}W^{e-nd}\prod_{i=2}^{n}\omega(v_i)$$

isometric copies of $\lambda\Gamma$ in A_N when $N \geq N_0$. [79] For any particular $\lambda > \lambda_0$, we can choose an $N > 3\lambda^4$ and find a corresponding A_N on which this estimate holds. By letting N go to infinity, we see A in fact contains infinitely many isometric copies of $\lambda\Gamma$. □

[79] We use $N/3$ instead of N to account for the discrepancy between the cube $\mathcal{C}_N = [1, N]^d$ in Lemma 1.10 and the cube $Q_N = [-N, N]^d$ in Theorem 4.39.

Theorem 4.41. *Let $A \subset \mathbb{Z}^d$ with $\delta^*(A) > 0$ and $d > 40$. Let Γ be a distance tree of triangles in $\mathbb{Z}^d$. Then there exist $\lambda_0 = \lambda_0(A, \Gamma) \in \sqrt{\mathbb{N}}$ and $q = q(A) \in \mathbb{N}$ such that A contains an isometric copy of $q\lambda\Gamma$ for every $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.*

Proof. Let $A \subset \mathbb{Z}^d$ with $\delta^*(A) > 0$. Fix η and m from Theorem 4.40. By Lemma 1.8, there exist $q < m^{1/\eta}$ and $\underline{s} \in [q]^d$ such that the q-scaled $\underline{s}$-restricted set $A' := R_{q,\underline{s}}(A)$ satisfies $\delta^*(A') \geq \delta^*(A)$ and is η-uniformly distributed with respect to m. Theorem 4.40 then says that A' contains an isometric copy of $\lambda\Gamma$ for every $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda > \lambda_0$. Scaling A' back to A with the transformation $\underline{x} \to q\underline{x} + \underline{s}$ gives the result. $\square$

Bibliography

[Bir62] B.J. Birch. "Forms in Many Variables". In: *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 265 (1321 1962), pp. 245–263.

[Bou86] J. Bourgain. "A Szemerédi Type Theorem for Sets of Positive Density in $\mathbb{R}^k$". In: *Israel Journal of Mathematics* 54.3 (Oct. 1986), pp. 307–316.

[Bul17] K. Bulinski. "Spherical recurrence and locally isometric embeddings of trees into positive density subsets of $\mathbb{Z}^d$". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 165.2 (June 2017), pp. 267–278. ISSN: 1469-8064.

[Cas18] B. Casselman. *Hensel's Lemma*. `https://personal.math.ubc.ca/~cass/research/pdf/Hensel.pdf`. [Online; accessed 15-April-2023]. 2018.

[Cona] K. Conrad. *A Multivariable Hensel's Lemma*. `https://kconrad.math.uconn.edu/blurbs/gradnumthy/multivarhensel.pdf`. [Online; accessed 15-April-2023].

[Conb] K. Conrad. *Hensel's Lemma*. `https://kconrad.math.uconn.edu/blurbs/gradnumthy/hensel.pdf`. [Online; accessed 15-April-2023].

[CLM21] B. Cook, N. Lyall, and Á. Magyar. "Multilinear maximal operators associated to simplices". In: *Journal of the London Mathematical Society* 104.4 (2021), pp. 1491–1514.

[FKW90] H. Fürstenberg, Y. Katznelson, and B. Weiss. "Ergodic Theory and Configurations in Sets of Positive Density". In: *Algorithms and Combinatorics*. Springer Berlin Heidelberg, 1990, pp. 184–198.

[GT08] B. Green and T. Tao. "The primes contain arbitrarily long arithmetic progressions". In: *Annals of Mathematics* 167.2 (2008), pp. 481–547.

[Gre69] M.J. Greenberg. *Lectures on Forms in Many Variables*. Math Lecture Notes Series. W.A. Benjamin, 1969. ISBN: 9780805335538.

[IP19] A. Iosevich and H. Parshall. "Embedding Distance Graphs In Finite Field Vector Spaces". In: *Journal of the Korean Mathematical Society* 56.6 (2019), pp. 1515–1528.

[LN96] R. Lidl and H. Niederreiter. *Finite Fields*. 2nd ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996.

[LM20a] N. Lyall and Á. Magyar. "Distance graphs and sets of positive upper density in $\mathbb{R}^d$". In: *Analysis & PDE* 13.3 (Apr. 2020), pp. 685–700.

[LM20b] N. Lyall and Á. Magyar. "Distances and trees in dense subsets of $\mathbb{Z}^d$". In: *Israel Journal of Mathematics* 240.2 (Oct. 2020), pp. 769–790.

[LM22] N. Lyall and Á. Magyar. "Weak hypergraph regularity and applications to geometric Ramsey theory". In: *Transactions of the American Mathematical Society, Series B* 9.5 (Mar. 2022), pp. 160–207.

[Mag08] Á. Magyar. "On distance sets of large sets of integer points". In: *Israel Journal of Mathematics* 164.1 (Mar. 2008), pp. 251–263.

[Mag09] Á. Magyar. "k-Point configurations in sets of positive density of $\mathbb{Z}^n$". In: *Duke Mathematical Journal* 146.1 (Jan. 2009).

[MSW02] A. Magyar, E. M. Stein, and S. Wainger. "Discrete Analogues in Harmonic Analysis: Spherical Averages". In: *The Annals of Mathematics* 155.1 (Jan. 2002), p. 189. ISSN: 0003-486X.

[MT16] Á. Magyar and T. Titichetrakun. "Almost prime solutions to diophantine systems of high rank". In: *International Journal of Number Theory* 13.06 (Dec. 2016), pp. 1491–1514.

[Mye18] S.L. Rydin Myerson. "Quadratic forms and systems of forms in many variables". In: *Inventiones mathematicae* 213.1 (Feb. 2018), pp. 205–235. ISSN: 1432-1297.

[Par17] H. Parshall. "Point configurations over finite fields". PhD thesis. University of Georgia, 2017.

[Sal32] H. Salié. "Über die Kloostermanschen Summen $S(u, v; q)$". In: *Mathematische Zeitschrift* 34 (1932), pp. 91–109.

[Ste93] E.M. Stein. *Harmonic Analysis. Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Vol. 43. Princeton Mathematical Series. Princeton, NJ: Princeton University Press, 1993.

[Tao20] T. Tao. *Lecture Notes 8 for 247B*. `https://www.math.ucla.edu/~tao/247b.1.07w/notes8.pdf`. [Online; accessed April-2023]. 2020.

[Vin27] I.M. Vinogradov. "Démonstration analytique d'un théorème sur la distribution des parties fractionnaires d'un polynôme entier". In: *Bulletin de l'Académie des Sciences de l'URSS. VI série* 21 (4 1927), pp. 567–578.

[Wei48] A. Weil. "On some exponential sums". In: *Proc Natl Acad Sci U. S. A.* 34 (1948), pp. 204–207.

[Zie06] T. Ziegler. "Nilfactors of Rm and configurations in sets of positive upper density in $\mathbb{R}^m$". In: *Journal d'Analyse Mathématique* 99.1 (Dec. 2006), pp. 249–266. ISSN: 1565-8538.